Impacts of a Violence Prevention Program for Middle Schools

Findings After 3 Years of Implementation

NATIONAL CENTER FOR
EDUCATION EVALUATION
AND REGIONAL ASSISTANCE

Institute of Education Sciences

Impacts of a Violence Prevention Program for Middle Schools

Findings After 3 Years of Implementation

May 2011

Authors
Suyapa Silvia
Jonathan Blitstein
Jason Williams
RTI International

Chris Ringwalt
Pacific Institute for Research and Evaluation (PIRE)

Linda Dusenbury
Collaborative for Academic, Social, and Emotional Learning (CASEL)
(*formerly* Tanglewood Research, Inc.)

William Hansen
Tanglewood Research, Inc.

Project Officer Name
John Rice
Institute of Education Sciences

NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

Institute of Education Sciences

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Q. Easton
Director

National Center for Education Evaluation and Regional Assistance
Rebecca Maynard
Commissioner

May 2011

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions, recommendations, or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

To order copies of this report,
- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 1398, Jessup, MD 20794-1398.
- Call in your request toll free to 1-877-4ED-Pubs. If 877 service is not yet available in your area, call 800-872-5327 (800-USA-LEARN). Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY) should call 800-437-0833.
- Fax your request to 301-470-1244 or order online at www.edpubs.org.

This report is also available on the IES website at http://ncee.ed.gov.

Alternate Formats
Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, call the Alternate Format Center at 202-205-8113.

This page intentionally left blank.

Contents

Appendixes

This page intentionally left blank.

List of Tables

This page intentionally left blank.

List of Figures

This page intentionally left blank.

Acknowledgments

This page intentionally left blank.

Disclosure of Potential Conflicts of Interest

The research team for this evaluation consists of a prime contractor, RTI International, and two subcontractors, Pacific Institute for Research and Evaluation (PIRE) and Tanglewood Research, Inc. RTI and PIRE formed the evaluation team, while Tanglewood Research oversaw implementation of the two programs. None of these organizations or their key staff members has financial interests that could be affected by findings from the evaluation of the two school-based violence prevention programs considered in this report. No one on the Technical Working Group, convened by the research team to provide advice and guidance, has financial interests that could be affected by findings from the evaluation.

This page intentionally left blank.

Executive Summary

This is the second and final report summarizing findings from an impact evaluation of a violence prevention intervention for middle schools. This report provides findings from the second and third years of the 3-year intervention. The U.S. Department of Education (ED) contracted with RTI International and its subcontractors, Pacific Institute for Research and Evaluation (PIRE) and Tanglewood Research, Inc., to conduct an evaluation of a hybrid intervention model that combines a curriculum-based program, Responding in Peaceful and Positive Ways (RiPP [Meyer and Northup 2002a, 2002b, 2006]), and a whole-school approach, Best Behavior (Sprague and Golly 2005). The combined intervention was administered over the course of 3 successive years. Using a randomized control trial design (with entire schools randomly assigned either to receive the intervention or not), the evaluation assessed the intervention's effects on both the full sample of students as well as on students at high risk for committing violence. Tanglewood Research provided implementation oversight along with site-based liaisons and coordinated training and technical assistance for staff in intervention schools. The developers of the two programs that constitute the intervention—Prevention Opportunities and University of Oregon—provided the program materials and conducted staff training. An earlier report presented findings after 1 year of program implementation (Silvia et al. 2010).

Study Background

ED's Office of Safe and Drug-Free Schools (OSDFS) administers a variety of state and national programs under the Safe and Drug-Free Schools and Communities Act (SDFSCA, Title IV-A) that are focused on efforts to develop and maintain safe, disciplined, and drug-free schools. Drug and violence prevention activities under these programs are carried out in elementary and secondary schools, as well as in institutions of higher education. While there is now a lengthy set of school-based drug prevention curricula that have been evaluated using rigorous designs, much less evidence is available concerning effective violence prevention strategies in school settings.

The need for evidence-based violence prevention programs is particularly critical for middle schools, whose students experience the highest rate of school-based violence relative to students in other grades. The most recent data available from the National Crime Victimization Survey show that in the 2006−07 school year, 4.3 percent of students aged 12 through 18 reported that they had been victims of a crime at school (DeVoe, Bauer, and Hill 2010). Data also indicate that the rate of victimization in 2007 for nonfatal violent crimes at school for students aged 12 through 14 was 67 incidents per 1,000 students, compared with the rate for students aged 15 through 18, which was 49 incidents per 1,000 students (Dinkes, Kemp, and Baum 2009). Students were also more likely to experience a violent event in middle schools (41 per 1,000) than in elementary (26 per 1,000) or secondary (22 per 1,000) schools. Data also indicate that bullying is a significant problem. In the 2007–08 school year, 44 percent of middle schools, compared with 21 percent of primary schools and 22 percent of high schools, reported weekly or more frequent incidents of bullying (Dinkes, Kemp, and Baum 2009).

Violence prevention strategies in schools can be divided into two broad types: curriculum-based programs and whole-school (or environmental) strategies. Curriculum-based programs are implemented in a classroom setting and typically aim to improve students' social and problem-solving skills for dealing with conflict and managing violence. Whole-school (or environmental) approaches seek to influence the school environment through a variety of strategies, such as

increasing supervision of the school grounds, clarifying rules and consequences for student behavior, establishing reward systems to encourage positive behaviors, and training staff in classroom management.

While evaluations of curriculum programs have yielded statistically significant results, their effect sizes are modest. A meta-analysis of school-based violence prevention evaluations from a mix of experimental and quasi-experimental designs reported an average effect size of 0.10 (Cohen's *d*) for classroom-based social skills programs (Wilson and Lipsey 2005). RiPP (Meyer and Northup 2002a, 2002b, 2006) is an example of such a curriculum and is one of two prevention programs selected for this study. RiPP has been subjected to three discrete evaluations by the program's developers. One of these compared one classroom receiving RiPP to a nonmatched comparison classroom in the same school. In the second study, eight schools self-selected either to implement RiPP or be in the control condition. The third study used an experimental design. The two nonexperimental studies found that students who were exposed to RiPP reported significantly less physical aggression and lower levels of peer provocation than students in the comparison group but reported no significant differences in nonphysical aggression or drug use. Findings regarding self-reported delinquent behaviors were mixed (Farrell, Valois, and Meyer 2002; Farrell, Valois et al. 2003).

The single experimental evaluation of RiPP took place in 27 classes of 6th-graders in three urban middle schools. The evaluation found that after 1 year of exposure to RiPP, students reported fewer serious fight-related injuries and more participation in peer mediation compared with students in the control group but no difference on weapons-related violence or threats to teachers. School records showed fewer in-school suspensions and disciplinary violations for violent offenses for the students exposed to RiPP but no differences in out-of-school suspensions. The statistically significant outcomes were not maintained either 6 months or 1 year later (Farrell, Meyer, and White 2001). In addition, intervention classrooms in two of the study schools were subsequently randomized either to receive an additional year of RiPP or not to receive an additional year. In this case, based on school records, there were fewer violent offenses for the RiPP group 1 year after the treatment group received 2 years of RiPP and the control group received 1 year of RiPP (Farrell, Meyer et al. 2003).

A few school-based programs have sought to prevent violence by means of a whole-school approach such as the schoolwide Positive Behavior Support (PBS) approach (Sugai and Horner 1994; Sprague, Sugai, and Walker 1998) and the whole-school component of Positive Behavioral Interventions and Supports (PBIS). Three evaluations of the whole-school approach examined changes in student outcomes without using any comparison group (Barrett, Bradshaw, and Lewis-Palmer 2008; Bradshaw, Mitchell, and Leaf 2010; Metzler et al. 2001), and a fourth used an experimental design in 63 elementary schools. The investigators from the latter study reported that the program led to an improvement in school staff members' perceptions of school safety and an increase in the proportion of 3rd-graders meeting state reading assessment standards (Horner et al. 2009). In addition, two evaluations have examined student outcomes after combining characteristics of both whole-school and curriculum-based approaches. One of these evaluations studied the approach using nonmatched, comparison schools (Sprague et al. 2001), and the other evaluation used an experimental design with a sample of eight schools (Orpinas et al. 2000). Neither study reported statistically significant differences on the targeted behavioral outcomes.

The study's research questions for impacts and implementation are provided in table ES-1.

Table ES-1. Research questions

Intervention implementation
• Is delivery of the violence prevention program consistent with its design and intended implementation?
• What are the average costs of the program for participating schools and students?

Student outcomes and impacts
• Are there differences in the degree of violence and victimization in schools that implement the violence prevention program, relative to schools that do not implement it?
• What is the impact of the violence prevention program on students who are at elevated risk for violence?
• What are the outcomes of the violence prevention program for students with maximum program exposure or dosage?

The Intervention

Two research-based programs were selected through an open competition and advice from a panel of experts in the field of violence prevention: the RiPP program (Meyer and Northup 2002a, 2002b, 2006) was chosen as the curriculum-based component of the intervention, and the Best Behavior program (Sprague and Golly 2005) (a formalized version of schoolwide PBS [Sugai and Horner 1994; Sprague, Sugai, and Walker 1998]) was selected as the whole-school component. These two approaches are considered complementary in that they target both individual- and school-level change mechanisms. RiPP is similar to other social skills programs in that it aims to increase social competence and improve problem-solving skills to reduce violence. While whole-school programs such as Best Behavior typically feature a schoolwide committee to review rules and policies for student behavior, Best Behavior (and PBS, upon which it is based) also includes a reward system to reinforce appropriate behavior.

RiPP and Best Behavior were rated the highest by outside expert reviewers among a set of 16 curricular and whole-school programs submitted for consideration by program developers of middle school violence prevention programs. Criteria for program selection included the program's developmental appropriateness, overall quality of the approach and potential for reducing violence, theoretical foundation, and any outcome or process evaluation results, if available. The design of the intervention as a hybrid of the two types of programs was based on recommendations presented by a group of technical advisors in a design paper commissioned by ED prior to the current study (Bos, Weinstock, and Frankenberg 2004). The group's recommendations were the result of discussions held during a series of technical working group meetings and were informed by several literature reviews and commissioned papers on the subject. Experts concluded that a combination of the two broad types of violence prevention strategies—curricular and whole-school approaches—offered the strongest potential for impacts. This conclusion was based on the experts' judgment that the effect size was likely to be low for any one program, particularly a universal curriculum program, and that a whole-school strategy might boost that effect.

While both RiPP and Best Behavior are implemented by school staff, in this study technical assistance was made available throughout the implementation period by on-site implementation liaisons trained and hired by the implementation subcontractor, Tanglewood Research. Liaisons (e.g., former school teachers or administrators) were hired to facilitate, coach, and monitor the progress and delivery of both programs.

The Curriculum: RiPP Program

RiPP (Meyer and Northup 2002a, 2002b, 2006) is a universal, social-cognitive violence prevention program that focuses on the reduction of situational and relationship violence. The goal of the curriculum is to promote effective social-cognitive problem-solving skills; motivation and self-efficacy for using those skills; and school norms that support those attitudes and skills while reducing the appeal and perceived utility of violent behaviors and related attitudes. By targeting these attitudes and skills, the program is designed to increase social competence and thereby reduce violent behavior.

The RiPP curriculum consists of 16 lessons (each lasting 50 minutes) per year in grades 6 through 8. Each lesson builds on the previous lessons in a cumulative fashion. Similarly, each grade-level curriculum expands on the concepts taught in the previous year. In year one of program implementation, all students in the 6th through 8th grades received lessons designed for use in 6th grade (RiPP-6) because all students were required to receive the foundational lessons before the more advanced RiPP materials. In the second year of program implementation, 6th-graders received the RiPP first-year curriculum, while the 7th- and 8th-graders both received the RiPP second-year curriculum. In the third year of the project, each grade received its grade-specific curriculum. Thus, by the end of the 3-year project, the study sample (6th-graders in the first year of the project) received all 3 years of the curriculum.

The lessons in the RiPP program introduce and then reinforce the problem-solving model. The lessons comprise a variety of activities and strategies, including team building, social-cognitive problem solving, repetition and mental rehearsal, small group work, role playing, rehearsal of specific social skills for preventing violence, and didactic learning. Most lessons contain between four and six of these activities and are estimated to take between 5 minutes and 15 minutes per activity. Each activity is scripted and tied to a specific objective. Most lessons make use of the student workbook as part of the activities.

The Whole-School Intervention: Best Behavior Program

Best Behavior (Sprague and Golly 2005) provides a standardized staff development program that is designed to develop and administer effective school rules and discipline policies at both schoolwide and classroom levels to decrease school violence and antisocial behavior. The complete Best Behavior program is designed to be implemented on an approximately 3-year timeline, as individual school capacity dictates. The program is implemented by a school management team made up of teachers and administrators. Best Behavior involves intervention strategies at the school and classroom levels, including the following:

- review and refinement of school discipline policies;
- use of positive reinforcement and recognition for prosocial behaviors, both schoolwide and in individual classrooms;
- clarification and teaching of behavioral expectations for student behaviors; and
- systematic collection and review of patterns of discipline referrals to guide decisionmaking and planning.

The Best Behavior management team is expected to create a systematic approach to developing schoolwide positive behavior supports. This approach includes four broad sets of activities, the first three of which should be implemented beginning in the first year. First, the team

is to conduct a schoolwide needs assessment to identify reasonable goals. Needs assessments are to be repeated annually. Second, the team is to define rules and expectations, with general rules (e.g., be safe, be respectful, be responsible) supported with expectations for all settings within the school environment. Rules and expectations are to be taught on a regular basis by all teachers and staff. Third, the team is to develop and support a positive behavior reinforcement system in which students are to be given token rewards for obeying rules and meeting expectations. Finally, in the second year of program implementation, the team is to develop a data-based decisionmaking process for identifying and addressing the needs of high-risk students.

During annual training, Best Behavior prompts each school to develop its own specific strategies for achieving goals and objectives for each year while specifying the particular components that schools must implement. Best Behavior provides training on typical program implementation to enable each team to implement the various components specific to the needs of its individual school. For example, a sample lesson plan for teaching "be respectful" is provided, but teams are encouraged to adapt this prototype lesson to fit the language and unique cultural features of their school. As another example, all schools develop a reinforcement system, but the nature of the reinforcement system varies across schools (e.g., pizza parties vs. field trips). Therefore, while all schools implement each component of the program, the specific nature of the activities varies by school.

Best Behavior was adapted to reinforce and complement the RiPP curriculum. As part of the integration with RiPP, Best Behavior recommended that a RiPP teacher be included on the school management team. In addition, Best Behavior trainings included a review of RiPP concepts and encouraged team members to think about ways of reinforcing these concepts schoolwide. RiPP lessons also reinforced Best Behavior concepts, such as the school rules.

Evaluation Design

The study was designed as a group-randomized control experiment in 40 middle schools serving grades 6 through 8. Twenty schools were randomly assigned to receive the combined intervention, and 20 schools were randomly assigned to serve as control schools (with no intervention beyond that which schools were already implementing). The intervention was delivered schoolwide by school staff trained annually by the program developers. Best Behavior was implemented by school management teams formed at each school, while RiPP was delivered by teachers in specific classes (e.g., science, social studies). School staff also received technical assistance and were monitored throughout implementation by trained site liaisons under the guidance of Tanglewood Research.

The study was designed to estimate the impact of the RiPP and Best Behavior programs relative to what is being offered in the control schools. Most schools have ongoing violence prevention activities; therefore, the study compared the combined Best Behavior and RiPP intervention to extant school programs or the status quo. Schools that were implementing the Best Behavior or RiPP intervention (or a few other very similar programs listed on national registries as exemplary or model programs) prior to the study were excluded.

Participating Schools

Following recruitment of 13 districts and 40 schools, random assignment to condition was conducted within district, among pair-matched sets of schools based on the percentage of students receiving free or reduced-price lunches. The sites were geographically dispersed and represented a

range of district enrollment sizes. A majority of the districts were located in large urban or suburban areas with only three sites in rural districts. All participating schools were middle schools that included only grades 6 through 8. The average enrollment in these schools was 871 students and ranged between 462 students and 1,404 students. Minority students composed 65 percent of the student body, on average, and ranged between 15 percent and 100 percent. The average percentage of students receiving free or reduced-price lunches was 56 percent and ranged between 16 percent and 97 percent. There were no statistical differences between the intervention and control groups on these characteristics. Thirty-six schools (18 matched pairs) completed all 3 years of the project. For both pairs that did not complete all 3 years of the project, the control school was dropped from the study as a result of its paired intervention school ending its participation. The analytic sample for this report is comprised of data from these remaining 36 schools.

Student Sample

The analysis that estimated impacts after 3 years of the program used data from the full 8th-grade student body, including those who entered the study schools as 6th-graders or 7th-graders and remained as 8th-graders and those who were new in spring 2009 (third follow-up). Nineteen percent of the students in the third follow-up sample were not in the original sample at baseline, while 81 percent were in both samples. These data were used to answer questions about the effects of the intervention across the general student population after 3 years of program implementation. Sixty-one percent (5,854) of students enrolled in the study schools completed a survey at the third follow-up that provided data for the impact analysis after 3 years of the program. Student demographic characteristics at baseline for the 36 schools that participated for all 3 years of program delivery indicate that minority students composed 72 percent of the sample in intervention schools and 61 percent of the sample in control schools. Forty-nine percent of the students in each group were male, and 60 percent in each group lived in single-adult households. A two-tailed *t*-test, obtained from multilevel regression models using the school as the unit of analysis, indicated that none of the mean demographic characteristics was statistically different between students attending intervention schools and those attending control schools.

To address the research question regarding how the program impacts students at high risk for violent behaviors, we identified a subset of students at high risk for violent and aggressive behaviors, based on student responses to the fall 2006 baseline survey. Student self-report was considered critical for this study because many of the behaviors of concern often take place outside of the classroom and may not be reported to the school administrators. In addition, students have the best knowledge of their own behaviors and experiences. On the other hand, some students may hesitate to disclose information about their own behaviors, especially if they have concerns about the confidentiality of the data. The survey administration procedures used in this study were designed to make students comfortable with providing self-report data (see details in the "Data Collection and Outcome Measures" section below). A large study of the test-retest reliability of the Youth Risk Behavior Survey, which asks students about numerous risky behaviors, found that three of eight violence and victimization items had at least "substantial" reliability (kappa statistic ≥ 61 percent) and that five had at least "moderate" reliability (kappa statistic ≥ 41 percent) (Brener et al. 2002).

The high-risk sample was divided into two subgroups. *High-risk perpetrators* were defined as students who had self-reported perpetration of any one of eight serious acts of violence at school at

least once in the past 30 days.[1] Approximately 25 percent of the total treatment and control group sample, or 1,923 students, were assigned to the high-risk perpetrator group at baseline. *High-risk nonperpetrators* were defined as students who self-reported endorsing violence or reacting inappropriately to anger but who did not self-report any of eight serious acts of violence. Approximately 3 percent of the total treatment and control group sample, or 230 students, were assigned to the high-risk nonperpetrator group.

Unlike the remainder of the students in the full sample, the high-risk samples were to be followed longitudinally; the research team attempted to survey the students in the high-risk subgroups even if they left one of the study schools and were attending another school in the same district at each follow-up. If a student went to a different district, the student was not followed. Seventy-one percent of the students identified as high risk at baseline in the 36 schools that participated for 3 years completed a survey at the third follow-up (N = 1,357).

Teacher Sample

Secondary outcome data were collected from teachers through an annual survey conducted in spring. This survey was administered to a random sample of 24 teachers (stratified by grade) at each of the intervention and control schools participating in the study; a new sample was selected each year. Eligible teachers included all full-time classroom teachers and could include RiPP teachers who delivered the curriculum in intervention schools.[2] Ninety-eight percent (N = 854) of the sampled teachers completed a survey in the third year. Among responding teachers in intervention schools, 30 percent were RiPP teachers.

Data Collection and Outcome Measures

Data for the study's outcome measures to estimate intervention impacts were collected through student surveys in both intervention and control schools. Student surveys provided data to address the main impact research questions regarding school violence. Peer nomination measures by which data about individual students are simultaneously collected from many different peers were not considered due to the time-consuming and impractical nature of this type of collection within the context of a large multiyear effort. For this report, the baseline data collection for students occurred in fall 2006, and the follow-up to estimate impacts after 3 years occurred in spring 2009. Teacher data were collected through a survey administered to a random sample of 24 full-time teachers at each school in spring of each year to assess other program impacts besides main outcomes (school climate, victimization, feelings of safety). In addition to outcome data, the study team collected implementation data through the teacher survey, class records, annual school prevention coordinator and teacher interviews, and classroom observations. Annual interviews were conducted with the following: one person at each intervention and control school with the most knowledge about existing violence prevention activities, three randomly selected RiPP teachers at each intervention school to ask about teachers' experiences with teaching RiPP, and three randomly selected members of the school management team at each intervention school to ask about their experiences with implementing Best Behavior. Finally, observations of RiPP sessions were

---

[1] The eight serious acts of violence were: (1) pick a fight with another student; (2) hurt another by hitting, pushing, slapping, or shoving; (3) throw something at another student to hurt them; (4) take, damage, or destroy someone else's belongings; (5) try to force another student to do something they did not want to do; (6) threaten another student with a weapon such as a knife, gun, or club; (7) bring a weapon such as a knife, gun, or club to school; and (8) injure another student with a weapon such as a knife, gun, or club.

[2] Schools identified a subject or subjects, such as health or social studies, through which RiPP could be delivered to all 6th-graders through 8th-graders. Teachers in those subjects were designated as RiPP teachers and trained to deliver the curriculum.

conducted by the evaluation team in three randomly selected classrooms (one classroom per grade level) in each intervention school.

Student survey administration procedures were designed to address potential issues with reliability of self-report data on topics of a sensitive nature, such as violence and victimization. To safeguard the confidentiality of the students' data and encourage candid responses from students, the following measures were used: (1) prior to beginning the survey, students were read standard instructions to advise them how the survey should be completed and how the data would be used and safeguarded; (2) students were advised that the survey was voluntary and that they could skip any questions they did not wish to answer; (3) students responded to the survey by filling in circles for each item, so no handwritten responses by which students might be identified were required; (4) special labels with peel-away portions left only a bar code and no name on the survey booklets; (5) members of the evaluation team administered the survey, and no school staff were allowed to circulate among the students while the survey was in progress; (6) seating was arranged so that students had an empty seat between them, when possible; and (7) completed surveys were placed in large envelopes that were then sealed and taken from the school by the evaluation team.

The primary outcomes are student violence and student victimization, both measured through student surveys. For each of these two outcomes, two additional subindices were created to better understand any differences between intervention and control schools with regard to specific types of violence. A second set of indices was created to examine possible secondary effects from the intervention, beyond the primary effects. These were: (1) student safety concerns; (2) teacher safety concerns; (3) teacher victimization; and (4) student prosocial behaviors. Finally, a third set of indices was created to examine possible intermediate effects from the intervention. The theoretical model for the combined intervention predicts that changes in these areas would precede changes on the primary outcomes and included: (1) student perceptions of behavior expectations; (2) student attitudes toward violence; and (3) student self-reported coping strategies.

Analytic Approach for Estimating Program Impacts

The study team evaluated program impacts using multiple regression models that predicted each outcome's measure (e.g., violence, victimization) as a function of condition (intervention vs. control) and relevant covariates (e.g., demographic characteristics, school characteristics) using a mixed-effects regression model based on multilevel equations. Primary outcomes include self-reported counts of violent behavior and victimization occurring in the past 30 days.

The full student sample and gender subgroup analyses used a matched nested cross-sectional model (matched analysis). Under this model, students are nested in schools; schools are nested in pairs *and* in experimental condition; and pairs are crossed with experimental condition (i.e., each pair is represented at each level of condition). The covariate models for students in the full sample predicted the average response at follow-up, adjusting for the following covariates: baseline school mean of the response, school size, and individual demographic variables (gender, race/ethnicity, and number of parents in the household). For the gender subgroup analyses, the adjusted models included a gender-by-condition interaction effect.

The statistical models employed to assess program outcomes among high-risk youth are different from those employed to assess program outcomes on the general population of students. For the high-risk youth, the interest is in whether or not the RiPP and Best Behavior intervention led to individual change across time. To address this question, nested cohort models using difference-in-difference estimation were developed to assess changes on self-reported measures of

aggression and victimization among high-risk youth in intervention schools relative to changes among high-risk youth in control schools, for both subpopulations of high-risk, nonperpetrator students and high-risk, perpetrator students. These models use data collected on the same sample of students at each measurement occasion. The repeated measures models for the high-risk subsamples contained the student's treatment condition (intervention vs. control), data collection wave, wave-by-condition interaction effect, gender, race/ethnicity, number of parents in household, and school size. Estimated program impacts reflect the net difference of the within-group change from pretest to third follow-up for intervention versus controls.

To examine teacher outcomes, we employed multivariate models where teachers are nested within schools and schools are nested within matched pairs randomized to experimental condition. Hierarchical linear models account for the correlation of teachers within schools and for schools within matched pairs assigned to condition. The models predicted the average response at follow-up, adjusting for school size.

Implementation Findings

The key descriptive findings regarding implementation of the curriculum portion of the program across all years include the following:

- **In a majority of intervention schools, students were exposed to the full set of 16 RiPP lessons in each of the 3 years of implementation.** Between 61 percent and 72 percent of schools delivered all 16 lessons to all classrooms in each year of program implementation, while another 17 percent to 22 percent of schools delivered all lessons in at least three-fourths of these classrooms.

- **The curriculum was not fully delivered with fidelity.** In year three, teachers in 44 percent of the intervention schools were observed to deliver lessons with few or no deviations from the written lesson plan (e.g., adding or modifying activities or changing the activity sequence), according to classroom observations by the evaluation team. With regard to teachers delivering lessons with few or no deviations from the prescribed teaching strategies (e.g., using role plays or small group discussions), teachers met this second criterion in 56 percent of schools in the third year.

- **Interviewed RiPP teachers cited challenges with using one or more of the prescribed teaching techniques or approaches.**[3] Eighty-eight percent of teachers interviewed in year three mentioned difficulties with implementing at least one of five RiPP techniques or approaches, and 27 percent mentioned difficulties with implementing three or more of the five techniques or approaches.

- **The extent to which students were engaged with the curriculum declined by year three.** The evaluation team observed the same cohort of students receiving the curriculum over 3 successive years. These students were found to be engaged during the lesson activities, exercises, and discussions in 89 percent of the intervention schools in year one and 69 percent in year three.

---

[3] RiPP teachers were to use the following teaching techniques and approaches: make RiPP real (tie it to students' daily lives); role plays; small group work, discussion, and brainstorming; encourage self-talk by students; use *Review* and *Closure* to begin and end sessions.

The key descriptive findings regarding the implementation of the whole-school portion of the program across the 3 years include the following:

- **Principal support and commitment for the whole-school portion of the intervention was mixed in year three.** Principals at 72 percent of the intervention schools were rated as supportive in year three, according to liaisons who helped implement the program. In addition, liaisons reported that 50 percent of the principals in year three used their leadership to promote the program. In the third year of implementation, slightly more than half (56 percent) of the principals were regularly involved with the school management team.

- **By the end of the third year, the majority of intervention schools had instituted behavioral rules and rewards. In addition, a majority of teachers agreed that the rules were well defined and clear with regard to the behaviors being targeted.** By the end of the third year, 83 percent of intervention schools had developed and posted school rules in the school, 78 percent had developed and instituted a token reward system for adhering to school rules, and 78 percent had developed lesson plans and taught the school rules in classrooms. In addition, 87 percent of teachers at intervention schools agreed or strongly agreed that school rules were clearly defined. However, a smaller percentage (64 percent) of the teachers surveyed in year three agreed or strongly agreed that that it was clear what consequences would follow when school rules were broken.

- **Among the cited challenges with implementing Best Behavior were finding time to implement the program, obtaining teacher buy-in, maintaining student interest, and funding the rewards program.** School management team members interviewed in year three mentioned difficulties with finding the time to implement the program.[4] Others talked about issues with low teacher buy-in, a lack of student interest in the rewards offered through the reward system, and continued problems with funding the reward system.

- **By design, no control school implemented RiPP or Best Behavior during the 3 years of the study.** However, there were various violence prevention activities already in place in the participating schools. Between eight schools and nine schools in the intervention group and between six schools and seven schools in the control group administered classroom-based education other than RiPP across the 3 years. The types of programs implemented included gang resistance programs, character education programs, and individual presentations that were not part of a curriculum (most often, speakers, a video, or a lesson) focused on specific topics such as bullying, harassment, and dating violence.

---

[4] Key program practices are as follows: (1) the school management team meets regularly; (2) rules of behavior are defined and adopted; (3) rules are posted throughout the school; (4) plans are developed for teaching rules; (5) rules and expectations are taught; (6) a reinforcement system is developed; (7) the reinforcement system is implemented school-wide; (8) discipline data are collected and reviewed; (9) focus is given to the needs of high-risk youth; and (10) a schoolwide needs assessment is conducted periodically.

Impact Findings

These are the main findings regarding intervention impacts after 3 years.

- **There were no statistically significant differences between intervention and control schools on self-reported student violence or victimization (figure ES-1).** On average, 8th-graders in the intervention schools reported engaging in 2.84 violent acts at school in the past 30 days, compared with 8th-graders in control schools, who reported engaging in 2.69 violent acts at school in the past 30 days. On average, 8th-graders in intervention schools reported being victimized 4.14 times in the past 30 days, while 8th-graders in control schools reported being victimized 4.16 times in the past 30 days.

- **There were no statistically significant program impacts on violence or victimization for students who were at risk for engaging in violence but who had not previously done so (figure ES-2).** On average, 8th-graders in the intervention schools who were categorized as being at a high risk for violence but who had not self-reported any of eight serious acts of violence ever at baseline (nonperpetrators) reported at the third-year follow-up that they had engaged in 3.30 (change from baseline of 1.28) violent acts at school in the past 30 days. This is compared with high-risk, nonperpetrator 8th-graders in the control schools who reported at the third follow-up that they had engaged in 3.12 (change from baseline of 1.35) violent acts at school in the past 30 days. In addition, high-risk, nonperpetrator 8th-graders in the intervention schools reported being victimized an average of 4.30 (change from baseline of 3.60) times at school in the past 30 days. This is compared with high-risk, nonperpetrator 8th-graders in the control schools who reported being victimized an average of 3.79 (change from baseline of 3.41) times at school in the past 30 days.

- **There were no statistically significant program impacts on violence or victimization for high-risk students who had previously engaged in violence (figure ES-3).** On average, high-risk, perpetrator 8th-graders in the intervention schools reported at the third follow-up that they had engaged in 3.95 (change from baseline of 5.36) violent acts at school in the past 30 days. This is compared with high-risk, perpetrator 8th-graders in the control schools who reported at the third follow-up that they had engaged in 3.90 (change from baseline of 4.92) violent acts at school in the past 30 days. Those in intervention schools reported being victimized 4.90 times at school in the past 30 days, on average (change from baseline of 6.29), while those in control schools reported an average victimization rate of 5.05 (change from baseline of 6.10).

- **There were no statistically significant impacts on either secondary or intermediate outcomes.** In addition, after 3 years of exposure to the RiPP and Best Behavior intervention, student measures for secondary outcomes—including student safety concerns, teacher victimization and safety concerns, and student prosocial behaviors—did not differ between students in intervention schools and students in control schools. Also, there were no statistically significant differences on intermediate outcomes—that is, where the program logic model predicts change would be observed before it would be observed on the outcome measures. These include student self-reported coping strategies, student perceptions of behavior expectations, and student attitudes toward violence.

Figure ES-1. Main program impacts on self-reported violence and victimization, after 3 years of program delivery

Incidents, past 30 days (Adjusted)

■Intervention   ■Control
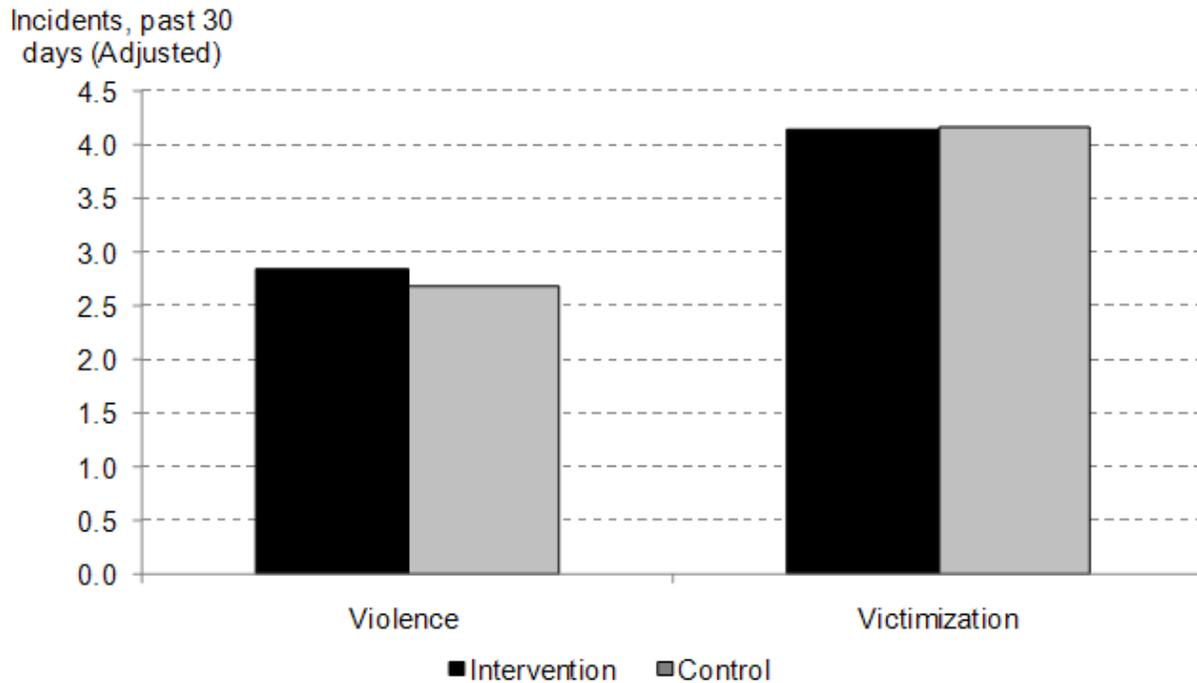
NOTE: *N* = 2,784 students clustered in 18 intervention schools and 3,070 students clustered in 18 control schools. Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, treatment condition (intervention vs. control), race/ethnicity, sex, number of parents in household, and school size. Program impacts were estimated as a model-adjusted event rate ratio (ERR) for intervention versus controls at follow-up, with 95 percent confidence limits. There are no statistically significant program impacts.
SOURCE: Student survey, fall 2006 (baseline) and spring 2009 (third follow-up).

Figure ES-2. Main program impacts on self-reported violence and victimization, after 3 years of program delivery: High-risk, nonperpetrator subgroup (Via repeated measures)
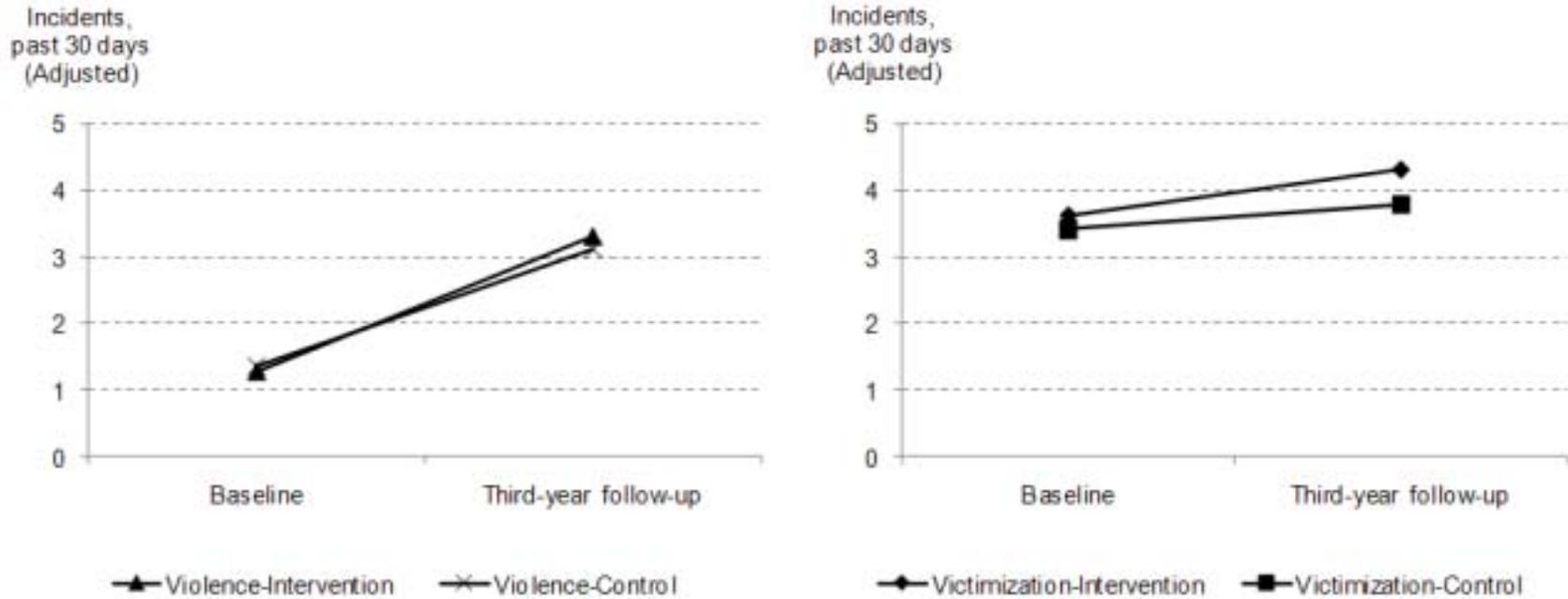
NOTE: *N* = 70 students clustered in 18 intervention schools and 74 students clustered in 18 control schools. Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included gender, race/ethnicity, number of parents in household, and school size. Program impacts (with 95 percent confidence limits) were estimated via difference-in-difference models comparing change across time in the intervention versus control group. There are no statistically significant program impacts.

SOURCE: Student surveys limited to a high-risk subgroup of 6th-graders surveyed in fall 2006 and followed up in spring 2009.

Figure ES-3. Main program impacts on self-reported violence and victimization, after 3 years of program delivery: High-risk, perpetrator subgroup (Via repeated measures)
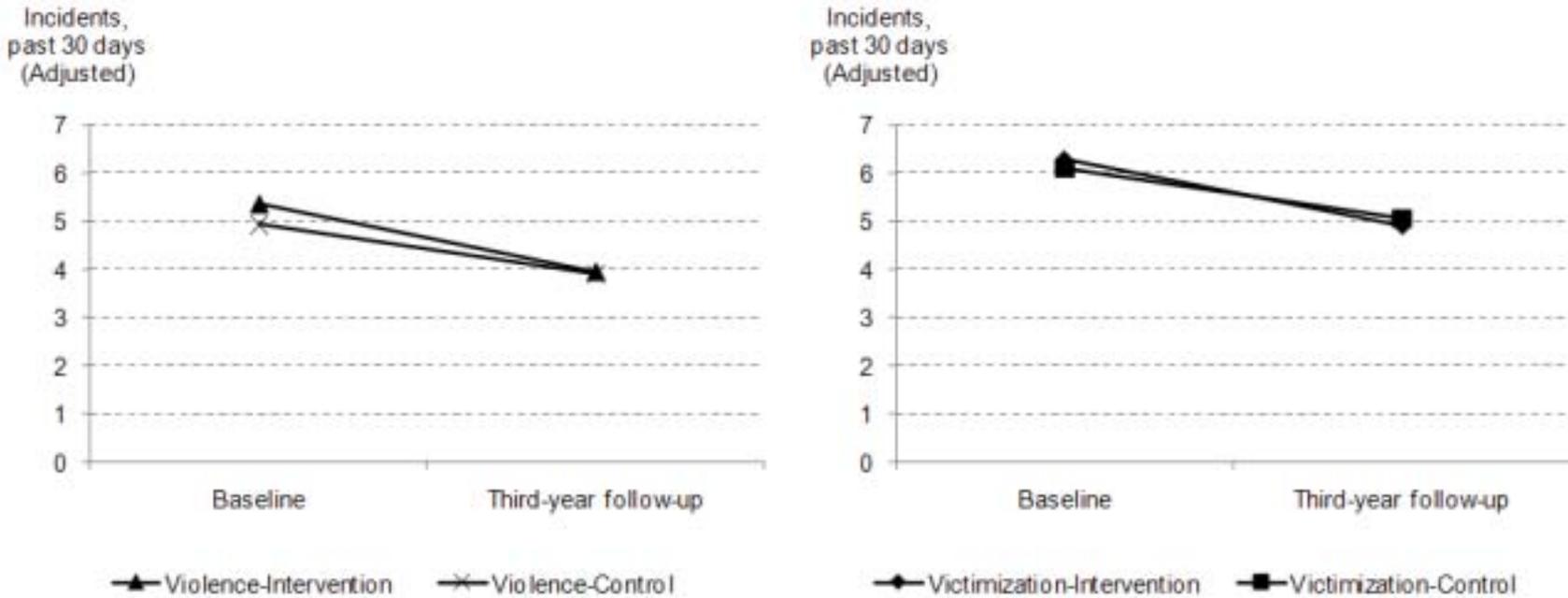
NOTE: *N* = 552 students clustered in 18 intervention schools and 661 students clustered in 18 control schools. Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included gender, race/ethnicity, number of parents in household, and school size. Program impacts (with 95 percent confidence limits) were estimated via difference-in-difference models comparing change across time in the intervention versus control group. There are no statistically significant program impacts.

SOURCE: Student surveys limited to a high-risk subgroup of 6th-graders surveyed in fall 2006 and followed up in spring 2009.

Chapter 1.
Overview of the Study

RTI International and its subcontractor, Pacific Institute for Research and Evaluation (PIRE), conducted an impact evaluation of a middle school–based violence prevention program under contract from the National Center for Education Evaluation and Regional Assistance of the U.S. Department of Education. This study examined the effects of a hybrid intervention model that combined curriculum-based and whole-school approaches. Using a group-randomized trial design, the evaluation assessed program effects on student violence and victimization. The program was administered over the course of 3 successive school years by Tanglewood Research, Inc., which assisted in the process by which the programs under study were selected and also coordinated training and technical assistance for staff in intervention schools. The developers of the two programs that made up the intervention—Prevention Opportunities and University of Oregon—provided the program materials and conducted staff training; both were subcontractors on this study. This report provides findings from the second and third years of the three-year program. An earlier report presented findings after 1 year of program implementation (Silvia et al. 2010).

This chapter provides an overview of the federal programs in the U.S. Department of Education that address school violence and the nature and extent of the problem of violence in the nation's middle schools. It also summarizes the research literature on the effectiveness of violence prevention programs, particularly in middle schools. The chapter then describes the interventions that were implemented and evaluated in the present study and presents the key research questions that were addressed. Finally, the chapter summarizes the findings after the first year of program implementation. Later chapters report findings from the second and third school years.

1.1 Federal Legislation Supporting Prevention Efforts

The Drug-Free Schools and Communities Act (DFSCA) enacted by Congress as Subtitle B of Title IV of the Anti-Drug Abuse Act of 1986, has provided the largest source of federal support for alcohol and other drug use prevention efforts for school-aged youth. In response to increasing concerns over school safety, Congress reauthorized the DFSCA as the Safe and Drug-Free Schools and Communities Act (SDFSCA) of 1994, Title IV of the Elementary and Secondary Act of 1965, and added violence prevention as a supported activity under the legislation. Concerns for an orderly and safe school environment were also reflected in the Goals 2000: Educate America Act, P.L. 103-227, 108 Stat. 125–209 (1994), which included as one of its eight national education goals, a specific goal for safe, disciplined, and alcohol- and drug-free schools.

The U.S. Department of Education's Office of Safe and Drug-Free Schools (OSDFS) administers a variety of state and national programs under SDFSCA that are focused on efforts to develop and maintain safe, disciplined, and drug-free schools. Drug and violence prevention activities under these programs are carried out in elementary and secondary schools, as well as institutions of higher education. A study conducted in 2000 found that 90 percent of districts that receive SDFSCA funding implemented curricula that target youth violence (Hantman and Crosse 2000).

The No Child Left Behind Act (NCLB) of 2001, which reauthorized the SDFSCA under Title IV, Part A, further specified that SDFSCA funding be used to develop, implement, and evaluate scientifically valid, evidence-based programs to prevent violence and the illegal use of drugs

among students (Section 4121). NCLB legislation also included the "Principles of Effectiveness," which was originally developed by OSDFS in 1998 (U.S. Department of Education 1998). The third of these principles directs schools receiving SDFSCA funds to administer prevention programs that have yielded empirical evidence of reductions in drug use or violent behavior. With support from the National Institutes of Health—and particularly from the National Institute on Drug Abuse—a lengthy set of school-based drug prevention curricula now have been evaluated using randomized trials (National Institute on Drug Abuse 2003). However, much less rigorous evidence is available concerning effective violence prevention strategies in school settings, particularly for middle schools, whose students (as described in the next section) experience the highest rate of school-based violence relative to students in other grades.

1.2 The Problem of Violence in the Nation's Middle Schools

Violent behavior may be defined as the threat, attempted use, or actual use of physical force that results in physical or nonphysical harm. Violence, as well as other forms of nonphysical aggression, is prevalent in a wide variety of social and physical contexts, including the nation's schools. As defined in *Violence in U.S. Public Schools* (Miller 2003), school violence encompasses events that occur on or near school property, on school buses, and at school-sponsored events. While school violence may be lethal, it far more commonly involves fighting, bullying, physical attacks, and verbal threats made by students or their parents against teachers or other students.

The most recent data available from the National Crime Victimization Survey show that in the 2006–07 school year, 4.3 percent of students aged 12 through 18 reported that they had been victims of any crime at school (DeVoe, Bauer, and Hill 2010). Three percent reported being victims of theft, 1.3 percent reported a simple assault, and 0.4 percent reported a serious violent victimization, which includes rape, sexual assault, robbery, and aggravated assault.

Students are most likely to be subjected to violent behaviors in middle school grades. Based on youth self-reports, the rate of victimization in 2007 for nonfatal violent crimes at school for students aged 12 through 14 was 67 incidents per 1,000 students, compared with the rate for students aged 15 through 18, which was 49 incidents per 1,000 students (Dinkes, Kemp, and Baum 2009). In addition, data recorded by schools indicate that students were more likely to experience a violent event in middle schools (41 per 1,000) than in elementary schools (26 per 1,000) or secondary schools (22 per 1,000). Data recorded by schools also indicate that bullying is a significant problem. In the 2007–08 school year, 44 percent of middle schools, compared to 21 percent of primary schools and 22 percent of high schools, reported weekly or more frequent incidents of bullying (Dinkes, Kemp, and Baum 2009).

1.3 Review of the Literature Concerning the Effectiveness of School-Based Violence Prevention Strategies in Middle Schools

Violence prevention strategies in schools can be divided into two broad types: curriculum-based programs and whole-school (or environmental) strategies. Curriculum-based programs are implemented in a classroom setting and typically aim to improve students' social and problem-solving skills for dealing with conflict and managing aggression. Whole-school (or environmental) approaches seek to influence the school environment through a variety of strategies, such as increasing supervision of the school grounds, clarifying rules and consequences for student behavior, establishing reward systems to encourage positive behaviors, and training staff in classroom management.

1.3.1 Curricular Approaches

Violence prevention curricula that target students in middle school settings include Second Step (Committee for Children 1990), Too Good for Violence (Mendez Foundation 1995), and Responding in Peaceful and Positive Ways (RiPP [Meyer and Northup 2002a, 2002b, 2006]), the latter of which is one of two prevention programs selected for this study. While evaluations of curriculum programs have yielded statistically significant results, their effect sizes are modest. A meta-analysis of school-based violence prevention evaluations from a mix of experimental and quasi-experimental designs reported an average effect size of 0.10 (Cohen's *d*) for classroom-based social skills programs (Wilson and Lipsey 2005).

RiPP has been subjected to three discrete evaluations by the program's developers. One of these compared one classroom receiving RiPP to a nonmatched comparison classroom in the same school. In the second study, eight schools self-selected to either implement RiPP or be in the control condition. The third study used an experimental design. The nonexperimental studies found that students who were exposed to RiPP reported significantly less physical aggression and lower levels of peer provocation than students in the comparison group, but reported no significant differences in nonphysical aggression or drug use. Findings regarding self-reported delinquent behaviors were mixed (Farrell, Valois, and Meyer 2002; Farrell, Valois et al. 2003).

The single experimental evaluation of RiPP took place in 27 classes of 6th-graders in three urban middle schools. The evaluation found that after 1 year of exposure to RiPP, students reported fewer serious fight-related injuries and more participation in peer mediation compared with students in the control group but no difference on weapons-related violence or threats to teachers. School records showed fewer in-school suspensions and disciplinary violations for violent offenses for the students exposed to RiPP but no differences in out-of-school suspensions. The statistically significant outcomes were not maintained either 6 months or 1 year later (Farrell, Meyer, and White 2001). In addition, intervention classrooms in two of the study schools were subsequently randomized either to receive an additional year of RiPP or not to receive an additional year. In this case, based on school records, there were fewer violent offenses for the RiPP group 1 year after the treatment group received 2 years of RiPP, and the control group received 1 year of RiPP (Farrell, Meyer et al. 2003).

1.3.2 Whole-School Approaches

A few school-based programs have sought to prevent violence by means of a whole-school approach. An example is the schoolwide Positive Behavior Support (PBS) approach (Sugai and Horner 1994; Sprague, Sugai, and Walker 1998). PBS is a schoolwide intervention that teaches school staff to recognize, monitor, and reward good school behavior and to provide consistent sanctions for rule violations. Best Behavior (Sprague and Golly 2005), which constitutes the second of the prevention programs selected for this evaluation, is a formalized, written expression of the PBS model. While Best Behavior has yet to be evaluated as such, it is based on the components and strategies that have been tested as part of PBS.

There have been two evaluations of student outcomes for PBS. One evaluation found reductions in students' self-reported aggressive social behaviors, increases in perceptions of school safety, and decreased disciplinary referrals over time in schools using PBS. However, the analyses used to obtain these findings either did not use data from the comparison group or did not conduct statistical tests for group differences (Metzler et al. 2001). There has been one published randomized study of PBS that took place in 63 elementary schools. The investigators reported that the program

led to two statistically significant outcomes: an improvement in school staff members' perceptions of school safety and an increase in the proportion of third graders meeting state reading assessment standards (Horner et al. 2009).

There have been two evaluations of the whole-school component of a related program titled Positive Behavioral Interventions and Supports (PBIS); however, neither evaluation used a rigorous design to examine student outcomes. For example, one evaluations reported reductions in student suspensions and office discipline referrals over time in 21 elementary schools using PBIS, but did not compare these outcomes from the PBIS schools to those from a control or comparison group (Bradshaw, Mitchell, and Leaf 2010). In a second study, investigators used a convenience sample of 182 schools and found that office discipline referrals were significantly lower in these schools compared to the national average (Barrett, Bradshaw, and Lewis-Palmer 2008).

1.3.3 Combining Curricular and Whole-School Approaches

Several violence prevention approaches have combined characteristics of both whole-school and curriculum-based programs. One study evaluated PBS in combination with the Second Step curriculum using a nonmatched, comparison-group design with 15 schools. While intervention schools showed larger reductions in office discipline referrals and greater knowledge of social skills compared to nonmatched comparisons schools, tests of statistical significance were not conducted (Sprague et al. 2001). Another evaluation of a combination approach used an experimental design with eight schools. The program did not produce any detectable impacts (Orpinas et al. 2000).

1.4   Rationale for the Study

As summarized in earlier sections, the rationale for this study is based on several important considerations. First, there is a lack of rigorous evidence available concerning effective violence prevention strategies in school settings. Such is not the case for drug prevention programs, many of which have been subjected to randomized trials and other rigorous designs. Second, the need for information about evidence-based violence prevention programs is particularly critical for middle schools. As described earlier, students in middle schools experience more violence and are victimized at higher rates than students in other grades. Further, bullying behavior is much more prevalent in middle schools than in primary schools or high schools. This study aims to contribute to the knowledge about evidence-based violence prevention programs for middle schools.

Prior to the current study, ED commissioned a design paper (Bos, Weinstock, and Frankenberg 2004), and convened technical working groups of experts in the fields of school-based violence prevention, research design, and statistics to consider designs and a focus for the study. A recommendation was made to focus on a combination of curricular and schoolwide environmental approaches for violence prevention. The rationale was that the effect sizes of any one program were likely to be low, especially for a single curriculum program; therefore a combination of strategies would likely boost the effect size. Based on those recommendations, a hybrid intervention that combined a classroom-based curriculum and a whole-school approach was implemented for this study.

The intervention programs for this evaluation were selected through an open competition. RiPP and Best Behavior represented the most promising choices for curricular and whole-school approaches, respectively, among a set of 16 programs submitted for consideration by program developers of middle school violence prevention programs. Criteria for program selection included

the program's developmental appropriateness, overall quality of the approach and potential for reducing violence, theoretical foundation, and any outcome or process evaluation results, if available.

1.5   The Intervention

The school-based violence prevention intervention that is the focus of this evaluation combines two discrete approaches: a curriculum-based model to facilitate students' social competency, problem-solving, and self-control skills; and a whole-school model that targets school practices and policies through the systemic reorganization and modification of school management strategies, disciplinary policies, and enforcement procedures. These two approaches are considered complementary, in that they target both individual- and school-level change mechanisms; together, they offer the opportunity for synergistic benefits. Through an open competition and with the advice of a panel of experts in the field of violence prevention, the RiPP program (Meyer and Northup 2002a, 2002b, 2006) was chosen as the curriculum-based component of the intervention, and the Best Behavior program (Sprague and Golly 2005) was selected as the whole-school component. As universal prevention programs, RiPP and Best Behavior (as implemented in this study[5]) are not designed as interventions to be used with students who are already exhibiting serious violent behavior in school. Nonetheless, the study also examined outcomes among a subset of students at elevated risk for violence to investigate the potential impact of the program on these students.

Both of these programs were implemented as a combined intervention for 3 consecutive years in a number of middle schools. The program developers provided training for teachers and staff each year, and the implementation subcontractor, Tanglewood Research, provided teachers and staff with ongoing support for the duration of the evaluation. Figure 1 presents a theory-of-action model for the combined intervention.

The study also gathered information on all violence prevention strategies, other than the treatment, that are potentially related to the study's outcomes of interest, from both intervention and comparison schools. The study compared the combined Best Behavior and RiPP intervention to the extant school programs or the status quo. Most schools have ongoing violence prevention activities; however, schools that were already implementing RiPP or Best Behavior (or a few other very similar programs) were excluded from consideration at the recruitment stage.

---

[5] An additional component of Best Behavior, not included in this study, provides support systems for individual students.

Figure 1.   Theory-of-action model for the intervention
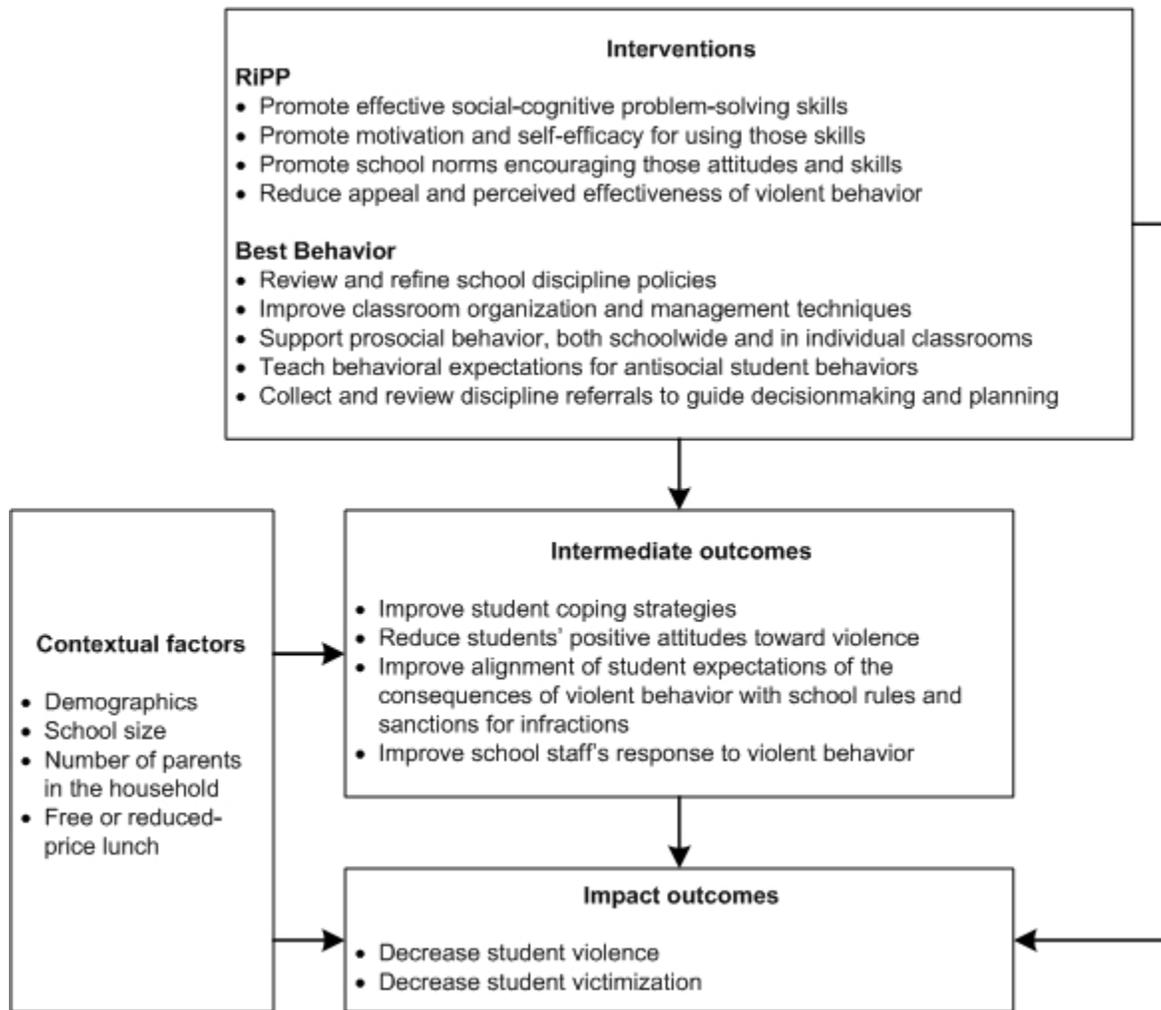
1.5.1 The Curriculum: RiPP Program

RiPP (Meyer and Northup 2002a, 2002b, 2006) is a universal, social-cognitive violence prevention program that focuses on the reduction of situational violence (precipitated by sociological factors such as poverty, overcrowding, alcohol and drug use, community norms, and access to weapons) and relationship violence (arising from interpersonal relationships). The goal of the curriculum is to promote effective social-cognitive problem-solving skills; motivation and self-efficacy for using those skills; and school norms that support those attitudes and skills while reducing the appeal and perceived utility of violent behaviors and related attitudes. By targeting these attitudes and skills, the program is designed to increase social competence and thereby reduce violent behavior.

For this study, the RiPP curriculum comprises 16 grade-specific lessons, each 50 minutes in length, delivered over the course of a school year to 6th- through 8th-grade students. Through repeated exposure to this problem-solving model, increased awareness of nonviolent options, and opportunities for reflection and practice, participants learn how to choose the prosocial strategy in any given situation that is most likely to provide desired short- and long-term outcomes.

In this study, the RiPP curriculum is delivered by classroom teachers in specific subjects selected by each participating school for implementation of RiPP (e.g., science, health), as opposed to the outside professionals for whom the curriculum was originally developed. The classes are purposely selected to be ones in which all students participate; thus, all students are exposed to the intervention. In addition to implementing the RiPP curriculum, teachers are trained to use social modeling and classroom management techniques to promote prosocial behaviors and reduce the incidence of violent behaviors. Classroom management techniques alter the classroom context and shift the reward contingencies (i.e., consequences) associated with inappropriate behavior. Prosocial modeling involves the demonstration of nonaggressive methods for conflict resolution. This suggests that teacher behavior in the classroom setting is an important experiential component of the intervention program. A more detailed description of RiPP appears in chapter 3.

1.5.2 The Whole-School Intervention: Best Behavior Program

The purpose of the Best Behavior program (Sprague and Golly 2005) is to improve discipline in schools, and the program is designed to be implemented by a school management team made up of teachers and administrators. The whole-school component involves intervention strategies at the school and classroom levels, including the following:

- review and refinement of school discipline policies;

- instruction on classroom organization and management techniques;

- use of positive reinforcement and recognition for prosocial behaviors, both schoolwide and in individual classrooms;

- clarification and teaching of behavioral expectations for student behaviors; and

- systematic collection and review of patterns of discipline referrals to guide decisionmaking and planning.

The Best Behavior developers provided extensive training to all school management team members at each school prior to the implementation and in follow-up years and were available throughout the study for assistance with resolving problems as they arose. Per the program design, only the members of the school management team were trained in Best Behavior. Through meetings and other dissemination efforts, the school management teams involved and brought awareness of the program to the entire faculty and staff as well as to the student body. Ongoing support throughout implementation was provided by site liaisons. Best Behavior was implemented schoolwide for 3 years under the guidance of a school management team that was developed for this purpose. A more detailed description of Best Behavior appears in chapter 3.

1.6 Key Study Research Questions

The study's primary research question is as follows: "Does the degree of violent behaviors differ in schools where the violence prevention intervention is implemented, relative to schools that do not receive the intervention?" The study includes both a process evaluation, which is designed to assess the fidelity with which the program was implemented and to provide contextual information, and an impact evaluation, which will assess improvement in program outcomes. Specific research questions, presented in table 1, guided each of these components.

Table 1.   Research questions

Intervention implementation
• Is delivery of the violence prevention program consistent with its design and intended implementation?
• What are the average costs of the program for participating schools and students?

Student outcomes and impacts
• Are there differences in the degree of violence and victimization in schools that implement the violence prevention program, relative to schools that do not implement it?
• What is the impact of the violence prevention program on students who are at elevated risk for violence?
• What are the outcomes of the violence prevention program for students with maximum program exposure or dosage?

1.7   Findings After 1 Year of Implementation

Findings from year one were published in an earlier report (Silvia et al. 2010). There were no statistically significant differences between the students in intervention and control schools on self-reported violence and victimization, both overall and for specific types of violence and victimization. In addition, there were no statistically significant impacts on self-reported violence and victimization between high-risk youth at intervention schools and high-risk youth at control schools. Finally, there were no statistically significant impacts on either secondary outcomes, such as prosocial behaviors, or on intermediate outcomes, such as student perceptions of behavior expectations or student attitudes toward violence.

The curriculum portion of the program was implemented largely as intended during the first year. For example, it was delivered in its entirety to all assigned classrooms in 70 percent of the schools. In 65 percent of the schools, curriculum teachers adhered to the scripts, and in 55 percent of the schools teachers adhered to prescribed teaching strategies. However, the implementation findings were mixed for the whole-school program. For example, the majority of intervention schools instituted the key steps of the whole-school program: 75 percent of schools had posted school rules in the school, 75 percent had set up a reward system for adhering to school rules, and 50 percent had the school rules taught in the classroom. In addition, 84 percent of full-time 6th-grade through 8th-grade teachers surveyed in intervention schools indicated that the rules developed from the program were well defined. However, the teams of school administrators and teachers who were charged with instituting the whole-school program met an average of five times during the year and not monthly, as stipulated by the program. Also, school principals (whose role is key in sustaining the program) were found to be supportive in half of the intervention schools.

1.8   Organization of This Report

This report is organized to provide a detailed description of the methods and data collected for the study and the findings after the second and third years of implementation of the combined intervention. Chapter 2 presents the study design, describes the sample selection, and outlines the analytic approaches. Chapter 3 describes the implementation of the program, including progress made through the second and third years, training and technical assistance, challenges, treatment contrast, and program costs. Chapter 4 presents findings from the impact analyses for both intermediate and impact outcomes. Finally, chapter 5 presents findings based on additional, exploratory analyses.

Chapter 2.
Study Design, Sample Selection, Measures, and Analytic Approach

This study examines the implementation and impacts of a violence prevention intervention in middle schools that was delivered over 3 years to students in grades 6 through 8. The key question is whether or not there are observed differences in levels of school violence and victimization in schools that delivered the intervention relative to control schools that did not receive the intervention. This chapter presents the study design, methods for selecting schools and respondents, and sample characteristics and response rates for the third year; it also provides an overview of data collection and the approaches to the data analysis. Corresponding sample characteristics and response rates for the second year appear in appendix F; there were no substantive differences between the second and third year sample demographics or response rates.

2.1 Study Design

Designed as a group-randomized control experiment, the study was conducted in 40 middle schools serving only grades 6 through 8. Twenty schools were randomly assigned to receive a combined curriculum plus whole-school violence prevention intervention, and 20 schools were randomly assigned to serve as control schools (with no intervention beyond that which schools were already implementing). The intervention was delivered schoolwide over 3 years by school staff trained by the program developers. School staff also received technical assistance and were monitored throughout implementation by trained site liaisons under the guidance of the implementation subcontractor.

A plan for RiPP delivery was developed in consultation with the developer, based on the logical progression of the program over the 3 years. In year one of program implementation, all students in the 6th through 8th grades received lessons designed for use in 6th grade (RiPP-6) because all students were required to receive the foundational lessons before the more advanced Responding in Peaceful and Positive Ways (RiPP) program (Meyer and Northup 2002a, 2002b, 2006) materials. In the second year of program implementation, 6th-graders received the RiPP first-year curriculum, while the 7th- and 8th-graders both received the RiPP second-year curriculum. In the third year of the project, each grade received its grade-specific curriculum. Thus, by the end of the 3-year project, the study sample (6th-graders in the first year of the project) received all 3 years of the curriculum.

Student outcome data were collected from the 2006 6th-grade cohort at each intervention and control school prior to program delivery in fall 2006 and again at the end of each of the 3 years of program delivery, in spring 2007, spring 2008, and spring 2009. Teacher outcome data were collected at the end of each school year. In addition to outcome data, the study collected implementation data annually from all intervention schools and documented violence prevention activities, policies, and programs that were offered in addition to the intervention, in both intervention and control schools.

The impact analysis is based on a pretest/posttest control group design; this design provides protection from a number of potential sources of bias and is generally recommended over designs that rely solely on information collected at the end of an intervention (Shadish, Cook, and Campbell

2002). The study evaluates program impact using multiple regression models that predict each outcome measure (e.g., violence, victimization) as a function of condition (intervention vs. control) and of relevant covariates (e.g., baseline measures, demographic characteristics). The inclusion of covariates related to the outcomes but unrelated to program exposure improves the precision of the test of the program effect by reducing unexplained variation. Because students are nested within schools and schools are nested within condition, the study estimates these effects using multilevel regression equations and software. A second impact analysis focuses on a subset of students at high risk for violence, based on student responses to the fall 2006 baseline survey. These analyses address a second key research question concerning the impact of the intervention on high-risk students. Additional, exploratory analyses examine program effects for gender subgroups and differences in outcomes based on program exposure and level of program implementation.

2.2   Description of the Process for Selecting Programs

Two research-based programs were selected through an open competition and advice from a panel of experts in the field of violence prevention: the RiPP program (Meyer and Northup 2002a, 2002b, 2006) was chosen as the curriculum-based component of the intervention, and the Best Behavior program (Sprague and Golly 2005) (a formalized version of Positive Behavior Support [PBS] [Sugai and Horner 1994; Sprague, Sugai, and Walker 1998]) was selected as the whole-school component. In this section, we describe the process for identifying and selecting these two programs.

2.2.1 Solicitation Process

To identify the most promising approaches to violence prevention for middle schools, the research team developed a solicitation that invited program developers to submit information about their approaches for review. The solicitation was disseminated in various ways, including mass e-mails to all known program developers of violence prevention programs for middle schools (approximately 46 programs), an advertisement in the October 6, 2004, issue of *Education Week* (http://www.edweek.org/ew/index.html), and distribution of a flyer at two national conferences on violence prevention ("Preventing Violence and Related Health-Risking Social Behaviors in Adolescents: An NIH State of the Science Conference," held October 13–15, 2004; "Persistently Safe Schools: The National Conference of the Hamilton Fish Institute on School and Community Violence," held October 27–29, 2004). The solicitation included a description of the goals and merits of this project and defined the criteria for inclusion. The solicitation described the types of middle school violence prevention programs sought, including: (1) classroom instruction models, (2) models of systemic change and reorganization, and (3) hybrid programs that integrate both classroom instruction and systemic change. The solicitation asked developers to submit with their proposals a description of the program's model and content, data on program usage, and any evidence of effectiveness. In addition, the solicitation requested information on the program's requirements for training, technical assistance, and program management and the developer's capacity to assist the contractor's implementation team with these activities. If the program was not a middle school approach or had not yet been tested in a middle school environment, the developer was asked to provide a description of how the program would be modified to fit that setting. A total of 16 programs were submitted for consideration.

2.2.2 Review Process

The primary criteria for program selection included having an approach that had strong potential—based on theory, supportive empirical data, or preferably both—to reduce violence and conflict, including aggression, bullying, and other forms of inappropriate behavior among middle school students. Approaches that had this potential were expected to have the following characteristics:

- an understandable and well-defined theory or logic model describing how actions taken by a school's faculty and staff are expected to affect students' competencies and motivations, the school environment, or both and how these changes would translate into changes in violent behavior;

- outcomes from past feasibility and pilot tests with explanations of how the approach is appropriate or will be modified to be appropriate for middle schools;

- manualization (in at least a preliminary form) or sufficiently clear documentation of the approach so that manuals can be produced that will allow the approach to be replicated systematically in many sites; and

- a plan for training and technical support.

Three independent experts in school-based violence prevention served as reviewers in the program selection process. Raters were provided with criteria for evaluating such qualities as developmental appropriateness, cultural appropriateness, ease of implementation, the extent of integration within and across settings, and potential for engaging students. Review criteria included the theoretical or conceptual foundation of the program and process and outcome evaluations, if any. Further, raters judged the overall quality of the approach and, in light of existing theory and evidence, the potential of the approach to reduce violence. Reviewers were trained in a 3-hour conference call, with documents available for each reviewer to examine and refer to as training proceeded. The reviewers reported no conflicts of interest with regard to any of the programs that were submitted.

The review of programs involved a two-stage process: (1) initial review of all documents provided by programs to determine a short list of likely qualified programs, and (2) a final review of programs based on documentation plus information gathered from applicants in response to questions generated from the initial review. Each program was analyzed by two reviewers. Paired reviews were discussed to ensure there was as much consistency among reviewers as possible about each program.

Recommendations presented by a group of technical advisors in a design paper commissioned by ED prior to the current study (Bos, Weinstock, and Frankenberg 2004), indicated that a combination of curricular and whole-school approaches offered the strongest potential for impacts (see section 1.4 for additional details). Unfortunately, none of the highest-rank programs could be described as a hybrid model. We selected two of the highest-rated programs—one that was a curriculum-based strategy, the other a whole-school approach—for this project. We consulted with the developers of each of the two programs to ensure their willingness to have each of their programs be part of the same evaluation and to be evaluated collectively, not separately. Developers also agreed to alter slightly the language used in the program so that the violence prevention terminology was consistent between the two programs.

2.3   Criteria and Process for District and School Selection and Random Assignment

This section describes the process by which districts and schools were identified and recruited into the study and, in the case of schools, subsequently assigned randomly to either the intervention or control group. Districts and schools were selected on the basis of a number of inclusion and exclusion criteria that help to ensure a sample of schools that would remain viable throughout the anticipated experimental period. This approach creates a purposive, rather than a truly random, sample of schools and is common in group-randomized trials.

2.3.1 District and School Selection

The sampling frame of schools was constructed using publicly available information from the Common Core of Data (CCD) (http://nces.ed.gov/ccd/districtsearch/index.asp) regarding poverty (based on the percentage of students in each school eligible for free and reduced-price lunches), enrollment size, locale, and race/ethnicity. To address the analytical objectives of this study and manage the many operational challenges, the sampling frame was limited to regular (public, noncharter, and nonmagnet) schools including at least grades 6 through 8 with a 6th-grade population of at least 250 students. The research team required at least 250 students in the 6th grade in each school to obtain a sufficiently large high-risk subsample. School districts with fewer than three eligible schools were excluded because they yielded insufficient matched pairs of schools to include in the study. Moreover, the sampling frame was designed to oversample urban, high-poverty, and high-minority schools, as these factors are strongly correlated with school violence (DuRant et al. 1994; DuRant et al. 1997; DuRant et al. 1999). Because data on actual rates of school violence are not generally available for individual schools, the aforementioned characteristics were used as proxies for identifying schools with elevated levels of violence and aggression.

Recruitment and randomization proceeded as illustrated in figure 2. Working from the sampling frame, districts were first screened to confirm their eligibility and to identify any approval processes required to contact sampled schools. Districts were deemed ineligible to participate if there was a current or planned district mandate for all middle schools to implement curriculum-based or whole-school violence prevention programs that were similar to RiPP or Best Behavior. A total of 248 school districts were found eligible to be recruited for the study. Of these, 107 districts refused to participate, 94 districts did not respond with a decision, and 34 districts were not contacted because the study's target number of schools had been reached. Thirteen districts agreed to allow the study team to contact their schools.

On receipt of district approval to contact schools, schools were mailed study materials and called to invite their participation in the study. In a majority of cases, the decision to participate in the study was left up to the individual schools. Information regarding 6th-grade enrollment and the percentage of students eligible for free and reduced-price lunches was requested to verify school-level eligibility. The study excluded schools that were currently implementing or were planning to implement programs similar to RiPP or Best Behavior or schools for which future redistricting plans (e.g., changing feeder patterns) would negatively impact the 6th-grade enrollment levels for the ensuing academic year. The study also excluded schools that could not accommodate 16 lessons, each lasting 50 minutes, for all 6th- through 8th-graders and that could not identify placement for the RiPP curriculum in the normal academic day. During recruitment, it was made clear that schools would be required to abide by random assignment. If selected for the control group, schools had to

agree to refrain from implementing similar violence prevention programs listed on any of the national registries for model violence prevention programs during the 3 years of the study.

Figure 2.   Flow diagram detailing recruitment and randomization
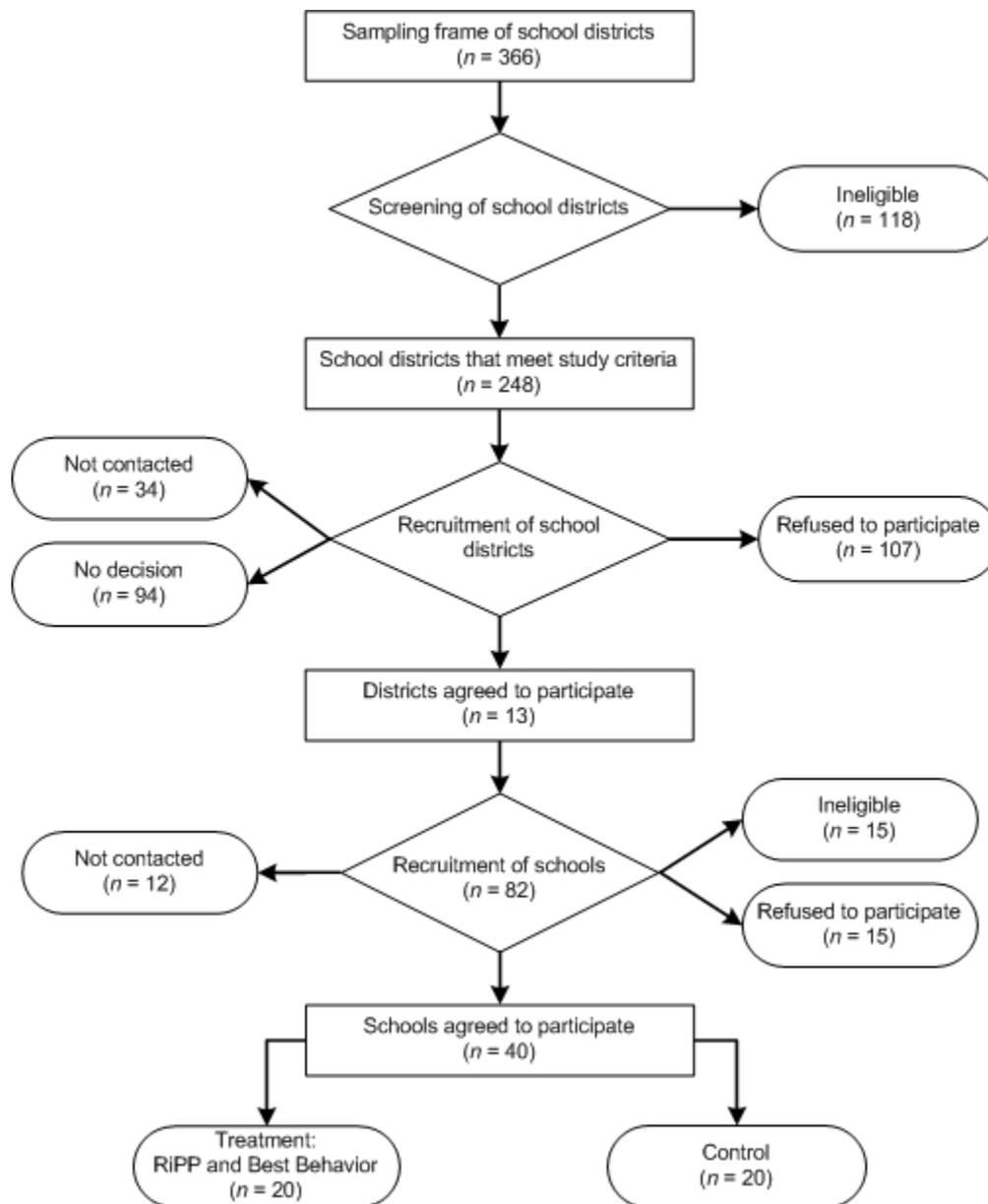
SOURCE: Data are from recruitment records maintained by the research team.

Of the 82 middle schools that met initial eligibility criteria for the sampling frame, an additional 15 schools were found ineligible during the recruitment stage, 15 schools refused to participate, and 12 schools were not contacted because they exceeded the number of schools needed to form matched pairs from the same district.

All schools in a district that met the criteria and agreed to participate were accepted into the study, if at least two such schools could be identified in the same district. Prior to randomly assigning schools to the intervention or control group, the research team required the school

principals and district superintendent to sign a memorandum of understanding that described the study activities and detailed the duties and responsibilities of the schools, districts, and RTI International.

2.3.2 Random Assignment of Schools to Intervention

A critical feature of the impact evaluation is the random assignment of schools either to the intervention group, which receives the RiPP and Best Behavior intervention, or to the control group, which continues the schools' ongoing violence prevention efforts. Random assignment to condition was conducted within district, among pair-matched sets of schools. In group-randomized trials, matching is used to help ensure the distribution of potentially confounding influences and to increase the precision of the test of the program impact (Murray 1998). Because matching occurred prior to the collection of baseline information, schools were matched on the proportion of students who receive free or reduced-price lunches; this variable is often employed as a proxy for low socioeconomic condition and tends to be correlated with violence in the community, particularly in cities and in combination with social disorganization (e.g., economic and social flux, high turnover of residents, and a high percentage of single-parent families) (U.S. Department of Health and Human Services 2001).

In districts with exactly two schools recruited, those schools automatically formed a pair, and one school was randomly assigned to the intervention condition. In districts with more than two schools (even number), schools were matched using the percentage of free or reduced-price lunches, and then one school in the pair was randomly assigned to the intervention condition and the other to the control condition. In districts that included an odd number of recruited schools, a final pair was formed across two districts.

2.3.3 Characteristics of the Districts and Schools in the Study

In all, 13 districts were recruited into the study, each yielding between one and three pairs of schools. As shown in table 2, the sites were geographically dispersed and represented a range of district enrollment sizes. Districts were located predominantly in the South; only four of the schools were located in districts in the Northeast, and eight schools were located in districts in the West/Midwest. A majority of the districts were located in large urban or suburban areas, with only three sites in rural districts. Some attrition of schools occurred over the 3 years; one pair of schools did not participate in the second and third years, while a second pair of schools from a different district did not participate in the third year. In each case, the control school was dropped from the study as a result of its paired intervention school ending their participation.

All participating schools were middle schools that included only grades 6 through 8. As shown in table 3, the average enrollment in the 36 schools that participated for 3 years is 853 students and ranges between 462 students and 1,404 students. Minority students comprise 67 percent of the student body, on average, and range between 15 percent and 100 percent. The average percentage of students receiving free or reduced-price lunches is 53 percent and ranges between 16 percent and 89 percent.

Table 2.   Characteristics of participating districts

Region	District location	District enrollment range	Number of participating schools
Total			40
West/Midwest	Urban	20,000 and fewer to greater than 100,000	8
Northeast	Urban, Suburban	20,000 and fewer	4
South	Urban, Suburban, Rural	20,000 and fewer to greater than 100,000	28

SOURCE: *Search for Public School Districts: School Year 2005–06*, Common Core of Data (CCD), U.S. Department of Education, Institute of Education Sciences, National Center for Educational Statistics. Retrieved July 22, 2008, from http://nces.ed.gov/ccd/districtsearch/index.asp.

Table 3.   Baseline demographic characteristics of schools with 3 years of participation

Characteristic	All schools	Intervention	Control	Difference	*p*-value
Enrollment (36 schools)					
Mean	853	824	881	−57	0.42
Range	462–1,404	462–1,404	634–1,209		
Race/ethnicity minority (%)					
Mean percentage	66.91	70.08	63.73	6.36	0.52
Range	14.69–100.00	20.26–100.00	14.69–100.00		
Students eligible for free or reduced-price lunches (%)					
Mean percentage	52.88	53.83	51.94	1.89	0.77
Range	16.44–89.43	27.48–83.55	16.44–89.43		

NOTE: A two-tailed *t*-test was used to test differences between intervention schools and control schools. Statistical significance is indicated by * if the *p*-value is less than or equal to 0.05.

SOURCE: *Search for Public School Districts: School Year 2005–06*, Common Core of Data (CCD), U.S. Department of Education, Institute of Education Sciences, National Center for Educational Statistics. Retrieved April 29, 2009, from http://nces.ed.gov/ccd/districtsearch/index.asp.

Intervention and control schools were compared for baseline differences. As shown in table 3, there are no statistical differences between the groups on these characteristics, based on the 36 schools that remained in the study in the third year. As described above, schools were pairwise matched using the percentage of free or reduced-price lunches prior to random assignment.

2.4   Defining and Recruiting Students and Teachers Into the Sample

2.4.1 Defining the Student Full Sample

The school year 2006–07 student sample included all 6th-graders in both intervention and control schools. In school year 2007–08, the student sample included all 7th-grade students in the schools; this included students promoted from the schools' 6th-grade as well as students who transferred in to the study schools. In school year 2008–09, we followed the same approach.

Approximately 55 percent of the students surveyed in year three (2008–09) were present in the selected treatment and control schools throughout the study period and constitute the full exposure subset of students. The issue of full exposure (i.e., students present in study schools for the full study period 2006–09) is explored in section 5.3.

The research team obtained written parental consent from 6th-grade students at baseline in fall 2006. At each of three follow-up points in spring 2007, spring 2008, and spring 2009, the study attempted to survey all 6th-, 7th-, and 8th-grade students, respectively, who had prior consent and remained in the schools. In addition, the study sought parental consent for new students who had moved into study schools in the survey grades.

The student survey sample was designed to respond to the study's main impact questions while minimizing the burden of data collection at each school. As shown in table 4, impact analyses for the second and third years are based on the 7th- and 8th-grade classes, respectively, and include students who remained in the study schools and also new students enrolled in each of those years whose parents provided consent for the survey. These data will be used to answer questions about the effects of the intervention across the general student population after 2 years and 3 years of program implementation.

Table 4.   Impact data collection points for students

Data collection time point	Grade level		
	6th grade	7th grade	8th grade
Baseline (Year 1: fall 2006)	Census of 6th grade (N = 7,601 in 40 schools)		
First follow-up (Year 1: spring 2007)	6th-grade students remaining in school and new students (N = 7,351 in 40 schools)		
Second follow-up (Year 2: spring 2008)		7th-grade students remaining in school and new students (N = 6,500 in 38 schools)	
Third follow-up (Year 3: spring 2009)			8th-grade students remaining in school and new students (N = 5,928 in 36 schools)

NOTE: Data were collected at both intervention and control schools.
SOURCE: Data collection records maintained by the research team.

While all students in grades 6 through 8 participated in the RiPP and Best Behavior programs each year, the 6th-grade cohort in 2006 was exposed to the intervention for all 3 years. The data collection schedule for the student survey thus follows the grade-progression pattern of the intervention program, from 6th grade to 8th grade. This approach is consistent with the primary aim of examining the impact of the RiPP and Best Behavior program on the school population as a whole. It examines the impacts of the RiPP and Best Behavior program among students who attend

intervention schools, as compared with students in a matched group of schools that did not include the program.

2.4.2 Defining the High-Risk Student Subgroup

To address the research question regarding how the program impacts students at high risk for violent behaviors, we identified a subset of students at high risk for violent and aggressive behaviors, based on student responses to the fall 2006 baseline survey. The goal at baseline was to identify an average of 54 high-risk students per school and establish a cohort of 2,160 high-risk youth that would be tracked across the intervention period. Allowing for nonparticipation and loss to follow-up, this would provide a cohort of approximately 1,440 high-risk students at the 3-year follow-up data collection. Unlike the remainder of the students in the full sample, the research team attempted to survey the students in the high-risk subgroup even if they left one of the study schools and were attending another school in the same district at each subsequent data collection point. This is because the high-risk cohort was tracked longitudinally, while the full sample comprised cross-sectional samples each year.

The science supporting the identification of youth at heightened risk for violent perpetration is not well developed. In consultation with experts in the field of youth violence prevention, the study team used data from the baseline survey and created an index to identify youth who may be at elevated risk for exhibiting violent behavior.[6] In examining risk, the study considered both overt behaviors and self-reported attitudes and intentions regarding violence, using two complementary strategies to identify high-risk students. The two strategies are described in the next sections.

Identifying High-Risk Youth Based on Overt Behaviors

To identify students who had committed the most serious violent and aggressive acts and were, therefore, likely to perpetrate again in the future, we examined students' responses to eight perpetration items in the baseline survey. The study employed item response theory (IRT [Nunnally and Bernstein 1994]) to select the 8 items from among the 14 items in the baseline survey that asked about student perpetration of violent and aggressive acts (see appendix B for additional details). IRT was used here because it permitted us to examine the relationship of each component item from the violence index to an underlying, unobserved latent factor of perpetration. In this manner, each perpetration item could be examined for its strength of relationship to the perpetration factor as well as its severity, relative to other items. The analysis identified the following eight items (listed in descending order of IRT severity ranking) that mark significant violent behavior:

- injuring another student with a weapon such as a knife, gun, or club;

- bringing a weapon such as a knife, gun, or club to school;

- threatening another student with a weapon such as a knife, gun, or club;

- trying to force another student to do something that he or she does not want to do;

- taking, damaging, or destroying on purpose something that belongs to another student;

---

[6] Identification of the high-risk sample was required prior to implementing the intervention (within 1 month or 2 months of the beginning of the year). It was not feasible to use nominations for high-risk youth by teachers, peers, or school administrators because the 6th-grade sample used to identify the high-risk youth was new to the school in the fall of 2006 and would not have been well known to others.

- throwing something at another student to hurt him or her;

- actually hurting another student by hitting, pushing, slapping, or shoving; and

- instigating a fight with another student.

Students who reported committing any of these acts one or more times in the past 30 days were placed in the high-risk subgroup. Approximately 28 percent of the total sample of treatment and comparison students, or 2,135 students, had committed at least one of these behaviors in the past 30 days at baseline and were classified as high risk by this criterion.

Identifying High-Risk Youth Based on Attitudes and Intentions

To identify additional high-risk students based on attitudes and intentions toward violence, the study examined measures that indicate a cognitive predisposition toward violent behavior. As these constructs serve as theoretical mediators of program effects, they are relevant for identifying risk. Latent class analysis (LCA [McCutcheon 1987; Collins and Lanza 2009]) was employed to identify the additional high-risk students.

An initial factor analysis of the self-report student survey items identified three unidimensional constructs: attitudes endorsing violence (eight items), positive reactions to anger (nine items), and negative reactions to anger (three items). The study team next entered scale scores for each construct into an LCA, a latent variable technique similar to factor analysis. However, whereas factor analysis attempts to find common dimensions that group items, LCA attempts to find profiles of responses that group respondents. The groups formed in LCA are based on similarities in response patterns, so that item relationships are explained by class membership.

LCA produced a five-class solution, with one class (approximately 10 percent of the total sample, or 722 students) demonstrating a clear potential for future violence based on attitudes and intentions. The profile of responses in this class showed elevated positive attitudes toward violence and negative (self-reported) coping strategies, together with lower positive (self-reported) coping strategies.[7] Among the 722 students identified as high risk based on profiles of attitudes and intentions, 65 percent (466 students) were previously classified as high risk based on overt behaviors. Consequently, the remaining 256 students representing an additional 3 percent of the total sample were classified as high risk based on profiles of attitudes and intentions only.

Total High-Risk Subgroup

Combining the students identified based on behaviors and those indentified based on profiles of attitudes and intentions generated a subsample of 2,391 high-risk youth (31 percent of the total sample of treatment and control students surveyed at baseline). Based on the results of a priori sample size estimates, the study team anticipated the need for a minimum cohort of 36 youth per school (for a total sample size of 1,440) for purposes of statistical power. Tracking the full high-risk cohort would not be possible based on budget considerations. Accordingly, the researchers randomly selected 60 high-risk youth from schools with 71 or more such students identified. This resulted in a final sample of 2,153 high-risk students (28 percent of the total sample), of which 1,923

---

[7] To measure *coping strategies*, students were asked to report the likelihood of reacting in various ways when angered. Exploratory factor analysis indicated two independent dimensions of the coping measure: one represented nine positive or nonviolent reactions (e.g., walk away or ignore the situation/person, try to talk it out with the other person), and the other represented three negative or violent reactions (e.g., yell at the person, break something, hit or threaten to hurt the person). Accordingly, two composite variables were formed to represent positive or negative reactions.

students (25 percent of the total sample) were classified as high-risk based on perpetration and 230 students (3 percent of the total sample) were classified as high-risk based on profiles of attitudes and intentions toward violence but not perpetration. This approach provides a high-risk cohort that met the target minimum detectable effects (MDEs) for the high-risk sample (see appendix E).

Impact analyses examine the effects of the program for the total high-risk group as well as the effects for the high-risk youth who reported acts of violence in the past 30 days at baseline (labeled high-risk perpetrators) and the high-risk youth who exhibited risk based on their attitudes and intentions toward violence but who reported no acts of violence at baseline (labeled high-risk nonperpetrators). Separate analyses were conducted for perpetrators and nonperpetrators to explore whether or not the intervention affects these two groups of youth differently.

The impact on the subset of students identified as high risk for violence is assessed using an intent-to-treat model wherein they are included in analyses, whether or not they receive the intervention. This component answers questions about the effects of the intervention on individual change in students at high risk for violence, specifically, students who reported having engaged in aggressive or violent acts at baseline and students who had not already engaged in such acts but whose attitudes and intentions toward violence placed them at high risk for violence. The same subset of high-risk students identified in fall 2006 was assessed at the end of the first, second, and third years, based on data collected in spring 2007, spring 2008, and spring 2009. This report presents results after the second and third implementation years.

2.4.3 Selection of the Teacher Sample

Secondary outcome data (school climate, victimization, feelings of safety) were collected from teachers through an annual survey conducted in spring. This survey was administered to a random sample of 24 teachers for the 6th grade through 8th grade (stratified by grade) at each of the 40 middle schools participating in the study; a new sample from among the entire full-time teacher faculty was selected each year. Eligible teachers included all full-time classroom teachers for the 6th grade through 8th grade. In intervention schools, eligible teachers included, but were not limited to, RiPP teachers who delivered the curriculum in select classes such as science or health.

2.5 Response Rates

2.5.1 Student Participation

For the student survey, response rates were largely determined by three factors: parental consent, student absence at survey administration, and student willingness or ability to complete the questionnaire. Active parental consent was required, and the study team worked closely with each school to ensure that parents who did not return the consent form were sent reminder notices with replacement forms, as needed, to maximize the parental consent rates. Regarding student absence, the study tracked attendance at survey administration for all students for whom the study had parental consent, keeping a list of absentees. The research team scheduled one makeup session in each school for students who missed the regular administration. During survey administration, the study team used two-part identification labels with peel-away portions that left only a bar-coded label with no identifiers on the survey; this was intended to bolster students' belief in the confidentiality of the survey. The surveys were administered by trained RTI staff and field staff members.

The student response rates for the follow-up at the end of the third year of program implementation are shown in table 5 (second-year response rates are shown in table F-4, appendix F). In spring 2009, 68 percent of the 9,358 students in the 8th grade received written parental consent for participation in the study. A survey was completed by 90 percent of students with consent, for an overall response rate of 61 percent. Response rates across schools ranged from 40 percent to 89 percent at the third follow-up, with a median of 62 percent. Four schools from the same district exhibited four of the five lowest response rates (below 50 percent).

Table 5.   Student response rates, total sample—Year three

	Total 8th-grade enrollment (2008–09)	Percent consented of total 8th-grade enrollment	Percent surveyed of consented	Percent surveyed of total 8th-grade enrollment
Third follow-up (36 schools)				
Total sample	9,358	68	90	61
School median	248	69	92	62
School range	130–396	42–96	70–100	40–89

SOURCE: Data collection records maintained by the research team.

As shown in table 6, a two-tailed *t*-test indicated that the mean response rates between intervention (63 percent) and control groups (62 percent) were not statistically different at the third follow-up. Student mobility in and out of study schools was examined by measuring the rate of students who exited or entered the sample at the third follow-up. Across all schools, there was a 23 percent increase in new students for whom consent was attempted at the third follow-up; students for whom consent was subsequently obtained and who completed a survey were included in the analysis sample. Also, there was a 28 percent loss of sample students from the prior year due to students transferring out of study schools. These rates were not statistically different for intervention and control groups. Response and mobility rates by group for year two are shown in table F-6, appendix F.

Table 6.   Student response and mobility rates, by group—Year three

	Total	Intervention	Control	Difference	*t*-test	*p*-value
(Mean rate)						
Response rate at third follow-up[1]	62.47	63.33	61.61	1.72	0.41	0.6864
Rate of exiting students[2]	28.37	32.39	24.35	8.03	1.39	0.1725
Rate of entering students[3]	22.90	23.19	22.61	0.58	0.11	0.9122

[1] The response rate is defined as the percent of students surveyed of enrolled in year three.

[2] The rate of exiting students is the percentage of students who left the analysis sample at the third follow-up.

[3] The rate of entering students is the percentage of students who entered the analysis sample at the third follow-up. An attempt was made to obtain consent from all new students. Students from whom consent was subsequently obtained and who completed a survey were included in the third follow-up sample.

NOTE: A two-tailed test was used to test differences between intervention schools and control schools. Statistical significance is indicated by * if the *p*-value is less than or equal to 0.05.

SOURCE: Data collection records maintained by the research team.

Follow-up among the high-risk sample was handled differently. The impact analysis of high-risk youth employs a longitudinal design. Under this design, the sample of high-risk youth identified in the first year was treated as a cohort, and the study employed rigorous tracking methods to ensure that all youth in the high-risk cohort who remained in the school districts involved in the study provided data at each data collection period. Even so, some attrition was unavoidable. At the third follow-up, 71 percent of the 1,915 high-risk students in the 11 remaining districts completed a survey, including students who transferred to other schools in the participating districts, while 3 percent did not complete a survey due to refusal or repeated absences. Twenty-six percent were not surveyed because they moved out of the school district and could no longer be included; these data were missing in the impact analyses at the third follow-up. Among intervention and control schools, the average school response rate for high-risk students at the third follow-up was 70 percent for intervention schools and 72 percent for control schools; rates were not statistically different across the two groups ($t = -0.66, p = 0.51$).

2.5.2 Teacher Participation

For the teacher survey, the study monitored data collection returns and encouraged participation by contacting respondents who did not return their completed questionnaire. Ninety-eight percent of the 864 teachers sampled in spring 2009 responded to the survey, for a total of 845 teachers. The teacher sample comprised approximately equal numbers of teachers in intervention ($N = 424$) and control schools ($N = 421$). The average school-level response rate among intervention and control schools was 99 percent in both groups ($t = 0.23, p = 0.813$). Among responding teachers in intervention schools, 30 percent had taught the RiPP curriculum during the year, while the majority (70 percent) were non-RiPP teachers (i.e., full-time teachers for the 6th grade through 8th grade who did not teach RiPP).

2.6 Data Collection

We conducted a number of different data collection activities to address the study's research objectives. Data for the study's outcome measures to estimate intervention impacts were collected through student surveys and teacher surveys (which were also used to collect implementation data). Both were gathered in intervention and control schools. Program implementation data sources included interviews with violence prevention coordinators, violence prevention staff, and school management teams; teacher data on RiPP implementation; and observations of RiPP classroom activities. These data were used to assess fidelity of implementation, level of student exposure to the program, training of implementers, implementation of interventions other than the test intervention, and other issues relevant to implementation. Table 7 summarizes the data collection activities, including the schedule. Each of the data sources is described in more detail in the sections that follow.

Table 7.   Data collection activities

Instrument	Use of information	Condition	First year		Second year		Third year	
			Fall 06	Spring 07	Fall 07	Spring 08	Fall 08	Spring 09
Student survey	Outcome	Intervention and control schools	•	•		•		•
Teacher survey	Outcome and implementation	Intervention and control schools			•	•		•
Violence prevention coordinator interview guide	Implementation	Intervention and control schools			•	•		•
Violence prevention staff interview guide	Implementation	Intervention schools			•	•		•
School management team interview guide	Implementation	Intervention schools			•	•		•
Curriculum implementation records	Implementation	Intervention schools			•	•		•
Classroom observations	Implementation	Intervention schools			•	•		•
Monthly Implementation Progress Reports (Liaisons)	Implementation	Intervention schools	•	•	•	•	•	•
Year-End Implementation Reports (Liaisons)	Implementation	Intervention schools			•	•		•

2.6.1 Student Surveys

Student surveys were conducted in both intervention and control schools. Peer nomination measures by which data about individual students are simultaneously collected from many different peers were not considered due to the time-consuming and impractical nature of this type of collection within the context of a large multiyear effort. However, the survey included various scales that ask about students' relationships and interactions with peers, including: victimization by or aggression toward peers in school; attitudes toward interpersonal peer violence; and prosocial behaviors extended to or received by peers. Other questions asked about students' social competency skills, aggressive or disruptive conduct, and perceptions of safety at school and related avoidance behaviors. The survey took no more than one class period (45–50 minutes) to complete, on average. The student survey was administered to 6th-graders in the fall and spring semesters of the 2006–07 school year, to 7th-graders in the spring semester of the 2007–08 school year, and to 8th-graders in the 2008–09 school year. At each follow-up, the survey included students in the high-risk subsample if they remained at the same school; if they had left the school they attended during baseline, they were still surveyed, if possible, as long as the school they now attended was in the same district.

A copy of the student survey is included in appendix A. The student survey was constructed using scales from existing surveys, as shown in appendix C, table C-1. These scales were chosen to

measure the specific intermediate and impact outcomes reflected in the logic model (figure 1, chapter 1) under several domains:

- violence,

- victimization,

- safety concerns,

- behavioral expectations,

- prosocial behaviors,

- self-reported coping strategies, and

- attitudes toward violence.

Prior to its first use at baseline, the study team pilot tested the survey with nine students to determine that they were able to complete it in 45 minutes or less and that the items were not overly difficult and did not make respondents uncomfortable.

2.6.2 Teacher Surveys

The teacher surveys were completed by teachers in both intervention and comparison schools. Questions asked about teachers' perceptions of the level of disruptive behaviors in class, perceptions of school climate, experience with victimization, and feelings of safety in school. Teachers completed the survey during spring in 2007, 2008, and 2009; a new random sample of teachers was selected in each year. The survey is self-administered and, on average, took no more than 30 minutes to complete.

The teacher survey was constructed using scales from other surveys that addressed specific outcome and implementation measures of interest for this study. The sources for the outcome measures are shown in appendix C, table C-1. A copy of the teacher survey is included in appendix A.

2.6.3 Violence Prevention Coordinator Interview

The violence prevention coordinator interview provided information on existing violence prevention strategies in both intervention and comparison schools. The interview was used to gather detailed information regarding all of the schools' violence prevention efforts, including curricula, whole-school and policy-based programs, and environmental efforts (e.g., metal detectors, security cameras). One person identified by the school administrator as the individual with the greatest institutional knowledge of violence prevention efforts was interviewed.

The interview also asked about staff professional development for violence prevention and, in intervention schools only, the coordinator's view of program implementation, including how well RiPP and Best Behavior work together. The study team conducted the interviews once each year during winter in 2007, 2008, and 2009 so that respondents could describe activities that were ongoing. Additional details about the interview protocol are provided in appendix D.

2.6.4 Violence Prevention Staff Interviews

The violence prevention staff interviews provided information on implementation of RiPP. These interviews were conducted with teachers and other violence prevention staff implementing RiPP in intervention schools. Sixty-one percent of the 47 teachers interviewed in year two and

69 percent of the 53 teachers interviewed in year three had taught RiPP for at least 2 years; 69 percent of teachers in year two and 77 percent in year three had completed teaching over one-half of the lessons when the interviews took place.

The semistructured interview asked about RiPP implementation experiences, challenges, and impressions; the time taken by teachers to prepare, deliver, and follow up the intervention; fidelity and adaptation; training and technical assistance; and how RiPP and Best Behavior fit together. Data were collected once each school year during the 2006–07, 2007–08, and 2008–09 school years. The research team conducted these interviews as part of the same site visit during which the violence prevention coordinator interview was conducted. Data were analyzed using the qualitative software program ATLAS.ti 6.0 (Muhr and Friese 2009) and provided context to the implementation fidelity findings about teachers' experiences with RiPP. Additional details about the interview protocol and coding of the data are provided in appendix D.

2.6.5 School Management Team Interviews

The school management team interviews provided information on the implementation of the Best Behavior whole-school approach to violence prevention. The interview asked about the school management team member's background and violence prevention roles; how staff in each school have implemented Best Behavior; the time taken by each member for activities; fidelity of implementation; staff impressions on training and technical assistance received; and staff impressions on how well Best Behavior and RiPP have fit together and with other programs used at the school, particularly for programs related to violence prevention. Data were collected once each school year during the 2006–07, 2007–08, and 2008–09 school years as part of the same site visit during which other interviews were conducted. Data were analyzed using the qualitative software program ATLAS.ti 6.0 (Muhr and Friese 2009) and were used to provide more in-depth information about the school management teams' experiences implementing Best Behavior. Additional details about the interview protocol and coding of the data are provided in appendix D.

2.6.6 Curriculum Implementation Records

Assessment of RiPP implementation records provided information on implementation of the RiPP curriculum and was collected only in intervention schools. The study provided RiPP teachers with standard reporting forms to record implementation information, including sessions covered or topics addressed, and date of session. Reporting forms were collected after every fourth lesson during implementation in each intervention school in the 2006–07, 2007–08, and 2008–09 school years.

2.6.7 Classroom Observations (Evaluation Team)

The classroom observations conducted by the research evaluation team provided information on tangible features of the violence prevention program, such as adherence to program design, consistent delivery, and level of student participation. Evaluation staff members were trained in the use of a standardized observation form and protocol to ensure consistency of observations across classrooms and schools. During annual site visits—held each spring during program implementation—evaluation staff members conducted observations in intervention school classrooms in which RiPP had been implemented. Observations were designed to minimize disruption to classroom implementation of RiPP; each observation was based on one complete session of RiPP, which typically lasted for one class period. Appendix D provides details on the classroom observation protocol and efforts to increase coding consistency among observers.

2.6.8 Monthly Implementation Progress Reports (Liaisons)

The implementation team—composed of Tanglewood Research staff and site liaisons—conducted ongoing observations and monitoring of RiPP and Best Behavior programs. Site liaisons visited each school at least once per week to observe RiPP lessons, for the purpose of providing feedback and support to teachers, and attend Best Behavior meetings (when scheduled). Using a standardized recording form, site liaisons tracked the status of each school's progress on a monthly basis. For Best Behavior, liaisons rated the progress made on each of the program's components (not achieved, in progress, or achieved); recorded changes to the school management team; and, if the team met, noted any issues of concern and plans for the coming months. Liaisons also recorded the status of RiPP implementation (how many sessions were completed by each teacher), any issues or concerns for the month and plans for the following month. In addition, liaisons noted any issues affecting the entire school, such as a change in principal. Further details on the implementation report are provided in appendix D.

2.6.9 Year-End Implementation Reports (Liaisons)

At year end, the site liaisons rated each school on its progress with Best Behavior, based on the expected progression with the program's key elements (as stipulated by the program), and the extent of implementation of RiPP across all grades. The year-end reports provided a summary of the monthly progress reports for the year. These data provided additional information to that collected by the evaluation team regarding the fidelity and implementation of the programs. Appendix D provides additional information about the implementation report.

2.7 Measuring Program Fidelity

In this section, we outline the criteria used to measure program fidelity for this study. Program fidelity refers to the degree to which a given program is administered as intended by the program's developers (Ringwalt et al. 2009).

The criteria for measuring fidelity of RiPP implementation are presented in table 8 and include two aspects of adherence; program exposure; and student responsiveness. Adherence to the lesson plans involves following the lesson plan as prescribed, while minimizing any additions, omissions, or changes to the activities or the sequence in which the activities are presented. An example of poor implementation fidelity of this criterion is dropping an activity that the teacher finds challenging, either because of class behavior or unease with the interactive teaching method required. Adherence to the RiPP teaching techniques involves following the teaching methods prescribed by the curriculum, which could include class discussion, small group discussion, brainstorming, games, role plays, worksheets, and journal entries. By this criterion, replacing interactive teaching methods prescribed for the lesson with a teacher lecture would constitute low implementation fidelity. Exposure to the program was measured by assessing the extent to which all 16 of the curriculum's lessons were delivered to all three grades targeted. Finally, the study measured student engagement to gauge the extent of participant (student) responsiveness to the lessons.

Table 8.　RiPP fidelity criteria

RiPP fidelity criteria	Measure	Source
1a. Adherence to the scripted lesson plan	The extent to which the teacher follows the written plans for the lesson; minimizes the number of activities that are added, changed, or omitted; and presents the material in the sequence as written	Evaluation team classroom observations
1b. Adherence to the prescribed teaching techniques	The extent to which the teacher uses the teaching modalities outlined in the lesson plan	Evaluation team classroom observations
2. Exposure to the program	The number of RiPP lessons delivered by year's end	Curriculum implementation records
3. Student responsiveness	The extent to which students are engaged in lesson activities, exercises, or discussions during the lesson	Evaluation team classroom observations

Two fidelity criteria for Best Behavior are shown in table 9: implementation progress and saturation. Best Behavior is implemented by means of a school management team comprising a core group of staff that includes school administrators as well other faculty and staff. Over the course of 3 years, the school management team is expected to have an increasing role in implementing a broad set of key start-up and ongoing practices. The study measured implementation progress by assessing the extent to which the school, through the work of the school management team, achieved these key practices. A primary objective, beginning in the first year, was for schools to have functioning teams in place that met regularly (at least once per month) to address Best Behavior program implementation (practice 1). Another set of key objectives for Best Behavior involved schoolwide rules and expectations for behavior. School management teams were to create and define schoolwide rules, both for the school in general and for specific settings in the school (e.g., classroom, cafeteria, hallways) (practice 2), post the rules throughout the school (practice 3), develop plans for teaching the rules to students (practice 4), and finally, teach the schoolwide expectations to students directly and formally (practice 5). Another objective for Best Behavior was to develop (practice 6) and implement (practice 7) a positive reward system to reinforce desirable behaviors. Details about the system were to be communicated to the entire school staff so that the staff could easily understand and implement the reward system in their interactions with students. Any and all adults in the school should be aware of the system and be able to issue rewards or tokens at any time to any student. To monitor and evaluate student behavior patterns, schools were to gather, summarize, and review discipline data periodically (practice 8). An additional objective for Best Behavior in the follow-up years was to address the needs of high-risk youth who exhibit problem behaviors, by developing individualized self-monitoring and self-management plans that teach the students how to manage their own behavior (practice 9). A final objective was for schools to conduct a schoolwide needs assessment early in the year that provided the team with information on the school staff's perceived needs for improving the school discipline system (practice 10). These needs assessments were essential for formulating measurable goals that were made part of the action plans for the year.

A second fidelity criterion for Best Behavior, shown in table 9, refers to saturation, or the degree to which information concerning the schoolwide rules and expectations, as well as school

discipline and rewards policies, has reached the school staff outside of the school management team. Saturation was measured through annual surveys of random samples of teachers.

Table 9.   Fidelity criteria for Best Behavior

Best Behavior fidelity criteria	Measure	Source
Progress	The extent to which the school achieves each key program practice: (1) school management team meets regularly; (2) rules of behavior are defined and adopted; (3) rules are posted throughout the school; (4) plans are developed for teaching rules; (5) rules and expectations are taught; (6) a reinforcement system is developed; (7) the reinforcement system is implemented school-wide; (8) discipline data are collected and reviewed; (9) focus is given to the needs of high-risk youth; and (10) a schoolwide needs assessment is conducted periodically	Year-End Implementation Reports
Saturation	School staff members are aware of and understand the school's policies and rules for student behavior	Teacher Survey

2.8   Construction of Impact and Implementation Indices

This section presents details on the construction of scales and analysis variables. A combination of exploratory and confirmatory factor analytic methods was used to determine the best measurement of the various impact indices. Specific details on the factor analyses are described in appendix C.

Two main impact measures, obtained through the student survey, are used to assess the effectiveness of the RiPP and Best Behavior programs. The first measure assesses whether or not the student engaged in violent behaviors in the past 30 days. The second measure assesses whether or not the student was victimized in the past 30 days. Disaggregated measures of the two main impact scales include violence, with and without a weapon, and two aspects of victimization: overt and relational. Secondary and intermediate student outcomes measured in this study are student safety concerns, prosocial behavior, perceptions of behavioral expectations, student self-reported coping strategies, and attitudes toward violence. Other outcome measures were obtained through the teacher survey; these are teacher victimization, teacher safety concerns, and types of interactions with students.

In the following sections, we discuss the measurement model assumptions, followed by details on the construction of each of the measures. Additional details about each measure, including scale properties, can be found in appendix C.

2.8.1 Measurement Model

The survey measures for violence and victimization are composed of multiple items that ask how many times an event has happened, with a nonequivalent range for each response (i.e., "never," "once or twice," "several times," "often"). These responses are ordinal rather than continuous and were found to have distributions that violate normality assumptions at the item and composite levels. Typically, measurement models acknowledge the nonnormality of responses for individual

items and any composites formed from these items by discarding the continuous measurement framework and instead using categorical or ordinal measurements.

A series of comparative confirmatory analytic/item response theory models for categorical variables was undertaken to determine the optimal scoring for the composite measure of overall aggressive and violent behaviors. The multiple-ordered categorical factor model was used as the benchmark because this model used all the features of the data (i.e., the incidence and frequency of each response) in calculating the latent factor score. This optimal score was then compared with those achieved by various other coding schemes. Ultimately, a parsimonious sum score of dichotomized never/ever responses to violence items proved to be highly related to the optimal factor score while also providing interpretive clarity. (The sum score represents an index of how many aggressive and violent behaviors a student committed.)

As all items measuring violence and victimization exhibited a majority of "never" responses, and then a smaller group of positive responses concentrated in the "once or twice" category, dichotomizing each item allowed the assessment of the absence/presence of each event. A simple sum score was then created to form an index of the numbers of behaviors exhibited by each student. Analyses on dependent variables created in this manner have two significant benefits. First, any effects observed in evaluation models will provide estimates with an easily understood metric because the outcome is a count type variable. Second, as a count variable, these composite outcomes lend themselves to analysis models that use Poisson distributions and obviate the problem of nonnormality in the distribution of the outcomes of interest.

When data are modeled as discrete events, outcomes may be reported as a measure of event occurrence. Event rates (ERs) model the frequency of behaviors over a given period of time; in this study of violence prevention, students and teachers were asked about violent behaviors and experiences of victimization that had occurred in the preceding 30 days. Event rate ratios (ERRs) compare the incidence density of two groups. In other words, ERRs compare the number of reported events between a group of interest (intervention) and a group used as a reference (control). An ERR of 1.00 indicates the same frequency of occurrence in the two groups. Where ERRs are greater than 1.00, the intervention group reported a higher frequency of violent behaviors and experiences of victimization than the control group. Where ERRs are less than 1.00, the opposite is true; the intervention group reported a lower frequency of violent behaviors and experiences of victimization than the control group. In this report, ERRs below 1.00 are indicative of positive program impacts because they signify that the incidence of violence and victimization events is lower in the intervention group than in the control group. ERRs above 1.00 are indicative of negative program impacts because they signify that the incidence of violence and victimization events is higher in the intervention group than in the control group.

2.8.2 Outcome Indices

The student and teacher surveys are the primary sources for the outcome indices used in this study. The student survey was constructed from a number of scales that map to outcome domains of interest, based on the theory of action (figure 1, chapter 1). The teacher survey includes several scales that measure outcomes of interest as well. The indices fall under three categories of outcomes: primary, secondary, and intermediate. This section provides details on the construction of each of these outcome indices under each category.

Indices for Primary Outcomes

Primary outcomes are those that the program is explicitly designed to impact. In this case, the primary outcomes are student violence and student victimization. An index was created to measure each one, based on questions from the student survey. In addition, for each of these two indices, two additional subindices were created to better understand any differences between intervention and control schools with regard to specific types of violence.

Student Violence

A total of 14 items assessed the frequency with which a student performed violent or aggressive acts—such as threatening or hurting others, picking a fight, or bringing a weapon to school—in the past 30 days. Response options for this frequency were "never," "once or twice," "several times," and "often."[8] As noted above, these options do not reflect an even ordering of response, and so a polytomous (or ordered categorical) exploratory factor analysis model was used to explore the dimensionality of these items. Two factors emerged: one capturing general violence not involving weapons, and the other capturing weapons-related violence. The measure of student violence without a weapon was assessed with 11 items (range 0–11) asking students to indicate how often they had engaged in violence that did not involve a weapon, such as threatening to hurt another student or getting angry and yelling at another student. The measure of student violence with a weapon was assessed with three items (range 0–3) asking students how often they brought a weapon to school or threatened or injured someone with a weapon. Both measures of student violence were created by dichotomizing student responses (never vs. ever) and summing the number of types of violent behaviors occurring in the past 30 days.

Student Victimization

Student victimization (in the past 30 days) was measured with 13 items that asked how often students were victimized in certain ways, such as getting hurt with a weapon or being called an insulting name. Response options mirrored those detailed above for students' violent and aggressive behaviors. Exploratory factor analysis for ordered categorical responses indicated two dimensions. Then, a comparative analysis strategy was undertaken to determine the best scoring method for student victimization. A sum score of dichotomized (never vs. ever) experiences of victimization was determined to provide the best composite measure. One dimension was labeled "overt victimization" and comprised nine items (range 0–9), such as getting hurt with a weapon or being punched or shoved. The second dimension was labeled "relational victimization" and was based on four items (range 0–4) that asked students to indicate how often they were victimized in nonphysical ways, such as being called an insulting name, getting left out of a group, or being mocked. The measures were created by dichotomizing student responses (never vs. ever) and summing the number of types of victimizations occurring in the past 30 days.

Indices for Secondary Outcomes

These indices were created to examine possible secondary effects from the intervention (e.g., spillover effects), beyond the primary effects. These were constructed using items from either the student or teacher surveys.

---

[8] The terms used for response options were not explained or defined for students. The same terms were used in the version of the survey that was pilot tested with nine middle school students. No concerns were raised during the pilot test regarding the students' understanding of the terms.

Student Safety Concerns

The measure of student safety concerns was based on two items that asked students how often they worried that someone from the school would attack or hurt them or would bully them (range 1–4). Response options for these items were "never," "almost never," "sometimes," and "often." The analysis composite was formed as the mean of the two items.

Teacher Victimization

Teacher victimization was assessed with three items asking if teachers had been verbally abused, threatened, or physically attacked by students in the past 6 months. Response options were "never," "once," "2 to 5 times," and "more than 5 times." Only a single factor was possible with three items. Both exploratory and confirmatory models produced negative residual variances when items were treated as continuous. Recoding all three items into a never/ever binary response yielded a proper factor solution with good fit. A composite measure of teacher victimization was generated as a sum score of these recoded items. The final model indicated that dichotomization (never victimized vs. any victimization) provided optimal fit.

Teacher Safety Concerns

Teacher safety concerns were assessed with a total of six items asking about how often (never, almost never, sometimes, often) teachers felt unsafe in various indoor and outdoor school locations in the past 30 days. Exploratory factor analysis indicated a single factor for all six items, with estimation problems (negative residual variances) for any solution with two or more factors. The analysis composite was formed as the mean of all items (range 1–4).

Student Prosocial Behavior: Extended to Others

Nine items inquired about how often (never, once or twice, several times, often) in the past 30 days students performed positive behaviors toward their peers, such as inviting another student to participate in an activity or giving a compliment. Exploratory factor analysis indicated a single dimension of prosocial behavior extended to others. A composite measure was formed as the mean of all nine items (range 1–4).

Student Prosocial Behavior: Received From Others

Nine items paralleling those for students' prosocial behavior extended to others measured how often (never, once or twice, several times, often) students received prosocial behavior from peers in the past 30 days, such as an offer of help or a friendly gesture. Exploratory factor analysis indicated a single dimension. These items were averaged to form a summary scale (range 1–4).

Indices for Intermediate Outcomes

These indices were created to examine possible intermediate effects from the intervention. These were constructed using items from either the student or teacher surveys. The theoretical model for the combined intervention predicts that changes in these areas would precede changes in the primary outcomes.

Student Perceptions of Behavioral Expectations

Student perceptions of behavioral expectations were measured with 10 items from the student survey. Students reported on their level of agreement (strongly agree, agree, disagree,

strongly disagree) with statements such as "everyone knows what the school rules are" or "the school rules are strictly enforced." Exploratory factor analysis indicated that these items shared a common dimension, and a composite score was formed as the mean of these 10 items (range 1–4).

Student Attitudes Toward Violence

Eight student survey items were intended to measure student attitudes toward violence (e.g., "It's OK to hit someone who hits you first"). Response options were "strongly agree," "agree," "disagree," and "strongly disagree." Exploratory factor analysis indicated that most of these items shared a single common dimension. However, two items were found to be only very loosely associated with the common factor and, therefore, unrelated to the other attitudes items. The lack of fit of these two items was also indicated by subsequent confirmatory models and their more robust set of diagnostic indicators for model and item fit. These items ("Anyone who won't fight is going to be picked on," and "I don't need to fight because there are other ways to deal with being mad") were not included in the composite scale. This scale was formed as the mean of the remaining six items (range 1–4).

Student Self-Reported Coping Strategies

Coping strategies that students report using when faced with anger were assessed with a total of 12 items that inquired about the likelihood of reactions (very likely, likely, unlikely, very unlikely) a student might have when angry. Exploratory factor analysis indicated two independent dimensions: one measuring positive or nonviolent reactions to anger (e.g., walking away or ignoring the situation or person, trying to talk it out with the other person), and one measuring negative or violent reactions (e.g., breaking an object, yelling at the person); see appendix C for additional details. The nine positive (self-reported) coping reaction items were averaged to form a composite score (range 1–4). Simple averaging of the three negative items yielded a composite with a severely nonnormal distribution. Additional factor analytic models were estimated with competing measurement models of these items (e.g., with responses as continuous indicators; responses recoded into dichotomous indicators). The final best-fitting model with optimal distributional characteristics indicated that a sum score of the three negative self-reported coping items dichotomized as agree/disagree was optimal.

Teachers' Interactions With Students

Fifteen items on the teacher survey inquired about how often teachers used certain behaviors (e.g., maintain calmness or include victimized or isolated students in group activities) in reaction to student aggressors or victims. Response options were "did not have the opportunity," "never," "almost never," "sometimes," and "often." Factor analysis of these items indicated three distinct dimensions; each scale ranged from 1 to 5. The first scale, interactions with victims of violence, was composed of four items. The second scale also included four items and measured teacher interactions with aggressors. Five items measured classroom management behaviors and techniques and formed a third scale. Two items did not significantly contribute to any of these three factors and were omitted from composites. Scales were formed by averaging the relevant items for each subscale.

An additional set of 16 items assessed teacher perceptions of school consistency in enforcing behavior codes and rules. Teachers reported on their level of agreement (strongly agree, agree, disagree, strongly disagree) with statements such as "the school rules for student behavior are clearly

defined" or "teachers at my school consistently enforce the rules." A single factor was suggested by factor analysis, and a summary scale was formed by averaging all items (range 1–4).

2.8.3 Implementation Indices for Correlational Analyses

Implementation indices were created to measure implementation fidelity and to examine the correlation between levels of fidelity and study outcomes. These were based on the implementation fidelity criteria for RiPP and Best Behavior and were constructed using various sources of implementation data, including classroom observations, curriculum implementation records, progress reports, and the teacher survey.

RiPP Fidelity Indices

RiPP indices were constructed to assess the fidelity criteria for adherence, exposure, and student responsiveness. The adherence index was comprised of two measures obtained from items on the classroom observation protocol (see appendix D): adherence to lesson plans and adherence to teaching methods. For each item, rating scores for a school were averaged across three observations (i.e., three randomly selected RiPP classrooms in each school observed each year). The two measures were then standardized and the index constructed by averaging the two items.

Program exposure was measured using curriculum implementation records kept by teachers and by the evaluation team that tracked the number of lessons (out of 16) completed by each classroom in a year. Schools were classified into four levels based on the percentage of all classrooms in a school in grades 6 through 8 that completed all 16 lessons. The index was comprised of levels 1 to 4 for each school.

A single item from the classroom observation protocol was used to measure student responsiveness (see appendix D). The scale was constructed by averaging across the three independent observations.

Best Behavior Fidelity Indices

For Best Behavior, two indices were formed to assess the fidelity criteria for program progress and saturation. The index for program progress was constructed from ratings on the Year-End Implementation Reports for 10 program practices: (1) monthly management team meetings; (2) rules and expectations defined; (3) rules posted; (4) plan developed for teaching rules and expectations; (5) rules and expectations taught systematically; (6) reinforcement/reward system developed; (7) reinforcement/reward system implemented; (8) disciplinary data reviewed; (9) plans developed for high-risk students; and (10) needs assessments conducted. Each dimension was rated on a 3-point scale (i.e., not initiated, in progress, completed). A scale for program progress was constructed as the average across the 10 dimensions.

The index that measures saturation was derived from four items on the teacher survey that asked about various aspects of the school's rules and policies for student behavior. Teachers reported on their level of agreement (strongly agree, agree, disagree, strongly disagree) with the following statements: "the school rules for student behavior are clearly defined"; "the school rules emphasize reinforcing desired behaviors"; "the school rules emphasize consequences for undesired behaviors"; and "when a school rule is broken, it is clear to the school staff what consequences should follow." The scale was formed as the mean of the four items.

2.9　Model Specification

The study team developed a series of hierarchical, or mixed-effects, regression models to evaluate the RiPP and Best Behavior program outcomes. These models account for correlation among responses by allowing for the inclusion of multiple sources of random variation. This is done by creating a series of "nested" models that reflect the research design. The primary outcome models, for example, include student-level models (level one) nested within school-level models (level two). The models predict each outcome (e.g., violence, victimization) as a function of condition (intervention vs. control) and relevant covariates (e.g., demographic characteristics, school characteristics).

Section 2.9.1 provides additional detail on the sampling models and link functions that describe the statistical models used to assess program outcomes. The sampling models vary at level one, depending on the characteristics of the outcome measure, and determine the appropriate link function. Primary outcomes include self-reported counts of violent behaviors and victimizations. These outcomes are based on count data; accordingly, we employed generalized linear mixed models with a Poisson distribution and a log link function. Other program outcomes (e.g., safety concerns, attitudes toward violence) are based on scales that have a continuous measure. For these outcomes, we employed general linear mixed models with Gaussian (i.e., normal) distributions and an identity link function. All sampling models at level two and higher are assumed to conform to the assumptions of linearity (McCulloch and Searle 2001; Raudenbush and Bryk 2002).

Section 2.9.2 describes the structural models that detail the explanatory variables and the model coefficients. The structural model is assumed to be a linear and additive function of the outcome variable; for the Poisson models, the assumptions of linearity and additivity apply to the transformed outcome variable. These models are determined by the research question addressed rather than the characteristics of the outcome.

2.9.1　Sampling Models and Linking Functions

The sampling model describes the expectation and distributional characteristics of the outcome at each level of the model. For the variables that constitute the outcomes of interest for this evaluation, level-one sampling models vary according to the characteristics of the outcome under consideration.

For variables that express the outcome of interest as a continuous measure, the level-one sampling model can be expressed as:

$$Y_{i:kp} \mid \mu_{i:kp} \sim N\left(\mu_{i:kp}, \sigma^2\right). \tag{1}$$

This indicates that, given the predicted value $\mu_{i:kp}$, the outcome measure of student i ($i = 1 \ldots m$) located in the kth condition ($k = 0, 1$) and that p^{th} pair ($p = 1 \ldots 20$) is normally distributed with expected value of $\mu_{i:kp}$ and a constant variance, σ^2. The expectations of these values are expressed as:

$$E\left[Y_{i:kp} \mid \mu_{i:kp}\right] = \mu_{i:kp} \text{ and } \mathrm{Var}\left(Y_{i:kp} \mid \mu_{i:kp}\right) = \sigma^2 \tag{2}$$

for the mean and variance, respectively. When the outcome of interest follows a normal distribution, it can be expressed directly as a function of a set of explanatory variables. However, to simplify the expression of the structural models that follow, we note that:

$$\eta_{i:kp} = \mu_{i:kp} ,\tag{3}$$

which indicates that the modeled outcome $\eta_{i:kp}$ is equal to the expected value of $Y_{i:kp}$.

For variables that express the outcome of interest as count data constrained to be nonnegative and take on integer values (e.g., number of events), the level-one sampling model can be expressed as

$$Y_{i:kp} \mid \mu_{i:kp} \sim P\!\left(\omega, \lambda_{i:kp}\right)\tag{4}$$

indicating that the outcome of interest $(Y_{i:kp})$ is an event that follows a Poisson distribution and has an expected ER of $\mu_{i:kp}$ for student i $(i = 1\ldots m)$ located in the kth condition $(k = 0, 1)$ and the pth pair $(p = 1\ldots20)$, over a given period of time (ω_p). To simplify the following, we note that ω is a nonvarying time period in this study $(\omega_p = \text{past 30 days})$ that has no impact on estimation and drop it from further notation. According to the Poisson distribution, the expected event count and variance of $Y_{i:kp}$ are expressed as:

$$E\!\left[Y_{i:kp} \mid \mu_{i:kp}\right] = \mu_{i:kp} \;\; \text{and} \;\; \mathrm{Var}\left(Y_{i:kp} \mid \mu_{i:kp}\right) = \mu_{i:kp} .\tag{5}$$

As (5) indicates, the expected variance of the outcome under a Poisson distribution is equal to the expected value. However, this strict expectation is seldom obtained. Rather, $\mathrm{Var}\left(Y_{i:kp} \mid \mu_{i:kp}\right)$ is likely to be larger (overdispersed) or smaller (underdispersed) than $\mu_{i:kp}$. To account for this, the model can be generalized by the inclusion of a scaling factor $\phi_{i:kp}$. Accordingly, the estimated residual variance of the Poisson models will be estimated as $\varphi_{\bullet:kp}\mu_{i:kp}$.

The canonical link when the level-one sampling distribution is Poisson is the log link, which can be expressed as follows:

$$\eta_{i:kp} = \log\!\left(\mu_{i:kp}\right)\tag{6}$$

and indicates that the modeled outcome $\eta_{i:kp}$ is equal to the log(nl) of the predicted value of $Y_{i:kp}$.

The sampling distributions for level-two (and higher) models express the characteristics of the modeled random effects. Here, the term $\left(z_{0:q}\right)$ is used to indicate random effects. For all of the structural models presented below, random effects are assumed to follow a normal distribution with

$$z_{0:q} \mid \gamma_{0:q} \sim N\!\left(\gamma_{0:q}, \sigma_u^2\right).$$

2.9.2 Structural Models

The structural models are used to express the expectation of the outcome as the function of a series of explanatory variables. In general form,

$$g\!\left(\mu_{i:kp}\right) = x_{i:kp}\beta_{i:kp} + z_{0:kp}u_{0:kp} .\tag{7}$$

Here, $g\!\left(\mu_{i:kp}\right)$ is the expected value of the outcome; $x_{i:jk}\beta_{i:jk}$ is a shorthand representation for the set of fixed-effect covariates and coefficients; and $z_{0:kp}u_{0:kp}$ is a shorthand representation for the set of random-effect covariates and coefficients.

As noted in the previous section, when the outcome of interest is represented by a variable that has a continuous measure, $g(\cdot)$ is the identify link, and from equation (3) it follows that

$$E\left[Y_{i:kp}\right] = \eta_{i:kp}. \tag{8}$$

Similarly, when the outcome of interest is represented by count data, $g(\cdot)$ is the log link, and from equation (6) it follows that

$$E\left[Y_{i:kp}\right] = \exp\left(\eta_{i:kp}\right). \tag{9}$$

For outcomes that meet the assumptions of linearity, we employ general linear mixed models where the expectation for $Y_{i:kp}$ in equation (8) is the appropriate form. However, when the assumptions of linearity cannot be met, we will employ generalized linear models where the expectation for $Y_{i:kp}$ in equation (9) is the appropriate form.

2.9.3 Matched Nested Cross-Sectional Models for Full Sample

The matched nested cross-sectional model (matched analysis) is an example of a cross-classified hierarchical model. Under this model, students are nested in schools; schools are nested in pairs *and* in experimental conditions; and pairs are crossed with experimental condition (i.e., each pair is represented at each level of condition). One feature of the cross-classified model is that school, per se, does not appear in the model. This is due to the fact that, in the matched design, each school is uniquely located in a cell defined by the condition-by-pair interaction. Accordingly, noting that a student is nested in the pth pair and the kth condition is the same as noting that the student is nested in the jth school; this can be seen in equations (10)–(12).

To more fully understand estimation under this design, it must first be recognized that, with matching, the schools in one experimental condition are *not* fully exchangeable with the schools in the other experimental condition. Here, the intervention effect is estimated as an adjusted net difference among the pairs of schools. This means that the two schools within a pair are differenced, and the average of those differences is the measure of program impact. Accordingly, the test of the program impact assesses variation among the adjusted condition means against the variation among the condition-by-pair means. The analysis is conducted on the expectation that differencing the matched pairs will provide a more precise estimate of the program effect. Preliminary exploration of the main impact models did in fact confirm that, compared to an analysis model that ignored the matching structure, the analysis that included matching provides a lower standard error of program impact and, accordingly, greater statistical power.

As a cross-sectional model, this employs data collected from a census of students at baseline (preintervention) and another census of students at follow-up (postintervention). This approach is appropriate when the focus is on whether or not the program effects change in a population of youth in intervention schools relative to a population of youth in control schools.

General(ized) Hierarchical Linear Model Presentation

The structural model used to assess the effects of the RiPP program in the general population of students can be articulated as a two-level hierarchical linear model (HLM). The level-one model, the student-level model, describes the outcome of interest as a function of student-specific parameters. The level-two model expresses the intercept of the student-level model as a set of parameters describing the contextual factors (e.g., school effects) influencing student outcomes.

Student-level model (level one). In this model, $\eta_{i:kp}$ represents the response of the ith student nested in the pth pair and the kth condition. The model includes an intercept parameter $(\beta_{0:kp})$, which can be interpreted as the mean response across all students located in the jth school. The model also includes a set of student-specific covariates (gender [0 = boys, 1 = girls], race/ethnicity, and number of parents in the household [HH]). Any variation between the predicted value and the observed value is attributed to residual error $(\varepsilon_{i:kp})$ in the Gaussian model but is a function of the expected mean for the Poisson model:[9]

$$\eta_{i:kp} = \beta_{0:kp} + \beta_{1:kp}\text{GENDER} + \beta_{2:kp}\text{RACE} + \beta_{3:kp}\text{HH} + \varepsilon_{i:kp} \qquad (10)$$

School-level model (level two). At the school level, the intercept parameter from the student-level model is expanded. The school-level model (11) describes $\beta_{0:kp}$ as a function of the mean intercept across all schools $(\gamma_{0:00})$ and two school-level covariates that account for school size (SIZE) and the baseline mean value of the student outcome at the school level $(\overline{Y}_{(t-1)})$.

An indicator variable (COND) identifies each school as a member of either the intervention or control condition; its coefficient, $\gamma_{0:01}$, accounts for the sum of the marginal differences among the pairs of intervention and control schools. The model includes two random effects. The first random effect, $(u_{0:0\,p})$, allows for variation among the marginal differences of the p pairs of schools; the second random effect, $(u_{0:kp})$, allows for school-to-school variation among the cells defined by the pair-by-condition interaction.

$$\beta_{0:kp} = \gamma_{0:00} + \gamma_{0:01}\text{COND} + \gamma_{0:02}\text{SIZE} + \gamma_{0:03}\overline{Y}_{(t-1)} + u_{0:0\,p} + u_{0:kp} \qquad (11)$$

General(ized) Mixed Model Presentation

Through substitution, the school-level model can be incorporated into the student-level model, providing the mixed-effects model shown in equation (12). In this expression of the model, fixed effects are presented in standard typeface, and random effects are presented in **bold** typeface. Fixed effects associated with gammas (γ) represent school-level effects, while those associated with betas (β) represent student-level effects.

$$\eta_{i:kp} = \gamma_{0:00} + \gamma_{0:01}\text{COND} + \gamma_{0:02}\text{SIZE} + \gamma_{0:03}\overline{Y}_{(t-1)}$$

$$(12)$$

$$+ \beta_{1:kp}\text{GENDER}_{0:kp} + \beta_{2:kp}\text{RACE}_{0:kp} + \beta_{3:kp}\text{HH}_{0:kp} + \boldsymbol{u_{0:0\,p}} + \boldsymbol{u_{0:kp}} + \boldsymbol{\varepsilon_{i:kp}}$$

As shown in equation (12), $\boldsymbol{\varepsilon_{i:kp}} + \boldsymbol{u_{0:0\,p}} + \boldsymbol{u_{0:kp}}$ represents the total variation in the outcome, $Y_{i:kp}$, and provides the components that describe the intraclass correlation coefficient (ICC). ICC indexes the loss of precision associated with clustering and is defined as the proportion of total variation in the outcome measure associated with the schools nested within condition; it is specified as shown in equation (13):

---

[9] For the Poisson model, $\boldsymbol{\varepsilon_{i:kp}}$ is $\boldsymbol{\phi_{\bullet:kp}\mu_{i:kp}}$, as noted in section 2.9.1

Impacts of a Violence Prevention Program for Middle Schools

$$ICC = \frac{\sigma^2_{u_{0:kp}}}{\sigma^2_{\varepsilon_{i:kp}} + \sigma^2_{u_{0:0p}} + \sigma^2_{u_{0:kp}}} \tag{13}$$

2.9.4 Nested Cohort Models for Program Outcomes Among Selected High-Risk Youth

The second research question concerns the impact of the intervention on youth who are at higher risk to express violent and aggressive behaviors and, subsequently, likely to derive the most benefit from the program. A combination of behavioral and attitude or belief measures has been used to identify youth who may have a higher proclivity toward violence from the general student population; details on the criteria used to categorize high-risk youth are presented in section 2.5.

The models employed to assess program outcomes among high-risk youth are different from those employed to assess program outcomes on the general population of students. For the high-risk youth, the interest is in whether or not the RiPP and Best Behavior intervention led to individual change across time. To address this interest, nested cohort models using difference-in-difference estimation have been developed to assess changes on self-reported measures of violence and victimization among high-risk youth in intervention schools relative to changes among high-risk youth in control schools. Given the study's interest in intraindividual change among the high-risk youth, this modeling approach is better aligned with the research questions that this analysis is designed to address.

These models use data collected on the same sample of students at each measurement occasion. One benefit of the repeated measures design is the ability to capitalize on the correlation that occurs as a result of taking repeated measures on the same students. To the extent that the students' measures are highly correlated over measurement occasion, the model can be more efficient than a similar model that relies on cross-sectional data.

General(ized) Hierarchical Linear Model Presentation

The structural model used to assess the effects of the RiPP program among high-risk youth can be articulated as a three-level HLM. The observation-level model (level one) describes the outcome of interest as a function of initial status and change over time. The student-level model (level two) includes two models, one for each of the two parameters of the observation-level model. The school-level model (level three) also includes two models, one for each of the intercepts in the two student-level models.

Observation-level model (level one). In this model, $\eta_{ti:j:k}$ represents the response of the ith student measured on occasion t, nested in the jth school of the kth condition. The model includes two parameters, one describing initial status, $(\beta_{0i:j:k})$, and the other describing the incremental change in $\eta_{ti:j:k}$ associated with a one-unit change in the variable TIME. For this model, TIME is indexed as "0" for baseline measures and as "1" for follow-up measures, leading to the interpretation of $\beta_{1i:j:k}$ as a change, or growth, parameter. Any variation between the predicted value and the observed value is accounted for by residual error $(e_{ti:j:k})$ in the Gaussian model but is a function of the expected mean for the Poisson model:[10]

---

[10] For the Poisson model, $\varepsilon_{i:kp}$ is $\phi_{\bullet:kp}\mu_{i:kp}$, as noted in section 2.9.1

$$\eta_{ti:j:k} = \beta_{0i:j:k} + \beta_{1i:j:k}\,\text{TIME} + e_{ti:j:k}\,. \tag{14}$$

Student-level models (level two). At the student level, each of the parameters (β) from the observation-level model is expanded. The first student-level model in equation (15) describes $\beta_{0i:j:k}$, the initial status of the ith student in the jth school of the kth condition, as a function of the intercept value of all students in school j($\gamma_{00:j:k}$); a set of student-specific covariates (gender, race/ethnicity, and number of parents in the household [HH]); and a random effect ($u_{0i:j:k}$) that allows for student variation from intercept school value. The coefficients associated with these covariates are not of direct interest.

$$\beta_{0i:j:k} = \gamma_{00:j:k} + \gamma_{00:j:k}\text{SEX} + \gamma_{00:j:k}\text{RACE} + \gamma_{00:j:k}\text{HH} + u_{0i:j:k} \tag{15}$$

$$\beta_{1i:j:k} = \gamma_{10:j:k} + u_{1i:j:k} \tag{16}$$

The second student-level model in equation (16) describes $\beta_{1i:j:k}$, the change or growth over time, of the ith student in the jth school of the kth condition as a function of the mean slope associated with school j($\gamma_{10:i:k}$) and a random effect ($u_{1i:j:k}$) that allows for student variation from the school-specific slope. Given the structure of the data being modeled, $u_{1i:j:k}$ is not directly estimable separate from $e_{ti:j:k}$, as noted in the mixed model specification by the brackets [ ] in equation (19).

School-level models (level three). At the school level, the intercepts from the student-level models are expanded. The first school-level model in equation (17) describes $\gamma_{00:j:k}$, the initial status of the jth school of the kth condition as a function of the mean intercept value across all schools ($\lambda_{00:0:k}$). This model includes an indicator variable (COND) identifying schools as a member of either the intervention or control condition; its coefficient ($\lambda_{00:1:k}$) accounts for any difference in initial status between schools in the two conditions. An additional covariate that indexes school size (SIZE) is included to account for variability in the outcomes that may be associated with differences in school size; its coefficient ($\lambda_{00:2:k}$) is not of direct interest. A random effect ($u_{00:j:k}$) allows for school-to-school variation from the overall intercept value.

$$\gamma_{00:j:k} = \lambda_{00:0:k} + \lambda_{00:1:k}\text{COND} + \lambda_{00:2:k}\text{SIZE} + u_{00:j:k} \tag{17}$$

$$\gamma_{10:j:k} = \lambda_{10:0:k} + \lambda_{10:1:k}\text{COND} + u_{10:j:k} \tag{18}$$

The second school-level model in equation (18) describes $\gamma_{10:j:k}$, the change over time of the jth school of the kth condition as a function of the mean slope across all schools $\lambda_{10:0:k}$. This model includes an indicator variable (COND) identifying schools as a member of either the intervention or control condition; its coefficient ($\lambda_{10:1:k}$) accounts for any difference in mean slope between schools in the two conditions. The random effect ($u_{10:j:k}$) allows for school-to-school variation from the condition-specific mean slope.

General(ized) Mixed Model Presentation

The five models described above can be combined into the familiar mixed-effects model shown in equation (19). In this expression of the model, fixed-effect terms are presented in standard typeface, and random-effect terms are presented in **bold** typeface. Fixed effects associated with lambdas (λ) represent school-level effects, while those associated with gammas (γ) represent student-level effects.

$$\eta_{ti:j:k} = \lambda_{00:0:k} + \lambda_{00:1:k}\text{COND} + \lambda_{00:0:k}\text{TIME} + \lambda_{00:0:k}\text{TIME}*\text{COND}$$

$$+ \lambda_{00:2:k}\text{SIZE} + \gamma_{00:j:k}\text{SEX} + \gamma_{00:j:k}\text{RACE} + \gamma_{00:j:k}\text{HH} \tag{19}$$

$$+ \mathbf{u_{0i:j:k}} + u_{00:j:k} + u_{10:j:k}\mathbf{TIME} + \left[u_{1i:j:k}\mathbf{TIME} + e_{ti:j:k} \right]$$

In equation (19), $\boldsymbol{u_{1i:j:k}}\mathbf{TIME}$ is the component of variation associated with repeated measures within person at a given point in time; as previously noted, that component cannot be estimated apart from residual error in this model and is dropped from further notation. Thus, $\boldsymbol{u_{0i:j:k} + u_{00:j:k} + u_{10:j:k}}\mathbf{TIME} + e_{ti:j:k}$ represents the total variation in the outcome, $Y_{ti:j:k}$, and provides the components that describe ICC. ICC indexes the loss of precision associated with clustering and is defined as the proportion of total variation in the outcome measure associated with the schools nested within condition; it is specified in (20) as follows:

$$\text{ICC} = \frac{\sigma^2_{u_{00:j:k}} + \sigma^2_{u_{10:j:k}}}{\sigma^2_{u_{0ij:k}} + \sigma^2_{u_{00:j:k}} + \sigma^2_{u_{10:j:k}} + \sigma^2_{e_{ti:j:k}}} \tag{20}$$

2.9.5 Program Impacts on Teacher Behaviors

Multivariate models are specified to examine teacher outcomes for victimization and safety concerns. We examine these impacts given that, based on the theory of action, these are expected to show changes before we see changes in student impact outcomes. In these multilevel models, teachers are nested within schools, and schools are nested within matched pairs randomized to experimental condition. The statistical models assessing program impacts on teacher behavior are similar to those presented in equations (10) and (11) to examine main impacts on student behavior. The teacher behavior models include an indicator variable for condition (intervention vs. control) and a covariate controlling for school size.

2.9.6 Analytic Approaches for Mixed Model Regression

To account properly for the multiple sources of random variation that result from the randomization of schools to condition with measurements taken on students nested within those schools, the study specified multilevel regression equations using SAS PROC MIXED (SAS Institute 2004) and SAS PROC GLIMMIX (SAS Institute 2006) for general and generalized linear mixed models, respectively. These two procedures offer a flexible approach to modeling the longitudinal and multilevel regression models specified here. A primary strength of the mixed model approach is that multiple random effects may be modeled independently. Under the general linear mixed model, the random effects are assumed to be independent and normally distributed; the random effects necessary to avoid misspecification for each model are identified in the preceding subsection. The

analyses can be extended to non-Gaussian data in the generalized linear mixed model through the appropriate specification of an alternative error distribution and link function. The standard errors estimated and significance tests conducted account for the fact that schools (not students) are the units of random assignment.

The models were estimated using restricted maximum likelihood (REML) for general linear mixed models and the restricted pseudo-likelihood for generalized linear mixed models. These approaches provide parameter estimates by maximizing the probability that the predicted values agree with the observed data. They are iterative, similar to maximum likelihood (ML) estimation, but provide separate estimation for fixed and random effects. Separate estimation of the fixed and random components is less efficient, which may result in a slightly larger mean square error; however, estimates obtained in this manner produce less of a downward bias than ML estimates.

2.9.7 Exploratory Gender Subgroup Analyses

The structural model used to explore the differential effect of the violence prevention program on boys and girls among the general population of students can be articulated as a two-level HLM closely following the specification of the primary impact model. Here, however, an additional level-one, fixed effects model is added. This model allows for the specification of the cross-level interaction of gender and condition.

Student-level model (level one). The first level-one model describes the outcome of interest as a function of student-specific parameters and is similar to the level-one model used to assess primary impacts. In this model, $\eta_{i:kp}$ represents the response of the ith student nested in the pth pair and the kth condition. The model includes an intercept parameter $\left(\beta_{0:kp}\right)$, which can be interpreted as the mean response across all students located in the jth school. The model also includes a set of student-specific covariates (gender [0 = boys, 1 = girls], race/ethnicity, and number of parents in the household [HH]). Any variation between the predicted value and the observed value is attributed to residual error $(\varepsilon_{i:kp})$ in the Gaussian model but is a function of the expected mean for the Poisson model:[11]

$$\eta_{i:kp} = \beta_{0:kp} + \beta_{1:kp}GENDER + \beta_{2:kp}RACE + \beta_{3:kp}HH + \varepsilon_{i:kp} \tag{21}$$

The second level-one model is needed to express the interaction between gender and condition. In equation (22), we disaggregate the coefficient associated with gender $\left(\beta_{1:kp}\right)$ into two components. The first component $\left(\gamma_{1:00}\right)$ accounts for the overall impact associated with gender and captures the simple difference between girls and boys in the control group. The second component $\left(\gamma_{1:01}\right)$ is associated with condition. This term accounts for the interaction of condition and gender; its expression is made clear in the mixed model presentation:

$$\beta_{1:kp} = \gamma_{1:00} + \gamma_{1:01}COND \tag{22}$$

School-level model (level two). The level-two model expresses the intercept of the student-level model as a set of parameters describing the contextual factors (e.g., school effects) influencing student outcomes. This model is similar to the level-two model used to assess primary

---

[11] For the Poisson model, $\varepsilon_{i:kp}$ is $\phi_{\bullet:kp}\mu_{i:kp}$, as noted in section 2.9.1

impacts. The level-two model (23) describes $\beta_{0:kp}$ as a function of the mean intercept across all schools ($\gamma_{0:00}$) and two school-level covariates that account for school size (SIZE) and the baseline mean value of the student outcome at the school level ($\overline{Y}_{(t-1)}$).

An indicator variable (COND) identifies each school as a member of either the intervention or control condition; its coefficient, $\gamma_{0:01}$, accounts for the sum of the marginal differences among the pairs of intervention and control schools. Due to the presence of the gender-by-condition interaction, this term represents a simple main effect and should not be interpreted in the same way as the condition term in the main impact models. The model includes two random effects. The first random effect, ($u_{0:0\,p}$), allows for variation among the marginal differences of the p pairs of schools; the second random effect, ($u_{0:kp}$), allows for school-to-school variation among the cells defined by the pair-by-condition interaction.

$$\beta_{0:kp} = \gamma_{0:00} + \gamma_{0:01}\text{COND} + \gamma_{0:02}\text{SIZE} + \gamma_{0:03}\overline{Y}_{(t-1)} + u_{0:0\,p} + u_{0:kp} \tag{23}$$

Generalized Mixed Model Presentation. Through substitution, the school-level model can be incorporated into the student-level model, providing the mixed-effects model shown in equation (24). In this expression of the model, fixed effects are presented in standard typeface, and random effects are presented in bold typeface. Fixed effects associated with gammas (γ) represent school-level effects, while those associated with betas (β) represent student-level effects.

$$\eta_{i:kp} = \gamma_{0:00} + \gamma_{0:01}\text{COND} + \gamma_{0:02}\text{SIZE} + \gamma_{0:03}\overline{Y}_{(t-1)} + \gamma_{1:01} + \gamma_{1:00}\text{GENDER}_{0:kp}$$

$$+ \gamma_{1:01}\text{COND} * \text{GENDER}_{0:kp} + \beta_{2:kp}\text{RACE}_{0:kp} + \beta_{3:kp}\text{HH}_{0kp} + \boldsymbol{u_{0:0\,p}} + \boldsymbol{u_{0:kp}} + \boldsymbol{\varepsilon_{i:kp}} \tag{24}$$

The condition-by-gender interaction is clearly shown in this presentation. The coefficient associated with the interaction $(\gamma_{1:01})$ captures the additional multiplicative effect that allows the impact of the program to vary by gender.

2.9.8 Implementation and Program Exposure Analyses

Exploratory analyses presented in chapter 5 are designed to assess the association of implementation of the RiPP curriculum and Best Behavior program with program impacts and the association of program exposure with self-reported violence and victimization. This approach recognizes that schools differ in terms of capacity and commitment, management and resources, and that these factors may influence the ability of the prescribed program to achieve anticipated aims. It also recognizes that students' level of exposure to program components may play a key role in shaping behavior change.

Implementation Analyses

The purpose of the correlational analyses is to contextualize the impact findings and to evaluate whether or not variation in program implementation is related to program impacts. These analyses will be conducted at the school level ($n = 18$). Two types of association will be explored. First, we will assess the relationship among the five implementation indices (three for RiPP and two

for Best Behavior) described in section 2.8.3. These tests will explore the degree to which fidelity of implementation is correlated between the curricular and whole-school aspects of the intervention. Second, we will assess the relationship between implementation and program impacts. Bivariate associations will be assessed through correlational analysis. The Pearson product-moment correlation coefficient (r) will be used to examine the degree of association between two variables. All assessments of bivariate association will take into account the small number of independent units ($n = 18$) included in the analysis.

Exploratory Analysis Examining Maximum Program Exposure

An analysis designed to examine program outcomes among a subgroup of students who received "full dosage" via exposure to the RiPP curriculum and Best Behavior program across the 3-year study period is presented in section 5.3. The model is constrained to include only the 3,240 students who remained in the same school—in both the intervention and control schools—across the 3 years of program implementation (2006–09); this number represents 44 percent of the student sample included in the year one impact analyses. For this analysis, we employ the same model used in the primary impact models (section 2.9.3), limiting the included sample of students as described. By limiting the sample of included students in this manner, we are able to explore the association between maximum program exposure with students' self-reported violence and victimization outcomes. Because the differences in mobility rates between intervention and control schools were not statistically significant (see table 6), this analysis is appropriate. However, unlike the impact estimations for the entire sample, this analysis cannot determine causality because mobility was self-selected.

2.10 Observed Precision

Statistical power calculations guide a number of key decisions in the design and execution of the research study.[12] Power is the probability of observing a statistically significant difference, where such a difference exists. In other words, power indicates the probability that an intervention effect will be judged to be statistically significant, given the assumptions of the specified model. Statistical power is a function of sample size, the assumed magnitude of the intervention impact, and the anticipated level of random variation in the measure of the outcome. Based on the observed distribution of the variables employed to examine violence and victimization in the study, the created indices are expressed as counts of events. Accordingly, statistical models for violence and victimization have been estimated on the expectation of a Poisson distribution with a log link function.

The models employed for these analyses partition the total variation into within-subjects (i.e., school) variation, represented by the term $\varepsilon_{i:kp}$,[13] and between-subjects variation, represented by the terms $u_{0:0p}$ and $u_{0:kp}$ (variation among pairs of schools and variation between schools within a given pair, respectively). These components can be used to estimate ICCs based on equation (13), which is re-presented here:

---

[12] This section contains information on the level of precision actually observed in the study. Information on a priori estimates of precision, which were used to determine the evaluation's sample size, is in appendix E.

[13] For the Poisson model, $\varepsilon_{i:kp}$ is $\phi_{\bullet:kp}\mu_{i:kp}$, as noted in section 2.9.1

$$\text{ICC} = \frac{\sigma^2_{u_{0:kp}}}{\sigma^2_{\varepsilon_{i:kp}} + \sigma^2_{u_{0:0p}} + \sigma^2_{u_{0:kp}}} \, .$$

The estimated covariance parameters provided by SAS PROC GLIMMIX (SAS Institute 2006) are scaled in terms of the variance function of the outcome and must be adjusted before the ICCs can be calculated (Murray 1998). For the Poisson distribution, the variance function is λ, the mean of the outcome. The between-subjects components of variation are scaled as equation (25):

$$\sigma^2_{u_{0:kp}\,(\text{scaled})} \cong \frac{\sigma^2_{u_{0:kp}\,(\text{unscaled})}}{\lambda^2_{i:kp}} \text{ and } \sigma^2_{u_{0:0p}\,(\text{scaled})} \cong \frac{\sigma^2_{u_{0:0p}\,(\text{unscaled})}}{\lambda^2_{i:kp}}, \qquad (25)$$

so that the unscaled variance components are obtained by multiplying the scaled values by the square of the variance function. The residual component of variation is scaled as equation (26):

$$\sigma^2_{\varepsilon_{i:kp}\,(\text{scaled})} \cong \frac{\sigma^2_{\varepsilon_{i:kp}\,(\text{scaled})}}{\lambda^2_{i:kp}}, \qquad (26)$$

so that the unscaled residual error is obtained by multiplying the scaled values by the variance function. As table 10 indicates, ICCs are smaller than anticipated. The inclusion of covariates (baseline value of the outcome of interest, demographics) reduces the magnitude of ICC for all main and disaggregated outcomes, save the outcome looking at weapons-related violence.

To better understand the ICC and its role in these analyses, it is helpful to understand what the ICC represents. The ICC is a proxy for the various and unmeasured contextual factors that lead students in a particular school to be more similar to each other than to students in another school. The magnitude of the ICC is a function of the type of unit in which clustering occurs (e.g., families, schools, communities) and the degree to which contextual factors related to the unit impact the observed outcomes. In many studies conducted in schools where educational outcomes are assessed on students nested within classrooms or schools, and where classrooms or schools are nested within intervention conditions, the magnitude of the ICC can be quite large. When health behaviors are assessed in school settings, the magnitude of the ICC will often be much more modest. Though modest, a positive ICC still indicates that students in a given school are more similar to other students in their school than they are to students in different schools. This lack of independence at the student level leads to correlation in the data and an increase in the model variance which must be accounted for in a valid analysis.

Model-based estimates of MDEs are presented in table 10. MDEs indicate the smallest differences that would allow us to reject the null hypothesis with confidence (i.e., 80 percent statistical power). MDEs are presented as ERRs that compare the mean number of events in intervention schools to the mean number of events in control schools, When the two ERs are equal, ERR is 1.00. As the ER among intervention schools moves away from the ER in the control schools, ERR moves away from 1.00. Accordingly, the two ERRs listed for each outcome in table 10 provide an interval space around the null value of 1.00 and identify the minimum program impacts that would be viewed statistically significant. For the current study, ERRs below 1.00 indicate positive program effects, while those above 1.00 indicate negative program effects.

Table 10. **Observed intraclass correlation coefficients (ICCs) and minimum detectable effects (MDEs) reported as event rate ratios (ERRs) for self-reported violence and victimization—Year three**

	ICCs		ERRs as MDE[1]	
Self-reported student outcome	Unadjusted	Adjusted	Positive program effect	Negative program effect
Violence (All items)	0.00876	0.00397	0.88	1.14
Without a weapon	0.00983	0.00502	0.87	1.14
With a weapon	0.00174	0.00030	0.73	1.37
Victimization (All items)	0.00487	0.00194	0.92	1.09
Overt	0.00544	0.00402	0.89	1.12
Relational	0.01188	0.00582	0.90	1.11

[1] Values below 1.00 indicate positive program effects; that is, event rates (ERs) in intervention schools are below those in control schools. Values greater than 1.00 indicate negative program effects; that is, ERs in control schools are below those in intervention schools.
SOURCE: Student survey, spring 2009.

MDE for violent behavior (all items) in table 10, for example, indicates that to reject the null hypothesis and view the ER among intervention schools as different from the ER among control schools would require a ratio of at least 0.88 to 1.00 (positive program effect) or 1.14 to 1.00 (negative program effect). To further illustrate, consider the case of an intervention effect where the mean ER among the students in the control schools is 2.89 events in the past 30 days. Here, the mean ER among students in the intervention schools would have to be 2.54 events in the past 30 days *or less* to achieve statistically significant positive program effects.

Estimations used to derive MDEs presented in table 10 involved a number of assumptions having to do with values derived from the empirical model. These assumptions include the following:

- The scale parameter observed in data reflects the true population parameter.

- The model-based variance components are true for population.

- The parameter estimates are asymptotically normal.

These assumptions underscore one of the main differences between linear and generalized linear modeling. In the former, covariance parameter values are assumed constant and independent of location; this assumption is untenable for the latter. It is important also to bear in mind that ERRs are calculated based on coefficients that have been estimated in the natural logarithmic (ln) scale where $\ln(1.00) = 0.00$. This fact, a statistical necessity, means that direct translation of standardized MDEs to ERR-based MDEs is not straightforward.

Chapter 3.
Implementation of the Violence Prevention Program

This chapter provides details about the implementation of the combined intervention in the 18 intervention schools that participated during all 3 years. The chapter begins by describing the core elements of the two programs that constitute the intervention. The chapter then provides a description of the training and technical assistance to support implementation of the programs. This is followed by a discussion of how the programs were implemented, including fidelity of implementation and the challenges that were encountered. The chapter then provides a comparison of programs offered in intervention schools and control schools for violence prevention or similar goals. The chapter concludes with a discussion of program costs.

The key descriptive findings regarding implementation of the curriculum portion of the program are reported across all 3 years because the impact estimates in this report use baseline data from the beginning of year one and outcome data from the end of year three. Therefore, the impact estimates are cumulative across all 3 years of program implementation. The key implementation findings include the following:

- **In a majority of intervention schools, students were exposed to the full set of 16 Responding in Peaceful in Peaceful and Positive Ways (RiPP) lessons in each of the 3 years of implementation.** Between 61 percent and 72 percent of schools delivered all 16 lessons to all classrooms in each year of program implementation, while another 17 percent to 22 percent of schools delivered all lessons in at least three-fourths of these classrooms.

- **The curriculum was not fully delivered with fidelity in year three.** In year three, teachers in 44 percent of the intervention schools were observed to deliver lessons with few or no deviations from the written lesson plan (e.g., adding or modifying activities or changing the activity sequence), according to classroom observations by the evaluation team. With regard to teachers delivering lessons with few or no deviations from the prescribed teaching strategies (e.g., using role plays or small group discussions), teachers met this second criterion in 56 percent of schools in the third year.

- **Interviewed RiPP teachers cited challenges with using one or more of the prescribed teaching techniques or approaches.**[14] Eighty-eight percent of teachers interviewed in year three mentioned difficulties with implementing at least one of five RiPP techniques or approaches, and 27 percent mentioned difficulties with implementing three or more of the five techniques or approaches.

- **The extent to which students were engaged with the curriculum varied across the 3 years.** The evaluation team observed the same cohort of students receiving the curriculum over 3 successive years. These students were found to be engaged during the

---

[14] RiPP teachers were to use the following teaching techniques and approaches: make RiPP real (tie it to students' daily lives); role plays; small group work, discussion, and brainstorming; encourage self-talk by students; use *Review* and *Closure* to begin and end sessions.

lesson activities, exercises, and discussions in 89 percent of the intervention schools in year one, 94 percent in year two, and 69 percent in year three.

The key descriptive findings regarding the implementation of the whole-school portion of the program across the 3 years include the following:

- **Principal support and commitment for the whole-school portion of the intervention was mixed in year three.** Principals at 72 percent of the intervention schools were rated as supportive in year three, according to liaisons who helped implement the program. In addition, liaisons reported that 50 percent of the principals in year three used their leadership to promote the program. In the third year of implementation, slightly more than half (56 percent) of the principals were regularly involved with the school management team.

- **By the end of the third year, the majority of intervention schools had instituted behavioral rules and rewards. In addition, a majority of teachers agreed that the rules were well defined and clear with regard to the behaviors being targeted.** By the end of the third year, 83 percent of intervention schools had developed and posted school rules in the school, 78 percent had developed and instituted a token reward system for adhering to school rules, and 78 percent had developed lesson plans and taught the school rules in classrooms. In addition, 87 percent of teachers at intervention schools agreed or strongly agreed that school rules were clearly defined. However, a smaller percentage (64 percent) of the teachers surveyed in year three agreed or strongly agreed that that it was clear what consequences would follow when school rules were broken.

- **Among the cited challenges with implementing Best Behavior were finding time to implement the program, obtaining teacher buy-in, maintaining student interest, and funding the rewards program.** School management team members interviewed in year three mentioned difficulties with finding the time to implement the program. Others talked about issues with low teacher buy-in, a lack of student interest in the rewards offered through the reward system, and continued problems with funding the reward system.

- **By design, no control school implemented RiPP or Best Behavior during the 3 years of the study.** However, there were various violence prevention activities already in place in the participating schools. Between eight schools and nine schools in the intervention group and between six schools and seven schools in the control group administered classroom-based education other than RiPP across the 3 years. The types of programs implemented included gang resistance programs, character education programs, and individual presentations that were not part of a curriculum (most often, speakers, a video, or a lesson) focused on specific topics such as bullying, harassment, and dating violence.

3.1   Description of the Intervention

3.1.1 Overview

The comprehensive violence prevention program used for this project is designed for students enrolled in grades 6 through 8 in middle and junior high schools and is applicable to children from all socioeconomic, racial/ethnic, and cultural backgrounds. It was delivered in

intervention schools over 3 years. The program addresses behavior management for all students in a school. The two programs that make up the intervention, however, are universal programs and, as such,[15] are not designed specifically for students who are already exhibiting serious violent behavior in school. However, program impacts were examined for a high-risk subgroup, as the effects of universal programs on high-risk students would likely be of interest to practitioners.

For the purpose of this study, the concepts and skills taught through each part of the program—that is, RiPP (Meyer and Northup 2002a, 2002b, 2006) and Best Behavior (Sprague and Golly 2005)—are designed to be complementary and mutually reinforcing. RiPP was delivered by classroom teachers, and Best Behavior was designed to clarify policies and reinforce desirable behaviors schoolwide and was implemented through a team of school staff and administrators. Collectively, the two programs were designed to do the following:

- develop norms and expectations for nonviolent conflict resolution and positive achievement;

- diminish stereotypes, beliefs, and attributions that support violence;

- increase skills for nonviolent conflict resolution and self-management; and

- increase abilities to use appropriate violence prevention strategies.

At the same time, the program aims to help school staff do the following:

- make data-based decisions about school rules and policies;

- develop and implement effective school rules;

- establish a positive reinforcement system;

- establish effective classroom management methods; and

- effectively publicize rules and reinforcement systems to staff and students.

The two programs in this study were each modified. Prior to this project, Best Behavior had most frequently been used in elementary schools. It was therefore necessary for the developers to revise their approach to make it more acceptable to a middle school population. In general, the modifications focused on language used and rewards that were likely to be acceptable to middle school versus elementary youth. These modifications were primarily emphasized during training (described in section 3.2) rather than in the Best Behavior program manual.

In preparation for participating in the evaluation, the RiPP developers also revised their curriculum so that the number of lessons per year was evenly distributed across the 3 years. This was done to address concerns about the competing demands on classroom time within a given school year. The original RiPP curriculum design consisted of 25 lessons in the first year and 12 lessons in each of the next 2 years. The RiPP developers modified the curriculum so that each year's program consisted of 16 lessons. This program design modification resulted in redistributing the RiPP lessons across the 3-year implementation so that, at the end of 3 years, practically the same number of sessions would be delivered to the schools (48 lessons vs. 49 lessons).

The revised Best Behavior whole-school approach and the revised RiPP curriculum were piloted in a middle school in North Carolina in the years preceding full-scale implementation. Specifically, the revised Best Behavior approach and the revised first-year RiPP curriculum were

---

[15] An additional component of Best Behavior, not included in this study, provides support systems for individual students.

piloted during the 2005–06 school year. The second-year RiPP curriculum was piloted in the same middle school in North Carolina during the 2006–07 school year, and the third-year RiPP curriculum was piloted during the 2007–08 school year. The implementation team worked very closely with the school management team and teachers during the pilot years to gather information about how well the revised program was working in middle schools. Based on the feedback from teachers and administrators and on observations, the language in RiPP and Best Behavior was altered slightly to make the program more appropriate for students in middle school, and language was added to RiPP teachers' manuals to help facilitate teacher-student discussions.

3.1.2 Responding in Peaceful and Positive Ways (RiPP)

RiPP (Meyer and Northup 2002a, 2002b; 2006) is designed to promote social competency, problem solving, and self-control through weekly or biweekly classroom sessions where students have the opportunity to rehearse the use of problem-solving and violence prevention skills through experiential learning strategies, small group activities, and repetition. RiPP teaches students how to delay their reaction time under stress and conflict, to calm down, to examine their thoughts critically and consider alternative explanations and reactions, to recognize nonviolent options, and to understand the benefits of nonviolent responses.

Theoretical Model

The rationale for RiPP derives from Jessor's (1985) health promotion model, which views aggressive behavior as a function of individual and environmental characteristics. Consistent with the research-based literature on violence prevention, this model views individual thoughts and emotions as playing a central role in setting the context for violence; RiPP teaches participants how to delay their reaction time under stress and conflict, to calm down, and to empathize. Cognitive scripts (e.g., Huessmann [1988]) and negative attributions (Weiner, Graham, and Chandler 1982) also set the stage for violent reactions. RiPP teaches adolescents to examine their thoughts critically and then consider alternative explanations and reactions, recognize nonviolent options, understand the benefits of nonviolent responses, and practice positive cognitive scripts.

The theoretical model for RiPP is based on social learning theory, which recognizes the importance of family, peer, and environmental influences in setting the context for responses to conflict. The rationale for RiPP also derives from Crick and Dodge's (1994) work in social information-processing. Consistent with Crick and Dodge's research findings, RiPP teaches the problem-solving model in a way that is consistent with social information processing. Skills are taught and social knowledge is expanded and enhanced to encourage positive responses. Youth are also provided opportunities to clarify their values and identify their personal goals.

All of RiPP is anchored by an understanding of adolescent development, and the program recognizes the importance of peer relationships and acceptance by one's peers and of changing adolescent cognitions and emotions. The theoretical model for RiPP also recognizes the importance of situational factors (Friedlander 1993) in teaching students to recognize and assess dangerous situations. RiPP emphasizes nonviolence as a positive alternative strategy for violent responses. The program utilizes ground rules to help foster a positive school climate (Remboldt 1998).

Curriculum Design

The RiPP curriculum consists of 16 lessons (each lasting 50 minutes) per year in grades 6 through 8. RiPP comprises three grade-specific curricula, RiPP-6 (the first-year curriculum) for grade 6, RiPP-7 (the second-year curriculum) for grade 7, and RiPP-8 (the third-year curriculum) for

grade 8. Because the first-year (i.e., 6th-grade) RiPP curriculum serves as the foundation of all 3 years of the program, and because each year of the program is designed to build on the previous year's curriculum, a decision was made to give the RiPP first-year curriculum to all three grade levels in the first year of the project. The first-year curriculum can be viewed as an introduction to the 3-year program; as such, the content was appropriate to 7th- and 8th-graders as well as 6th-graders. In the second year of the project, 6th-graders received the RiPP first-year curriculum, while the 7th- and 8th-graders both received the RiPP second-year curriculum. In the third year of the project, each grade received its grade-specific curriculum. Thus, by the end of the 3-year project, the study sample (6th-graders in the first year of the project) received all 3 years of the curriculum.

Curriculum materials for each grade level consist of the following:

- a teaching manual with written lessons and instructions;

- individual student workbooks with activities keyed to specific lessons;

- posters of the RiPP problem-solving model to display in the classroom; and

- PowerPoint presentations for each lesson to assist teachers who wish to use them as part of their teaching.

Table 11 provides a list of the curriculum lessons and teaching techniques for RiPP-6, RiPP-7, and RiPP-8. Included with each lesson in the manual are instructions and recommendations for the teacher, including the session goal and objectives, a description of the underlying beliefs or theories and the program objectives addressed, materials and a discussion of the preparation needed, a definition of key concepts, and a discussion of other issues and considerations for the lesson. The lessons comprise a variety of activities and strategies, including team building, social-cognitive problem solving, repetition and mental rehearsal, small group work, role playing, rehearsal of specific social skills for preventing violence, and didactic learning. Most lessons contain between four and six of these activities and are estimated to take between 5 minutes and 15 minutes per activity. Each activity is scripted and tied to a specific objective. Most lessons make use of the student workbook as part of the activities.

Table 11.   Content and sequence of the 3-year RiPP curriculum lesson plans

Content	Interactive techniques
First-year RiPP curriculum	
Lesson 1. Introduction to Problem Solving. Lesson helps students get to know each other, introduces RiPP, and links RiPP to the Best Behavior whole-school rules. Lesson introduces SCIDDLE (Stop, Calm down, Identify the problem and your feelings about it, Decide among your options, Do it, Look back, and Evaluate) and RAID (Resolve, Avoid, Ignore, Defuse).	Games, experiential learning, and discussion
Lesson 2. Impact and Making RiPP Real. Increases students' awareness about the impact of violence on their lives.	Discussion
Lesson 3. First Response: Stop and Calm Down. Helps students understand physical reactions to anger and anxiety. Teaches a self-talk technique for calming down. Also teaches a deep-breathing technique.	Discussion, journaling, coaching, brainstorming
Lesson 4. Feelings and Information About Problems. Demonstrates how frustration can lead to strong feelings. Helps students become familiar with cues about their own and others' feelings.	Demonstration activity, small group work, journaling
Lesson 5. Identifying the Problem: Differences. Explores how differences can cause conflict. Introduces and explores perspective taking.	Demonstration activity, small group work
Lesson 6. Deciding Among Your Options: Solutions and Goals. Helps students understand how to judge when a solution to a problem is a good match for the goals they want to achieve.	Demonstration activity, discussion
Lesson 7. The Chain of Violence and the Web of Support. Helps students understand that violence leads to more violence. Provides strategies for resisting violence and suggests that each student has personal responsibility for contributing to the kind of community in which he or she wants to live.	Discussion, brainstorming
Lesson 8. Decide Option One: Avoid. Helps students understand that avoiding a person or situation is an option for managing danger. Students develop a personal safety plan.	Discussion, brainstorming, role play, small group work
Lesson 9. Decide Option Two: Ignore. Helps students understand that ignoring a person is an option for avoiding conflict. Also explores the self-talk strategy for deciding how to deal with a problem.	Discussion, brainstorming, role play, journaling
Lesson 10. Decide Option Three: When to Defuse. Explores how "expecting the worst" and "taking the bait" set a person up for a fight. Introduces the defuse strategy.	Discussion, small group work, role play, demonstration
Lesson 11. Decide Option Three: How to Defuse. Exposes students to a variety of techniques for defusing conflict situations. Includes opportunities to practice.	Discussion, small group work, role play
Lesson 12. Decide Option Four: Resolve. Exposes students to resolve techniques.	Discussion, journaling
Lesson 13. Doing It: Role Playing Resolve. Gives students the opportunity to practice resolve strategies through role play.	Discussion, small group work, role play
Lesson 14. Look Back and Evaluate One: Who Is Responsible? Helps students begin to evaluate factors responsible for conflicts. Helps students understand the powerful role of bystanders.	Discussion, small group work, role play

continued

Table 11.   Content and sequence of the 3-year RiPP curriculum lesson plans—**Continued**

Content	Interactive techniques
First-year RiPP curriculum—Continued	
Lesson 15. Doing It Four: Role Playing SCIDDLE. Provides an opportunity to practice applying SCIDDLE in different contexts.	Discussion, small group work, role play
Lesson 16. RiPP Wrap-up. Celebrates completion of RiPP year one. Invites students to sign a nonviolence pledge.	Discussion, journaling
Second-year RiPP curriculum	
Lesson 1. Introduction and Review. Teacher introduces the lessons for the year and develops ground rules. SCIDDLE and RAID are reviewed.	Discussion, games
Lesson 2. Review Stop and Calm Down. The goals of this lesson are to reinforce Stop and Calm Down. Introduces short-term and long-term relaxation techniques.	Discussion, demonstration, practice
Lesson 3. Review Identifying Feelings, Avoid and Ignore. Review the Identify the Problem and Feelings step of SCIDDLE. Also review Avoid and Ignore of RAID as appropriate responses.	Discussion, small groups, journaling
Lesson 4. Being a Positive Bystander. This lesson describes Diffuse as a strategy bystanders can use and the role they play.	Discussion, role play, journaling
Lesson 5. More Ways to Get Along. Helps students understand their role in making school a safe and fun place to learn.	Discussion, brainstorming, small groups
Lesson 6. The Importance of Friendship. This lesson takes a look at goals related to friendship and the importance of maintaining them. It is also a review for working in small groups.	Discussion, small groups, role play, journaling
Lesson 7. Respect for Others and for Yourself. The goal of this lesson is to get students to think about respect and disrespect and how using SCIDDLE and RAID is a way to show respect to themselves.	Discussion, small groups, demonstration
Lesson 8. Friendships, Respect, and Conflict. Goals of this lesson are to discuss how conflict works in friendships and making healthy choices when facing disagreements.	Discussion, brainstorming, small groups, demonstration, journaling
Lesson 9. Using Rules for Resolve in Friendships. Review and understand RULES for Resolve and how it can reduce conflicts in friendships.	Discussion, games
Lesson 10. Listening to Others. Increase students' awareness and value of listening.	Demonstration, role play, small groups, discussion
Lesson 11. Speaking So Others Can Listen. Lesson helps students use nonconfrontational language and the value of I-Statements over You-Statements.	Discussion, small groups, practice
Lesson 12. Looking Back: Did I Resolve? Gives students opportunity to practice the Rules for Resolve.	Discussion, small groups, role play
Lesson 13. Continuing With Resolve. Gives students practice in resolving conflicts by sharing role plays.	Discussion, role play
Lesson 14. Similarities and Differences Between Boys and Girls. Reinforces SCIDDLE and RAID in the real word. Helps students respect the difference between boys and girls.	Discussion, Where do I stand? activity, journaling

continued

Table 11. Content and sequence of the 3-year RiPP curriculum lesson plans—**Continued**

Content	Interactive techniques
Second-year RiPP curriculum—Continued	
Lesson 15. Ways to Resolve in the Family. The goal of this lesson is allow students to think of families and why parents set certain rules.	Discussion, role play
Lesson 16. Closure and Commitments. Lesson emphasizes what students have learned. Students make a commitment for nonviolence and individual pledges for nonviolence.	Discussion, small groups, pledge
Third-year RiPP curriculum	
Lesson 1. Responding in Peaceful and Positive Ways: Now and in the Future. Lesson reviews 6th- and 7th-grade programs and introduces 8th-grade program.	Discussion, game
Lesson 2. Attitude Is Always a Choice. Lesson helps students understand how attitudes can influence a situation and how to use Responding in Peaceful and Positive Ways using SCIDDLE.	Discussion, small groups
Lesson 3. Everyday Challenges, Opportunities, and Difficulties. Lesson explores challenges students face as they get older and how to maintain balance to prepare for expected difficulties.	Discussion, brainstorming
Lesson 4. Envisioning the Future. Students envision their future in 25 years through expressive writing.	Discussion, coaching
Lesson 5. Stereotypes and Your Future. Recalls lessons learned in previous activities and helps students identify stereotypes.	Discussion, small groups, journaling
Lesson 6. Gaining Experiences and Strategies: Power Now and in the Future. Explores experience and strategies of doing nothing versus taking positive risks.	Discussion
Lesson 7. Looking at SCIDDLE in a New Way. Learn SCIDDLE is an example of a positive cognitive script.	Discussion, role play, journaling
Lesson 8. Emotion-Focused Coping. Students learn effective ways to cope with a problem.	Discussion
Lesson 9. Expanding RAID and Forgiveness. Students express their feelings and practice reframing though expressive writing. Explores the benefits of forgiveness.	Discussion, journaling
Lesson 10. Forgiveness for Real. Lesson applies the five-step forgiveness model.	Discussion, small groups, journaling
Lesson 11. Creating the Future Work I Want to Do. Introduces use of personal strengths and education to prepare for future work as an adult.	Discussion
Lesson 12. Safety Nets and Lifesavers. Interprets the meaning of Feared and Desired Selves and how to express them through various exercises.	Discussion, brainstorming, journaling
Lesson 13. After-School Time Is a Gift and Resource. Lesson helps students identify beneficial experiences during after-school hours.	Discussion

continued

Table 11.   Content and sequence of the 3-year RiPP curriculum lesson plans—Continued

Content	Interactive techniques
Third-year RiPP curriculum—Continued	
Lesson 14. Rights and Responsibilities as a Citizen and Community Member. Students consider how they want their community to be and their rights and responsibilities.	Discussion, journaling
Lesson 15. A Grounded Future. This lesson reviews the concept of being grounded. Also prepares students for transition to high school.	Discussion
Lesson 16. Moving On Together. This lesson provides closure to RiPP. Students write plans to stay on track.	Discussion, journaling

During RiPP, students are instructed in the use of a social-cognitive problem-solving model, SCIDDLE (Stop, Calm down, Identify the problem and your feelings about it, Decide among your options, Do it, Look back, and Evaluate), and RAID (a specific set of options that include Resolve, Avoid, Ignore, and Defuse). RiPP is designed so that students will have the opportunity to rehearse the use of the problem-solving model and violence prevention skills through experiential learning strategies, small group activities, and behavioral repetition. Through repeated use of this problem-solving model, increased awareness of the nonviolent options, and opportunities for reflection and practice, RiPP aims to help students learn how to choose the prosocial strategy most likely to provide the desired short- and long-term outcomes in a given situation.

Role of the Teacher

The RiPP curriculum is designed to be taught by trained classroom teachers. RiPP can be taught in virtually any subject area, but RiPP teachers typically are in an academic subject such as social studies, health, or science. The role of the teacher during RiPP is to serve as a facilitator and coach. For RiPP lessons to be effective, RiPP teachers must establish a positive, respectful classroom environment where students feel safe participating in in-depth discussions about personal experiences and different perspectives and responses to situations.

The RiPP teacher facilitates discussion by asking open-ended questions (provided in the curriculum) and leading students to make desired points that increase their understanding of the range of positive options that exist in conflict situations. The RiPP teacher also sets up classroom demonstrations and other experiential activities that help students recognize and explore their perceptions about violence. RiPP teachers are asked to make material relevant by drawing on local and national news stories and by telling stories about their own lives (such as about minor conflicts they have experienced) to introduce and expand topics for discussion. RiPP teachers model nonviolent responses. It is important to high-quality implementation that teachers be supportive of the key components and philosophy of the RiPP program.

Teaching Strategies

RiPP lessons are designed to be highly interactive in order to demonstrate and model social-cognitive problem solving and to offer students opportunities to develop and master new skills through repetition, mental rehearsal, and role playing.

Teachers are expected to lead lively discussions and to actively engage students during classroom activities. Teachers are called on to lead brainstorming sessions and to guide students

through small group activities that promote team building. RiPP also relies heavily on role plays (or skill rehearsal) to demonstrate new skills and to provide opportunities for practice and mastery.

The goal of brainstorming is to actively engage students as they create a list of relevant answers to whatever prompt they have been given. Through brainstorming, teachers are supposed to learn what students already know. The teacher may also learn what students erroneously think of as true. Brainstorming is designed to be fast paced and to be used in conjunction with another interactive method, such as discussion or as part of role play and rehearsal (Bosworth and Sailes 1993).

The goal of small groups/cooperative learning is to encourage students to increase learning by allowing students to gain from each others' efforts, talents, and strengths. Cooperative learning is designed to build teamwork and cooperation by putting students in small groups and assigning them tasks to which all contribute something of their own. The quality of work is judged at the team (not the individual) level with the goal of creating a positive pressure to perform.

Role plays, also called *skill rehearsal*, are designed to give students opportunities to practice new skills. Role plays can serve different functions. Role play and rehearsal are important steps in mastering new skills. The goal of the role playing is to help students gain a mastery of technique.

3.1.3 Best Behavior

Best Behavior (Sprague and Golly 2005) provides a standardized staff development program aimed at improving school and classroom discipline and decreasing school violence. Best Behavior is designed to develop and administer effective school rules and discipline policies at both schoolwide and classroom levels to decrease school violence and antisocial behavior. It includes whole-school, common area, classroom, and individual student interventions.[16] It begins by creating a school management team, setting up meetings, developing positive rules and expectations, developing lesson plans to teach those rules and expectations, and developing a positive reinforcement system. Over time, the system encourages the team to use data to drive decisions and to develop a team and strategies for dealing with high-risk youth who may not be reached by the positive behavior support system.

Theoretical Model

Best Behavior is based on the Positive Behavior Support (PBS) approach (Sugai and Horner 1994; Sprague, Sugai, and Walker 1998) developed at the University of Oregon and the National Center on Positive Behavioral Interventions and Supports (http://www.pbis.org/). According to the program developers, the goal of the Best Behavior program is to facilitate the academic achievement and healthy social development of children and youth in a safe environment conducive to learning.

Best Behavior was developed using principles from education, public health, psychology, and criminology. It is based on a multiple-systems approach to address the problems posed by antisocial students, including challenging and violent forms of student behavior. The key concepts underlying the development of Best Behavior include the following: (1) clear definitions of appropriate, positive behavioral expectations provided for students and staff members; (2) clear definitions of problem behaviors and their consequences for students and staff members; (3) regularly scheduled instruction and assistance to enable students to acquire the necessary skills that will ensure desired behavior change; (4) incentives and motivational systems provided to encourage students to behave

---

[16] The individual student intervention component was not part of the program implemented for this study.

appropriately; (5) school staff committed to staying with the intervention over the long term to monitor, support, coach, debrief, and provide booster lessons for students, as necessary, to maintain the achieved gains; (6) staff who receive training, feedback, and coaching about effective implementation of the intervention; and (7) established systems for measuring and monitoring the intervention's effectiveness that are carried out regularly and shared with implementers to improve implementation and maintain motivation to stay with the intervention.

The key concept of clarifying and teaching behavioral expectations is based on the use of Direct Instruction (DI) (Engelmann 1968), an instructional method, in combination with modeling and role-playing concepts derived from social cognitive theory (Bandura 1986). Reinforcement of appropriate behaviors, another of the program's key concepts, is based on research indicating that providing positive reinforcement for prosocial behavior is essential for creating lasting changes in behavior (Patterson 1982).

Role of the Principal

Leadership by the principal is crucial to successfully launching and sustaining Best Behavior. For Best Behavior to succeed, it must be viewed as a priority in the school. Principals have many competing priorities and demands on their time, and it is reasonable for a principal to delegate primary responsibility for the day-to-day operations of the program to an assistant principal, guidance counselor, or other appropriate faculty member. However, it is critical for the principal to communicate enthusiasm for and commitment to the Best Behavior program and to recognize Best Behavior as an organizing philosophy in the school. While the principal may not attend every Best Behavior team meeting, the principal's presence in at least some of the meetings also communicates ongoing support and commitment for the program. Even if principals should delegate responsibility for the program to a subordinate, principals are expected to be actively involved in committing to the program initially, in forming the Best Behavior school management team, and in communicating ongoing support for the program at faculty meetings and in other contexts.

Role of the School Management Team

Best Behavior is implemented by means of a Best Behavior school management team which, according to the program guidelines, should include a building-level administrator and representatives from each major stakeholder group, for example, representative teachers from each grade, guidance counselors or school psychologists, and administrators. Usually, the school management team is formed by the school principal in consultation with the Best Behavior training and support staff. The Best Behavior team identifies the prevention of youth violence and the promotion of social and emotional competence as priorities for the school community. The management team is expected to meet monthly.

The Best Behavior management team is expected to develop a systematic approach to developing schoolwide positive behavior supports. Over 3 years, the school management team is expected to have an increasing, expanded role in implementing the program throughout the school. This includes four broad sets of activities to be completed over 3 years. First, the team is to conduct a schoolwide needs assessment to identify reasonable goals. Needs assessments are to be repeated annually. Second, the team is to define rules and expectations, with general rules (e.g., be safe, be respectful, be responsible) supported with expectations for all settings within the school environment. Rules and expectations are to be taught on a regular basis by all teachers and staff. Third, the team is to develop and support a positive behavior reinforcement system in which students are to be given token rewards for obeying rules and meeting expectations. Finally, the team

is to develop a data-based decisionmaking process for identifying and addressing the needs of high-risk students. Based on the Best Behavior guidelines (Sprague and Golly 2005), the first three activities should be implemented in the first year, while the fourth activity should be in place by the second year.

Best Behavior Program Steps

The program provides a series of steps that schools follow to implement the program along a recommended 3-year timeline. The program recommends that all steps be initiated at some point in the first year, with steps 10 and 11 initiated in years two and three, respectively. However, schools are allowed latitude in how quickly they complete these steps, reflecting each site's needs and each team's capacity. The steps are the following:

Year one

1. **Form a school management team.** A representative building leadership team is formed to guide program implementation. Monthly team meetings are scheduled. An action plan with clear goals and objectives, initiated at the training workshop, is developed.

2. **Conduct a schoolwide needs assessment to guide priorities for the discipline system.** All adults in the school complete a self-assessment survey to identify areas for improvement or development and to set goals and priorities. Needs assessments are to be repeated annually.

3. **Define three to five general behavior rules.** The school management team selects three to five positively stated schoolwide behavior rules (e.g., be safe, be respectful, be responsible). Rules are posted or made visible in all school settings (e.g., in hallways, classrooms, cafeteria, gym). The rules should be taught and reinforced on a regular basis by all teachers and staff.

4. **Define positive behavior expectations for each school setting (e.g., cafeteria, gym).** Positive behavior expectations state exactly what is expected from students within each school setting (e.g., what do safe, respectful, responsible look like in the cafeteria, gym, restrooms).

5. **Develop and implement token reward systems for positive student behavior.** The school management team develops a formal system for reinforcing students through token economies (e.g., tokens that can be redeemed at the student store for school items or for snacks, celebrations, or other prizes). The team is to constantly update a continuum of positive consequences, such as being first in line, leaving class 2 minutes early, free time, or computer time.

6. **Define problem behaviors and consequences.** Problem behaviors and corrective consequences (e.g., verbal correction, loss of privileges, extra work, parent contact) are clearly defined and explained to all students. Teachers judiciously use positive and corrective consequences to make clear to students the boundaries of acceptable and unacceptable behavior.

7. **Develop lesson plans for teaching expectations in classrooms.** Lesson plans are developed for teaching about behavioral expectations in all school settings.

8. **School staff are trained to teach behavioral expectations, use classroom management techniques, and respond to problem behaviors.** School staff are trained to model appropriate behaviors and to provide students with step-by-step instructions on how to behave in desired ways in different contexts. Teacher training on classroom management strategies also occurs. Also, the team is to instruct teachers in how to recognize signs of escalating behavior and provide guidelines for intervening early in the chain of events.

9. **School staff convey behavioral expectations to students.** Teachers review or teach expected behavior on a regular basis throughout the school year, using the lesson plans. Positive reinforcement is used with students at all times and by all adults for following the expectations.

Year two

10. **Collect systematic data and use to modify the program.** Data such as discipline referrals are regularly collected, summarized, and reviewed by the school management team. Such data may be used to identify problem areas in the school, determine if the discipline and reward system is working, and identify problem students.

Year three

11. **Develop individualized support systems to address needs of students with ongoing problem behaviors.** Best Behavior provides material designed to help the school management team address the needs of students with chronic behavioral problems through individualized programming.

Best Behavior provides detailed guidelines for achieving each component listed above. For example, rules developed by the school management team as part of Best Behavior should be positively stated (e.g., "Walk in the halls" rather than "Don't run"). In addition, rules should be posted in hallways, classrooms, the school handbook, and so forth. Finally, the program recommends that rules be taught directly to all students and reviewed 10 to 20 times a year. Best Behavior also provides training on typical program implementation to enable each team to develop, adapt, and implement the various components specific to the needs of its individual school. For example, a sample lesson plan for teaching "be respectful" is provided, but teams are encouraged to adapt this prototype lesson to fit the language and unique cultural features of their school (e.g., use of local slang; modeling behavior using the school's mascot).

3.2   Training, Technical Assistance, and Support

3.2.1 RiPP Training

In year one, an initial 2-day training was designed so that participating 6th-, 7th- and 8th-grade teachers were introduced to the program's theoretical framework and key concepts in aggressive and violent behavior, nonviolence options, and adolescent development. The problem-solving skills central to RiPP were introduced and modeled. Participants also had the opportunity to practice interactive teaching and reflect on various concepts through small group activities. Rather than learning to teach every lesson, teachers were introduced to the curriculum structure, lesson objectives, and activities while trainers used some of the lessons to illustrate key concepts and teaching strategies. The training workshop provided opportunities for discussion of practical issues and potential challenges with implementing a violence prevention curriculum.

In addition, the training workshop modeled how the RiPP curriculum should look in the classroom. The trainers relied heavily on interactive teaching techniques relative to didactic instruction. The interactive techniques trainers used were the same ones the teachers are expected to use in the classroom. In addition, the trainers modeled the kinds of examples that teachers are asked to use to make RiPP relevant to students in the classroom.

RiPP teachers were trained in the summer or fall prior to the beginning of the first school year and then in August or September prior to the beginning of the second and third years. Separate RiPP training workshops were held for each district (or several small districts at a time, if they were located in close proximity to one another). Each workshop in the first year lasted for 2 days and was led by two trainers working together. Participating 6th-, 7th-, and 8th-grade teachers attended the first-year training. In the second year, a 1-day training was held using a similar format and approach to prepare 7th- and 8th-grade teachers to implement the second-year RiPP curriculum. A single day was acceptable since the majority of 7th- and 8th-grade teachers had received the initial RiPP training in the prior year as part of the first-year implementation. In the third year, a 1-day training prepared 8th-grade teachers to implement the RiPP third-year curriculum since, again, most of the 8th-grade teachers had received training in the previous 2 years, as part of first- and second-year implementations. When teachers were new to the program in the second and third years, they received a general orientation as needed from the liaison and were provided with the videos from the first-year training for an introduction to the general concepts and methods of RiPP. The trainers for the workshops in all 3 years were the program developers or others who had direct experience with implementing RiPP in the classroom; all trainers had been involved for at least 5 years in training activities for middle school teachers. The need for alternative training of some teachers arose as a result of teachers being unavailable during the initial group training, staff changes, or staff not being identified at the time of the training. As a result, DVDs of a live training session were produced in the first year and distributed to all liaisons.

Overall, 181 teachers were trained to administer the RiPP curriculum in the first year, in the 18 schools that participated in all 3 years of the study. Eighty-two percent was the site-level average of teachers trained during the developer-led sessions, while 18 percent was the site-level average of teachers who were trained through alternative methods, such as reviewing a videotape of the training, having one-on-one sessions with the liaison or a trained teacher, and coteaching a sample of lessons (see table 12). In the second year, a total of 201 teachers were trained to administer the second-year RiPP curriculum because four schools opted in the second year to implement RiPP using the entire faculty rather than a core group of teachers in a select subject. This meant that all the teachers in those schools needed to be trained. In the second year, 79 percent of teachers were trained in developer-led workshops, 17 percent were trained using alternate methods, and 4 percent had been trained in the first year and did not receive follow-up training in the second year.

In the third year, a total of 146 teachers were trained to administer the RiPP curriculum. Fifty-seven percent were trained in developer-led workshops, 32 percent were trained using alternate methods, and 12 percent were trained in the previous year but received no follow-up training in that year.

Table 12. **Average site-level percentage of designated staff trained in the intervention over 3 years**

	Portion of the intervention	
Type of training	Curriculum[1]	Whole-school[2]
First year		
Group training led by developer	82	89
Alternative training[3]	18	11
Second year		
Group training led by developer	79	79
Alternative training	17	21
No follow-up training	4	0
Third year		
Group training led by developer	57	73
Alternative training	32	27
No follow-up training	12	0

[1] For the curriculum, *designated staff* includes teachers assigned to teach RiPP. This included 6th-, 7th-, and 8th-grade teachers in year one; 7th- and 8th-grade teachers in year two; and 8th-grade teachers in year three.

[2] For the whole-school portion, *designated staff* includes all members of the school management teams.

[3] Alternative training for the curriculum included review of a videotaped training session, one-on-one sessions with the liaison or a trained teacher, and coteaching of sample lessons with the liaison. Alternative training for the whole-school portion of the intervention included orienting by the team, the liaison, or individual team members; review of videos on Best Behavior approaches; and learning through participation in subcommittees.

NOTE: Data are based on 18 intervention schools that participated for 3 years.

SOURCE: Implementation team's attendance records for training.

3.2.2 Best Behavior Training

The goal of the 2-day training in the first year was to provide school management teams with an understanding of the philosophical framework underlying the positive behavior approach and to guide them in developing policies and procedures for managing student behavior schoolwide, in classrooms, and on an individual basis. The training was led by the program developer. School management teams were introduced to the key components of Best Behavior, including defining schoolwide rules and expectations, teaching behavioral expectations, designing a schoolwide recognition and reward system, creating a positive culture in the school, using office referral data to monitor behaviors, learning the foundations of classroom management, and understanding group and individual behaviors. Videos were used to illustrate methods for systematic supervision and monitoring of common areas in the school and to show how to respond to escalating behavior and to defuse anger and violence. Teams were also provided with handouts illustrating examples of rules and reward systems developed in other schools and were given time to work in small groups to begin developing these for their own school.

Best Behavior school management teams were trained between April and September 2006 in preparation for the first year. Separate Best Behavior training workshops were held for each district (or several small districts at a time, if they were located in close proximity to one another). Each workshop in the first year lasted for 2 days and was led by one of two program developers. In the second and third years, the trainings lasted for 1 day and were designed as much as possible to meet teams at the points to which they had advanced in the process of launching Best Behavior and to

move them forward. Because of changes in principals and other staff, it was often necessary during trainings in the second and third years to take a step back, however, and help teams reorganize, as well as to provide new team members with an orientation to the program. The second- and third-year trainings were conducted between April and October of each year, in 2007 and 2008.

Across the 18 schools that remained in the study after 3 years, a total of 197 school staff were trained in the skills required to form a functioning school management team during the first year. Of these, an average of 89 percent of staff at each site received training in a group setting, while the residual staff were oriented by the school management team, a member of the implementation team, or another management team member; reviewed videos of Best Behavior approaches (Smith and Sprague 2006; Colvin 2004); and were trained experientially by participating in subcommittees. In the second and third years, 79 percent and 73 percent, respectively, received developer-led training, as shown in table 12. For Best Behavior, the need for alternative training arose as a result of staff being unavailable during the initial group training, staff changes, or staff not being identified at the time of the training. Program developers provided schools with PowerPoint presentations and the Best Behavior manual (Sprague and Golly 2005) and also offered technical assistance.

3.2.3 Technical Assistance, Teacher Support, and the Role of the Implementation Liaisons

While both RiPP and Best Behavior are implemented by school staff, in this project technical assistance was also made available throughout the implementation period by on-site implementation liaisons. Liaisons were expected to facilitate, coach, and monitor the progress and delivery of both of the programs. Liaisons hired for the project were a group of experienced former or current educators. All but two were retired teachers or principals. All but 2 of the 11 liaisons were recommended by the principal or district coordinator; 1 liaison was recommended by another liaison; and 1 liaison was recruited through an advertisement in the local newspaper. All but two of the liaisons were well known to the faculty at those schools, and six of them frequently served as mentors for new teachers in the district. The liaisons were a highly stable group, with only 2 of the 11 liaisons changing during the course of the 3-year project.

Liaisons attended the RiPP training alongside the teachers from their schools. During RiPP lessons, liaisons observed each teacher several times during the school year. Following their observations, the liaisons offered feedback and suggestions and any support that the teachers might request. Where it was clear that teachers needed additional support, liaisons modeled the delivery of the curriculum by coteaching it with them. Many of the liaisons regularly attended teacher team meetings to discuss RiPP and to plan for its implementation. When new teachers became involved in the curriculum after the initial training, the liaison used videos of the RiPP workshop, along with one-on-one meetings, to prepare the new teachers to implement RiPP. The liaisons focused primarily on teachers who were new to the program or clearly were struggling with it.

With regard to Best Behavior, in the first year of the program the liaisons attended the Best Behavior training with the staff from their assigned schools and supported the school management teams to the extent that they were needed, especially in the following key areas:

<u>Meeting facilitation.</u> Liaisons ensured that school Best Behavior management team meetings occurred. In many cases, the liaison acted as an assistant to the management team leader—organizing meetings, creating agendas, and typing up minutes—and kept track of and facilitated team progress in completing the steps of the Best Behavior program. In addition

to attending the monthly meetings with the team, the liaisons usually had weekly contact with key Best Behavior staff.

Reinforcement system. Liaisons aided the teams in designing and implementing the schoolwide reinforcement system. For example, liaisons had primary responsibility for compiling data on the number of tokens distributed to students each month, giving out rewards to students, and participating in the reward celebrations (e.g., attending movies or dances, accompanying groups of students to off-campus lunches).

Educating faculty and staff about Best Behavior. At times, the liaison served as the spokesperson for the Best Behavior program, especially when educating the rest of the faculty. Liaisons also took an active role in supporting school rules and lesson plans, such as helping schools develop strategies for delivering lessons on the public address system.

While in the first year of the program the liaisons were to provide extensive support to help launch the Best Behavior program, generally serving in whatever capacity might be needed to support the team and their activities, in the second and third years of the project liaisons were to help ensure institutionalization of the program by gradually turning responsibility for the program over to team members themselves. Thus, liaisons were to play a more secondary role by the third year, although they were to continue to participate in meetings and monitor and support the reinforcement system and other aspects of the program.

3.3   Program Implementation

This section reports the findings for program implementation, assessed in intervention schools, according to the fidelity criteria outlined in section 2.7. Fidelity of implementation for RiPP is presented first, followed by fidelity of implementation for Best Behavior. The discussion is intended to be descriptive, as no statistical tests were conducted to test for significance of differences across years. The section begins by describing the contextual background for the intervention schools.

3.3.1 Context of Program Delivery

RiPP and Best Behavior were implemented in schools in which there were often other challenges. We identified two specific issues that site liaisons had documented in the monthly implementation progress reports as having resulted in challenges for program delivery: administrator turnover and failure to meet adequate yearly progress (AYP) on academic end-of-grade testing.

Administrator Turnover

The Best Behavior developers (Sprague and Golly 2005) underscore that a school principal's support and leadership are critical to having a successful Best Behavior team, and the absence or reluctant participation of the principal creates challenges for moving the program forward. Changes in principal leadership have the potential to disrupt and sometimes derail implementation of schoolwide programs like Best Behavior that rely on the principal for leadership and support. Over the course of the program's 3 years of implementation, there was a change of principal in approximately one-third (33 percent) of the schools in the first and third years and in 22 percent of schools in the second year. Altogether, 12 of the 18 schools changed principals at least once during the study; 3 schools changed principals twice. However, principal changes were perceived as being beneficial in some cases, according to the site liaisons. For example, they claimed that the new principal was more supportive than the outgoing principal.

Failure to Meet Adequate Yearly Progress

Eight of the 18 intervention schools were on academic probation and did not meet standards for AYP during the first year of the project, 7 intervention schools did not meet AYP in the second year, and 4 of the 18 intervention schools did not meet AYP in the third year, as documented on the liaisons' year-end implementation reports.[17] According to the monthly implementation progress reports as well as annual reports completed by site liaisons, administrators, and teachers in these schools expressed the need to focus on addressing academic performance and test scores, and several schools introduced new academic programs intended to improve test scores. Liaisons reported that this priority on academic performance seemed to result in less time and attention devoted to progress on Best Behavior and completion of all RiPP lessons by teachers.

3.3.2 RiPP Implementation

Table 13 displays the variety of classes selected by the schools for delivering the RiPP curriculum in each of the 3 years. Across the three years, RiPP was frequently taught as part of a core subject such as social studies or science (61 percent in year three). Whereas in the first year, other subjects chosen were health or physical education (44 percent), more schools opted to move RiPP out of these subjects and into language arts or extended learning, enrichment, or advisory periods in the last 2 years. While enrichment or advisory periods provided schools with greater flexibility for delivering the curriculum, these periods also were of a shorter duration than a class period, which meant that one RiPP lesson had to be taught over multiple periods. According to the program guidelines, RiPP lessons are designed to fit in a class period so as to allow sufficient time for opening (review) and ending (closure) activities as well as for the use of interactive teaching methods (e.g., small group discussion and role plays) that require extended time. Six out of 18 schools combined several subjects to deliver the program to all students; for example, RiPP was sometimes taught in health class during the first semester and in physical education in the second semester.

To ensure that all students in the school received RiPP, schools selected a delivery schedule that allowed the program to be repeated during the year for new groups of students. Each RiPP teacher taught multiple classes of students throughout the day. Regardless of whether the curriculum was taught once or twice per week, approximately one-half of the students in all the intervention schools received the program by the end of the first semester, and the remainder received the program by the end of the school year, in each year of the study.

---

[17] Among control schools, 11 of the 18 schools did not meet Adequate Yearly Progress (AYP) during the first year of the project, 7 schools did not meet AYP in the second year, and 5 schools did not meet AYP in the third year.

Table 13.   Placement of RiPP curriculum in classes—3 years

Class	Percent of intervention schools		
	Year one	Year two	Year three
Health or physical education	44	17	28
Social studies or science	67	50	61
Language arts; allied arts (e.g., art; technology education); or extended learning, enrichment or advisory period	28	67	61

NOTE: Percentages total more than 100 percent because schools could implement in more than one type of class.
Data are based on 18 intervention schools that participated for 3 years.
SOURCE: RiPP implementation records; liaison year-end reports.

Exposure to the RiPP Program Curriculum

One criterion for faithful RiPP implementation is delivery of the entire curriculum each year. Table 14 shows the extent to which intervention schools delivered the full program (all 16 RiPP lessons) in each year of implementation. These data were compiled from records kept by the RiPP teachers and collected by project staff after every fourth lesson during implementation. Delivery of the full program in a school varied to the extent that individual classes did not complete all lessons; however, most intervention schools delivered all RiPP lessons to a majority of the assigned classrooms. Specifically, in each of the 3 years, between 11 and 13 of the study's 18 intervention schools (or between 61 percent and 72 percent) delivered all 16 lessons to all classrooms.

Table 14.   Exposure to the RiPP curriculum—3 years

Percent of classrooms completing all 16 lessons	Number (percent) of intervention schools		
	Year one	Year two	Year three
Less than 100	5 (28)	7 (39)	6 (33)
100	13 (72)	11 (61)	12 (67)

NOTE: Data are based on 18 intervention schools that participated for 3 years.
SOURCE: RiPP implementation records.

Fidelity of RiPP Delivery in the Classroom

Data on RiPP delivery in the classroom were collected by the evaluation team during the site visits, which were conducted in the spring of each implementation year. During the site visit to each school, RiPP delivery was observed in three different classrooms under three different instructors. During the site visits, several observational measures of fidelity were collected (see appendix D) that concern the extent to which teachers' delivery of RiPP was aligned with the following program requirements: (1) adherence to scripted lesson plans; (2) adherence to prescribed teaching techniques; and (3) student responsiveness during the session. Fidelity alignment for adherence to prescribed teaching methods was coded as follows: *well aligned* was indicated by all three classes observed in the intervention school displaying the techniques; *moderately aligned* was indicated by two of the three classes displaying the techniques; and *poorly aligned* was indicated by one of the three classes displaying the techniques. The measures for adherence to scripted lesson plans and for student responsiveness were scored from one (not at all) to four (always/very much) for each of the three classes observed. For each intervention school, scores were averaged across the three classroom observations. Average

scores greater than or equal to three were coded as well aligned; scores between two and three were coded as moderately aligned; and scores below two were coded as poorly aligned.

Results from the evaluation team's classroom observation are displayed in table 15 by the percentage of intervention schools where the degree of alignment was observed. In each of the 3 years of program implementation, the curriculum was not fully delivered with fidelity, as indicated by observational data collected on the three measures. Specifically, 67 percent of schools were rated as well aligned with respect to teachers following lesson plans in year one; 56 percent in year two and 44 percent in year three were similarly rated. With regard to teachers using the correct teaching techniques for each RiPP lesson, 56 percent of schools in year one, 44 percent of schools in year two, and 56 percent of schools in year three were well aligned on this measure. Eighty-nine percent of schools in year one and 69 percent in year three were found to be well aligned on the measure of student responsiveness during lesson activities.

Table 15.   **RiPP curriculum fidelity assessments, based on evaluation team observations— 3 years**

		Percent of schools		
RiPP fidelity criteria	Alignment with program requirements[1]	Year one	Year two	Year three
Adherence to the scripted lesson plan	Well aligned	67	56	44
	Moderately aligned	33	44	56
	Poorly aligned	0	0	0
Adherence to the prescribed teaching techniques	Well aligned	56	44	56
	Moderately/Poorly aligned	44	56	44
Student responsiveness	Well aligned	89	—	67
	Moderately aligned	11	—	33
	Poorly aligned	0	0	0

— Not available; value suppressed to protect respondent confidentiality.
[1] Fidelity scores are based on three observations per school.
NOTE: Three RiPP teachers from each school, one from each grade level, were randomly selected to be observed during a RiPP session. Rating scores for a school were averaged or summed across the three observations. Data are based on 18 intervention schools that participated for 3 years.
SOURCE: Classroom observations.

Challenges Implementing RiPP

To better understand the context and fidelity of RiPP implementation, data were collected from three randomly selected RiPP teachers at each intervention school, one from each grade level, through interviews conducted at midyear by the evaluation team. Teachers were asked about the appropriateness of RiPP for their students and how well students responded to the lessons and strategies. Among the issues mentioned by teachers were that students had trouble understanding or applying some of the concepts presented in the lessons (e.g., the difference between the "assist" role and "tattle-tailing" or applying techniques such as "ignore"), and that students had difficulties with certain strategies (e.g., journal writing and role playing). Other challenges pointed to students' lack of interest in certain activities, such as writing, and students having trouble staying on task, particularly during role playing and small group activities.

Teachers were also asked about specific aspects of their experience with delivering RiPP. Interviewed teachers reported a lack of class time to cover all aspects of the lessons (43 percent).

When asked about challenges they may have faced in using specific RiPP teaching techniques or approaches,[18] 88 percent of interviewed teachers mentioned facing challenges with at least one specific RiPP technique or approach, and 27 percent indicated they encountered challenges with three or more of the five techniques. Teachers faced the most challenges with using role play (70 percent), but they also faced challenges with small groups (54 percent) and teaching self-talk to students (42 percent).

Interviewed teachers were asked about specific ways in which they had dealt with these issues and challenges. Over 82 percent of teachers in year three left some techniques out of the lessons, and 89 percent either changed the content of the RiPP lessons or changed the way RiPP was taught. For example, teachers mentioned changing the content by adding more relevant examples and adapting the lesson to fit students' needs, such as incorporating videos so that classes with predominantly Hispanic students could see that other cultures have similar problems as they do. Teachers mentioned changing the ways in which they taught RiPP due to time constraints, such as shortening the lesson. Teachers also mentioned modifying or leaving out certain techniques such as replacing role play with class discussion due to time constraints or because students "got carried away."

3.3.3 Best Behavior Implementation

School Management Team Composition and Continuity

School management teams ranged in size between 4 members and 19 members each year of the program, with an average of 10 members. Table 16 shows the site-level averages with regard to the composition of the Best Behavior school management teams over the 3 years of implementation. On average, non-RiPP teachers constituted 42 percent to 47 percent of teams each year, while RiPP teachers constituted 12 percent to 13 percent (to facilitate alignment between RiPP and Best Behavior); administrators accounted for 21 percent in the first year and 17 percent by the third year; counselors made up 12 percent to 13 percent each year; and others constituted 11 percent or less over the 3 years. According to the Best Behavior guidelines (Sprague and Golly 2005), school management teams should include a building-level administrator and representatives from each major stakeholder group (e.g., grade-level teachers, guidance counselors or school psychologists, and administrators). For this study, which implemented a hybrid intervention, teams also included representatives from the RiPP teaching staff.

---

[18] RiPP teachers were to use the following teaching techniques and approaches: make RiPP real (tie it to students' daily lives); role plays; small group work, discussion, and brainstorming; encourage self-talk by students; use *Review* and *Closure* to begin and end sessions.

Table 16.   Best Behavior: Composition of school management teams—3 years

	Average percentage across schools		
Team members	Year one	Year two	Year three
Administrators	21	18	17
Non-RiPP teachers	43	42	47
RiPP teachers	14	23	20
Counselors	12	12	13
Other	11	5	3

NOTE: Teams ranged in size between 7 and 16 members in year one, between 4 and 18 members in year two, and between 6 and 19 members in year three. Data are based on 18 intervention schools that participated for 3 years. SOURCES: RiPP implementation records; liaison year-end reports.

Interviews conducted annually by the evaluation team with members of the school management team at each school indicate that team membership shifted over time. Among the 49 team members interviewed in year two and the 53 team members interviewed in year three, 81 percent in year two and 55 percent in year three had been involved with Best Behavior for the entire length of time since the study began.

Principal Support of Best Behavior Program

Principals' support and enthusiasm for the program and their leadership in promoting and sustaining the program are important aspects of the role of the principal for Best Behavior. Kam, Greenberg, and Walls (2003) note the importance of the support of a school's principal to the success of school-based prevention interventions.

Table 17 summarizes the level of principal support and commitment, as measured by various aspects of support. Principals were perceived as supportive (by the liaisons) in 83 percent of schools in year one, 78 percent of schools in year two, and 72 percent of schools in year three. Seventy-two percent of schools in year one, 61 percent in year two, and 50 percent in year three were rated by the liaisons as having principals that used their leadership to promote the program. Attendance by the principal at the Best Behavior training workshop was recorded in 78 percent of schools in the first year, 39 percent in the second year, and 28 percent in the third year. In the third year of implementation, only around one-half (56 percent) of the principals were regularly involved with the school management team; this level of involvement was recorded in 56 percent of schools in year one and 50 percent in year two. The observed levels of principal support in each year could, at least partially, be a factor of the principal turnover rates reported in section 3.3.1.

Table 17.   Principal support of Best Behavior—3 years

Type of principal support	Percent of schools		
	Year one	Year two	Year three
Principal was supportive (as perceived by the liaisons)	83	78	72
Principal used leadership to promote program (as perceived by the liaisons)	72	61	50
Principal attended Best Behavior workshop	78	39	28
Member of school management team and attended meetings on occasion	56	50	56

NOTE: Data are based on 18 intervention schools that participated for 3 years.
SOURCE: Liaison year-end report.

School Progress on Best Behavior

Frequency of School Management Team Meetings

One aspect of fidelity for the Best Behavior program is that the team at each school meets at least monthly. The progress of this recommended practice is shown in table 18. According to the monthly implementation progress reports completed by the site liaisons, over the approximate 9 months available for meetings, the teams met an average of five times in the first year and an average of eight times each in the second and third years. Across schools, the frequency of meetings ranged from a low of 2 meetings to highs of 28 meetings and 19 meetings in years two and three, respectively. One or two schools each year held the highest number of meetings as a result of meeting weekly during periods of increased program activity. Across the 3 years, schools met an average of 21 times; the frequency of meetings ranged between 6 meetings and 60 meetings. Based on annual interviews conducted by the evaluation team, about one-half of the team members in years two and three indicated they had attended all of the team meetings.

Table 18.   Best Behavior: Frequency of team meetings—3 years

	Year one	Year two	Year three	Total
Average number of meetings	5	8	8	21
Range	2–13	2–28	2–19	6–60

NOTE: Data are based on 18 intervention schools that participated for 3 years.
SOURCE: Liaison monthly reports.

Needs Assessment Completed

Completion of a schoolwide needs assessment was an additional fidelity practice for the first 2 years. The needs assessment provided the team with information on the school staff's perceived needs for improving the school discipline system for the following year. A needs assessment was therefore not expected to be completed in the final year of the study. The status of this milestone is shown in table 19. According to liaisons' year-end implementation reports, by the end of year one a needs assessment was completed by all 18 schools, whereas only 72 percent completed an assessment in the second year.

Table 19.   Progress achieving Best Behavior key practices over 3 years

Measure	Percent of schools completing key practice		
	Year one	Year two	Year three
Needs assessment			
Needs assessment completed	100	72	†
Schoolwide rules			
Rules and schoolwide behavioral expectations are defined; rules are posted in the school	72	83	83
Lesson plans are developed for teaching about behavioral expectations in all school settings (e.g., cafeteria, gym, restrooms); rules and expectations are taught systematically	56	72	78
Rewards system			
A schoolwide system is defined for recognizing and rewarding appropriate, expected behaviors; reinforcement system is implemented	78	78	78
Discipline data			
Discipline data are gathered, summarized, and reviewed periodically	†	33	44
High-risk students			
Individualized self-management plans are developed to address the needs of youth who exhibit problem behaviors	†	†	50

† Not applicable for this year.
NOTE: Schools were rated on a 4-point scale, where 1 = not achieved/initiated, 2 = in progress,
3 = achieved/completed, 4 = substantially achieved. The percentages in this table include scores of 3 and above.
Data are based on 18 intervention schools that participated for 3 years.
SOURCE: Site liaison reports.

Creation of Schoolwide Rules and Rewards System

The creation of a schoolwide rules and rewards system, another fidelity practice, was measured through liaisons' observation of whether schools completed several tasks, which can be found in table 19. An early task for the school management team was to identify, adopt, and disseminate (by posting throughout the school) a set of rules of behavior by which all of the school could be guided. At the end of the first year, 72 percent of schools had completed this task; the percentage was 83 percent in the second and third years.

By year three, 78 percent had developed lesson plans to teach students and faculty about each of the selected rules and expectations and were teaching the rules and expectations systematically to students in the classroom. Regarding the reward system for recognizing appropriate behavior, 78 percent of schools developed and implemented the system by the end of the third year.

Use of Discipline Data and Attending to High-Risk Students

In years two and three, schools were to address two additional program practices. First, schools were to gather, summarize, and review student discipline data periodically. This process

would be used to monitor and evaluate student behavior patterns and potentially identify specific problem areas of the school, times of the day, classrooms, and individuals in need of further attention. As shown in table 19, one-third of the schools established this process by year two, while less than one-half (44 percent, or 8 out of 18 schools) had done so by year three.

An additional practice for Best Behavior in the third year was to address the needs of individual youth who exhibit problem behaviors. Teachers were to develop individualized self-monitoring and self-management plans that teach the students how to manage their own behavior. Table 19 indicates that one-half of the schools were providing this individualized attention to high-risk youth by the third year of implementation.

Saturation of Best Behavior and Clarity of School Rules and Rewards

For faithful implementation of Best Behavior, teachers need to be aware of the school rules, given that teachers will be responsible for making sure students adhere to the rules and will be rewarding students for positive behaviors. Teacher data collected from a sample of the general teaching staff were used to assess how broadly the school rules and expectations for behavior were communicated and instituted, from the teachers' perspectives. Table 20 displays results from the teacher survey administered in spring of each implementation year. In year one, 54 percent of the teachers indicated that when school rules were broken, it was clear to school staff what consequences would follow. Sixty-four percent of the teachers responded affirmatively in year three, indicating that the consequences for misbehavior were still unclear for one-third of the faculty. Other aspects of the school rules and rewards were better understood by the teachers. Specifically, in year three, 79 percent of teachers reported that school rules emphasized the consequences for negative behaviors, 84 percent said that their school rules reinforced desirable behaviors, and 87 percent said that rules for student behavior were clearly defined in their schools.

Table 20.   Best Behavior saturation and clarity of school rules in intervention schools—3 years

	Percent of teachers agreeing		
	Year one (N = 416)	Year two (N = 411)	Year three (N = 429)
School rules for student behavior are clearly defined	82	84	87
The school rules emphasize reinforcing desired behaviors	79	82	84
The school rules emphasize consequences for undesired behaviors	67	74	79
When a school rule is broken, it is clear to school staff what consequences should follow	54	59	64

NOTE: Response options ranged from 1 ("strongly agree") to 4 ("strongly disagree"); "strongly agree" and "agree" responses were combined. Data are based on 18 intervention schools that participated for 3 years.
SOURCE: Teacher survey, spring 2007, spring 2008, and spring 2009.

Interview data collected by the evaluation team during site visits provided additional information about the extent to which the teachers were oriented about key aspects of Best Behavior. In year three, 35 percent of RiPP teachers interviewed said that they received in-service trainings on specific activities related to the Best Behavior program, such as behavioral expectations/school rules, positive reinforcement/reward system, or classroom management techniques.

Challenges Implementing Best Behavior

In an effort to better understand the context for and challenges with implementation of Best Behavior, data were also collected from school management team members through interviews conducted annually by the evaluation team. Nearly all interviewed staff (98 percent) felt that the program was appropriate for their school and felt comfortable with the program (92 percent). Nevertheless, school management team members talked about facing various challenges with implementing Best Behavior, for example, difficulties with finding the time to implement the program. This is substantiated through the liaison year-end reports, which indicated that liaisons were often asked to take on responsibilities that were supposed to be fulfilled by members of the school management team. Other issues mentioned by interviewed school management team members were low teacher buy-in and a lack of student interest in the rewards offered through the reward system.

A general challenge to all intervention schools, at least initially, was funding for the student rewards. According to the liaisons' monthly implementation reports, all 18 intervention schools expressed concern during the first year of implementation about budget constraints that they felt either directly or indirectly affected implementation of the student reward system (e.g., prizes, parties). For example, some schools obtained funding for Best Behavior rewards from community sources. Other schools found it difficult to provide teacher coverage for fundraisers or other Best Behavior activities because teachers were told not to stay after school because the school could not compensate them. In yet other schools, there were academic booster programs that competed with Best Behavior for funding for rewards. Among the concerns expressed by the school management team members that were interviewed in the third year were continued problems with funding the reward system.

3.3.4 Coordination Between RiPP and Best Behavior

Interviews conducted with RiPP teachers and Best Behavior school management teams included questions that asked how aware they were of the other program, in what ways they had been involved with the other program, and what training they had received with regard to the other program. Nearly all interviewed RiPP teachers (96 percent) and school management team members (92 percent) had some level of awareness of both the RiPP and Best Behavior programs. Most of the RiPP teachers (96 percent) indicated they had received at least some information about the Best Behavior program in general, either through announcements made by administrators, discussions at faculty meetings, or through in-service trainings. Despite a lack of training in specific Best Behavior activities, 74 percent of the year three RiPP teachers had some direct involvement in the Best Behavior program, for example, with the reward system or by serving on the Best Behavior school management team.

Interviewed staff were also asked to what extent they believed the Best Behavior and RiPP programs worked together. Eighty-five percent of RiPP teachers and 85 percent of school management team members felt the RiPP and Best Behavior programs worked together, citing that they complement each other, that they promote a positive school environment, and that using RiPP leads to Best Behavior. Other comments from RiPP teachers elicited by this question, however, suggested that although the programs work together in theory, they need to be better integrated in order for students and teachers to see the connection between the two programs. Similarly, comments made by school management team members in response to this question included the suggestion that the two programs could better work together if, for example, teachers had fully implemented the Best Behavior program or if more teachers implemented the RiPP program.

3.4 Treatment Contrast

This section provides information concerning the variety of interventions, aside from RiPP and Best Behavior, in the study's intervention and comparison schools that is potentially related to the study's outcomes of interest. As table 21 indicates, between eight schools and nine schools in the intervention group and between six schools and seven schools in the control group administered classroom-based education other than RiPP across the 3 years of the study. The types of programs implemented included gang resistance programs, character education programs, and individual presentations that were not part of a curriculum (most often, speakers, a video, or a lesson) focused on specific topics such as bullying, harassment, and dating violence. No control school implemented RiPP or Best Behavior during the 3 years of the study.

With regard to violence prevention strategies other than curricula and whole-school approaches, seven or fewer intervention schools and seven or fewer control schools used non-classroom-based approaches, such as peer mediation programs, in each year of the study. Security measures, such as cameras and metal detectors, were used by 11 intervention schools and 13 control schools in year one; 14 intervention schools and 10 control schools in year two; and 15 intervention schools and 14 control schools in year three.

Table 21.   **Number of schools implementing strategies related to violence prevention over 3 years**

Violence prevention strategy	Intervention schools			Control schools		
	Year one	Year two	Year three	Year one	Year two	Year three
Classroom-based education						
RiPP (Intervention)	18	18	18	0	0	0
Other classroom-based education (e.g., Gang Resistance Education and Training [GREAT]; character education; individual lessons, videos or speakers on specific topics)	9	8	8	6	7	7
Whole-school reward/disciplinary approach						
Best Behavior (Intervention)	18	18	18	0	0	0
Non-classroom-based prevention						
Peer mediation, conflict resolution, student court	7	3	7	7	6	4
Security cameras or metal detectors	11	14	15	13	10	14

NOTE: Data are based on 18 intervention schools and 18 control schools that participated for 3 years.
SOURCE: Violence prevention coordinator interview, 2007, 2008, 2009.

3.5 Program Costs

In this section, we present estimates of the direct costs of implementing RiPP and Best Behavior as revised or modified for this study. These estimates are necessarily approximate because the true cost of program administration will depend on a variety of factors, including travel distances and costs for trainers, the number of participating schools in the school district, school size, number of grade levels participating (for RiPP), strategy for teaching the curriculum through either a core

group of staff or the entire faculty, and staff turnover. For the purpose of these calculations, we have made certain assumptions, as described below, based on the typical districts and schools in this study. In practice, schools and districts negotiate with the program developers for services, based on school characteristics and needs. Also, the calculations presented below do not include opportunity costs related to staff and student participation.

Costs for RiPP. Average costs for implementing RiPP over 3 years in a middle school with approximately 600 students are shown in table 22. Direct program costs for RiPP pertain to teacher training, teacher curricula, and student workbooks. Schools may also opt to receive an on-site visit by a RiPP trainer to provide technical assistance once per year. The total cost of the curricula, training, and materials would come to $22,800 over the 3 program years without the optional technical assistance, yielding per teacher costs of $675 and per student costs of $20. If on-site technical assistance is included, the total cost would be $28,800, the cost per teacher would amount to $1,050, and the cost per student would be $20. Note that these estimates do not assume that staff will be compensated for time spent on RiPP-related activities.

Table 22.   Average cost of implementing RiPP in middle schools similar to the study schools over 3 years

Cost category	First-year costs	Second-year costs	Third-year costs	Overall cost	Cost per teacher[1]	Cost per student[2]
Curriculum (set of three, for grades 6 through 8)	2,800	0	0	2,800	175	0
Teacher training	4,000	2,000	2,000	8,000	500	0
Student workbooks	3,000	3,000	6,000	12,000	0	20
Optional technical assistance[3]	2,000	2,000	2,000	6,000	375	0
Total without options	9,800	5,000	8,000	22,800	675	20
Total with options	11,800	7,000	10,000	28,800	1,050	20

[1] Calculations are based on 16 teachers: 8 teachers in year one, and 4 replacement teachers in each of years two and three.
[2] Calculations are based on 600 students, 200 students per grade.
[3] Costs include consultation fees and travel expenses for one RiPP consultant to conduct a 2-day visit each year.
NOTE: The following assumptions are used: (1) eight RiPP teachers deliver the curriculum to all three grades each year; (2) four new RiPP teachers require training in each of the second and third years; (3) the curriculum is purchased in year one and not replaced in follow-up years; and (4) one student workbook is needed in the first 2 years, and two are needed in the program's third year at a cost of $5 per workbook, per student.
SOURCE: RiPP program developers.

Costs for Best Behavior. Direct program costs for Best Behavior pertain to staff training, technical assistance, and program materials or resources, some of which are optional for schools. Best Behavior is implemented through a school management team; for the purposes of these calculations, we have assumed school management teams of 10 staff per school and a 3-year implementation period. The total costs, with and without options, as well as costs per team member and per student, are shown in table 23. Activities for the first year of implementation include an initial needs assessment visit, initial training, and follow-up technical assistance visit. Activities for the second year of implementation include a 2-day ongoing needs assessment visit, a 2-day training visit, and a 2-day follow-up technical assistance visit. Activities for year three would be identical to

year two; however, schools could choose to receive an additional train-the-trainers session for sustainability purposes. Although cost-neutral incentives are encouraged for use with the behavior reinforcement system (e.g., privileges, special celebrations), during each year of implementation, schools also incur costs associated with other types of student incentives and rewards.

Table 23. Average cost of implementing Best Behavior in middle schools similar to the study schools over 3 years

Cost category	First-year costs	Second-year costs	Third-year costs	Overall cost	Cost per team member[1]	Cost per student[2]
Best Behavior guidebook	850	255	255	1,360	85	0
Needs assessment, 2-day visit[3]	3,689	3,689	3,689	11,068	692	0
Team training, 2-day visit[3]	3,689	3,689	3,689	11,068	692	0
Follow-up support, 2-day visit[3]	7,379	3,689	3,689	14,758	922	0
Student incentives	1,500	1,500	1,500	4,500	0	8
Optional train-the-trainer, 2-day visit[3]	0	0	3,689	3,689	231	0
Optional videos[4]	350	0	0	350	22	0
Optional referral tracking system, School-wide Information System (SWIS)[5]	550	250	250	1,050	66	0
Total without options	17,108	12,823	12,823	42,754	2,391	8
Total with options	18,008	13,073	16,762	47,843	2,709	8

[1] Calculations are based on 16 school management team members: 10 staff in year one and 3 replacement staff in each of years two and three.

[2] Calculations are based on 600 students, 200 students per grade.

[3] Costs include consultation fees, travel expenses, and overhead fees for one Best Behavior consultant.

[4] Calculations are based on a one-time purchase of two videos, at an average of $175 per video.

[5] Costs for SWIS include a one-time training fee of $300 and annual license renewal fee of $250 per year.

NOTE: The following assumptions are used: (1) school management teams are comprised of 10 staff each year; (2) three replacement team members require guide books in the second and third years; and (3) needs assessment visits occur once per year, training visits occur once per year, follow-up support visits occur twice in the first year and once thereafter, and the train-the-trainer visit occurs once, in year three.

SOURCE: Best Behavior program developers.

Schools have the option to purchase an annual site license for the School-wide Information System, a tool for management and evaluation of student discipline data (May et al. 2010). Videos that focus on strategies for managing student behavior offer additional resources to staff and may be purchased separately.

Overall, the direct cost of implementing Best Behavior across 3 years, including all options, equals $47,843. This yields a per team member cost of $2,709 and a per student cost of $8. Without any options, the total cost would be $42,754, the per team member cost would be $2,391, and the cost per student would be $8.

This page intentionally left blank.

Chapter 4.
Impacts of the Violence Prevention Program
After 3 Years

This chapter presents the impact findings after 3 years of a program to reduce student violence and victimization in middle schools. Impact findings after 2 years of delivery, which did not differ substantively from impacts after 3 years of delivery, can be found in appendix H. In the sections that follow, we present the results of analyses conducted to assess the impact of the intervention. We focus first on main outcomes that the intervention seeks to influence: student violence and student victimization. In addition, we examine the impacts on several other outcomes. One group of outcomes corresponds to those that, while not directly targeted by the program, might be expected to change as a result of the program: student and teacher safety concerns, student prosocial behaviors, and teacher victimization. Another group of outcomes examined includes those hypothesized to be more immediately affected by the intervention, compared with the main outcomes of violence and victimization. These outcomes include students' attitudes toward violence, students' self-reported coping strategies for dealing with violence, students' clarity of understanding of the school rules, teacher expectations for student behavior, and school staff members' responses to student violence.

This chapter begins with a description of the student characteristics at baseline across intervention and control groups. It then presents the results of the impact findings for the full sample, for both main outcomes and other outcomes. This is followed by a presentation of the impact findings for the high-risk subsamples. Unadjusted means and standard deviations for all impact variables are included in appendix G.

4.1   Student Characteristics at Baseline

This section presents the demographic characteristics of the student sample and the baseline outcome measures, across intervention and control groups, for the 36 schools that participated for all 3 years of program delivery.[19] Table 24 shows the student demographic characteristics at baseline for the intervention and control groups.[20] These data were obtained through the baseline student survey, which was administered in fall 2006. Minority students composed 72 percent of the sample in the intervention group and 61 percent of the sample in the control group. Forty-nine percent of the students in both samples were male, and 60 percent lived in single-adult households. A two-tailed *t*-test applied to each variable indicates that the demographic characteristics are not statistically different between students attending intervention schools and those attending control schools.

---

[19] Four of the original 40 schools were no longer participating in the study at the end of the third year. For details, see section 2.3.3 in chapter 2.

[20] Demographic information for the high-risk student subgroup can be found in appendix B.

Table 24.   Baseline demographic characteristics of the student sample in schools with 3 years of participation

Characteristic	Intervention group	Control group	Difference	t-test[1]	p-value
Sample size, grade 6 (*N* = 36 schools)[2]	3,198	3,418			
Race/ethnicity (%)					
Hispanic	39.34	29.87	9.47	0.39	0.70
Black, non-Hispanic	24.25	21.84	2.41	0.28	0.78
White, non-Hispanic	27.85	39.42	−11.57	−0.77	0.45
Other non-Hispanic or mixed[3]	8.56	8.91	−0.35	−0.16	0.87
Gender (%)					
Male	48.98	48.99	−0.01	−0.04	0.97
Single-adult household (%)	59.69	60.30	−0.61	−0.37	0.72

[1] Adjusted for intraclass correlations.

[2] Data represent students in the 36 (of 40) schools that remained in the study across the 3 years of program implementation.

[3] This category includes American Indian or Alaskan Native, Asian, Native Hawaiian/other Pacific Islander, and multiracial.

NOTE: Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

SOURCE: Student survey, fall 2006 (baseline).

The voluntary aspect of participation in the study means that students from whom outcome data were collected may not be representative of the general middle school student population. To test for potential bias in the study sample, the baseline demographic characteristics of the 6th-grade study participants in the 36 study schools were compared with 6th-grade demographic data for the entire grade obtained from the Common Core of Data (CCD) (http://nces.ed.gov/ccd/districtsearch/index.asp) for the same year.[21] Comparisons for race/ethnicity should be viewed with caution, as the two sources of information used different ways to collect the data.[22] Table 25 shows the results of comparing the demographic characteristics for three categories of race/ethnicity and for gender (male). A two-tailed *t*-test applied to each variable indicates that the demographic characteristics are statistically different between sample students and the population of 6th-graders in those same schools. Higher percentages of Black and male students were observed in the population, compared with the sample of study participants. Similar differences were observed among intervention schools. Among control schools, there was a difference with respect to the percentage of Black students which was found to be significantly higher in the population. In all instances, the mean percentages differ by less than 4 points between sample and

[21] The baseline sample of 6th-grade students participating in the study represented 67 percent of the total number of students enrolled in the 6th grade, based on the response rate.

[22] There is no comparable figure available in CCD for the student survey item that asks about parents living in the household. The race/ethnicity category "other" in CCD is not defined in a comparable way to the item in the student survey, which includes mixed races. These two items were omitted from the analysis. It should also be noted that the race/ethnicity questions are not framed in a comparable manner in these two sources; in particular, the student survey asks about Hispanic ethnicity first and then about race, while the school data collected through CCD do not measure ethnicity separately from race.

population. These results reveal that the students from whom outcome data were collected in year three were not representative of the middle school student population on these specific characteristics.

Table 25.   Difference in baseline demographic characteristics between the student sample and population in schools with 3 years of participation

Characteristic	6th-grade population	6th-grade study participants	Difference	*t*-test	*p*-value
Total sample (*N* = 36 schools)					
Race/ethnicity (Mean %)					
Hispanic	31.61	31.80	−0.19	−0.45	0.65
Black, non-Hispanic	27.91	24.64	3.26	8.24	0.00*
White, non-Hispanic	35.40	34.61	0.79	1.67	0.10
Gender (Mean %)					
Male	50.62	48.81	1.81	2.96	0.01*
Intervention group (*N* = 18 schools)					
Race/ethnicity (Mean %)					
Hispanic	33.42	32.29	0.43	0.72	0.48
Black, non-Hispanic	29.25	26.05	3.20	5.84	0.00*
White, non-Hispanic	32.22	31.56	0.66	1.01	0.32
Gender (Mean %)					
Male	51.42	48.73	2.69	3.05	0.01*
Control group (*N* = 18 schools)					
Race/ethnicity (Mean %)					
Hispanic	29.80	30.56	−0.77	−1.32	0.20
Black, non-Hispanic	26.55	23.23	3.32	5.84	0.00*
White, non-Hispanic	38.60	37.69	0.91	1.33	0.20
Gender (Mean %)					
Male	49.86	48.86	1.00	1.17	0.26

NOTE: A two-tailed test adjusted for intraclass correlations was used to test differences between population estimates and study sample estimates. Statistical significance is indicated by * if the *p*-value is less than or equal to .05. Comparisons for race/ethnicity may not be valid because of differences in the way the data were collected.
SOURCE: Calculations for the "6th-grade study participants" are based on a student survey administered at baseline in fall 2006. Calculations for the "6th-grade population" were based on *Search for Public School Districts: School Year 2005–06*, Common Core of Data (CCD), U.S. Department of Education, Institute of Education Sciences, National Center for Educational Statistics. Retrieved April 29, 2009, from http://nces.ed.gov/ccd/districtsearch/index.asp.

We compared the baseline student measures for primary and secondary outcomes across the intervention and control groups to test whether there were any significant differences between the two groups. These data were obtained through the baseline student survey administered in fall 2006. As shown in table 26, students' responses were not statistically different between intervention and

control schools, with one exception.[23] Students in control schools were more likely than students in intervention schools to worry that someone from their school would attack, hurt, or bully them. These analyses suggest that randomization of schools to condition resulted in generally similar groups.

Table 26. **Baseline measures for the student sample in schools with 3 years of participation: Full student sample**

Measure	Baseline event rates or scale means (SE)		Difference	*t*-statistic	*p*-value[1]
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	1.95 (0.14)	1.85 (0.13)	0.10	0.51	0.61
Violence: Weapons-related	0.08 (0.01)	0.07 (0.01)	0.01	0.51	0.62
Violence: Not weapons-related	1.87 (0.13)	1.78 (0.13)	0.10	0.52	0.61
Victimization (All items)	3.88 (0.15)	4.09 (0.15)	−0.21	−1.00	0.32
Victimization: Overt	2.18 (0.09)	2.31 (0.10)	−0.12	−0.92	0.37
Victimization: Relational	1.69 (0.06)	1.79 (0.06)	−0.10	−1.08	0.29
Other outcomes[3]					
Self-reported coping strategies: Positive	2.88 (0.02)	2.91 (0.02)	−0.03	−1.43	0.16
Self-reported coping strategies: Negative[2]	0.58 (0.04)	0.58 (0.04)	0.00	−0.06	0.95
Attitudes toward violence	3.00 (0.03)	3.02 (0.03)	−0.02	−0.45	0.66
Prosocial behaviors: Extended to others	2.94 (0.03)	2.97 (0.03)	−0.03	−0.84	0.41
Prosocial behaviors: Received from others	2.78 (0.03)	2.79 (0.03)	−0.01	−0.26	0.80
Behavioral expectations	3.16 (0.02)	3.15 (0.02)	0.01	0.32	0.75
Safety concerns	1.86 (0.03)	1.96 (0.03)	−0.10	−2.53	0.02*
Sample size (Schools)	18	18			
Sample size (Students)[4]	3,198	3,418			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates and standard errors (SEs); *t*-statistic adjusted for clustering of students within schools used to test the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). General linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.2 percent to 3.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

---

[23] Unadjusted baseline means and standard deviations can be found in table F-1, appendix F.

Additional analyses examined the student outcomes at baseline for students identified as being at high risk for engaging in violent behavior in the future. Table 27 presents the outcome measures at baseline for students identified as high risk based on their attitudes toward violence but who have committed no violent acts (nonperpetrator group). No significant differences were found between intervention and control groups, with one exception: students in control schools were more likely than students in intervention schools to report that expectations for student conduct were made clear at school.

Table 27. **Baseline measures for the student sample in schools with 3 years of participation: High-risk subgroup (Nonperpetrator)**

| Measure | Baseline event rates or scale means (SE) | | Difference | t-statistic | p-value[1] |
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	1.23 (0.17)	1.21 (0.17)	0.01	0.06	0.95
Violence: Weapons-related	0.00 (0.00)	0.00 (0.00)	—	—	—
Violence: Not weapons-related	1.23 (0.17)	1.21 (0.17)	0.01	0.06	0.95
Victimization (All items)	3.55 (0.29)	3.87 (0.30)	−0.31	−0.76	0.45
Victimization: Overt	20.6 (0.19)	2.20 (0.19)	−0.14	−0.52	0.60
Victimization: Relational	1.53 (0.14)	1.69 (0.15)	−0.15	−0.78	0.44
Other outcomes[3]					
Self-reported coping strategies: Positive	2.04 (0.07)	2.10 (0.07)	−0.06	−0.58	0.56
Self-reported coping strategies: Negative[2]	2.27 (0.16)	2.32 (0.15)	−0.05	−0.25	0.80
Attitudes toward violence	2.18 (0.07)	2.21 (0.07)	−0.03	−0.27	0.79
Prosocial behaviors: Extended to others	2.64 (0.11)	2.64 (0.11)	0.00	0.02	0.98
Prosocial behaviors: Received from others	2.66 (0.09)	2.50 (0.09)	0.15	1.18	0.25
Behavioral expectations	2.83 (0.05)	3.03 (0.05)	−0.19	2.83	0.01*
Safety concerns	1.98 (0.10)	1.74 (0.10)	−0.06	−0.43	0.67
Sample size (Schools)	18	18			
Sample size (Students)[4]	97	106			

— Not available. Nonperpetrator subgroup defined as those youth who reported no violent behaviors at baseline.
[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.
[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and SEs, difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.
[3] Based on continuous scale measures (unless otherwise indicated). General linear mixed model used to estimate group-specific baseline scale means and standard errors (SEs), difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.
[4] Missing data ranged from 0.1 percent to 3.1 percent.
SOURCE: Student survey, fall 2006 (baseline).

Table 28 presents the outcome measures at baseline for students who were identified as high risk based on self-reported aggressive or violent acts (perpetrator group). Significant differences were found between intervention and control groups on four of the measures: high-risk students attending intervention schools reported more overall violence and violence without a weapon than high-risk students attending control schools. Significant differences were also found for positive self-reported coping strategies (students in intervention schools made greater use of these strategies) and concerns over safety (students in control schools were more concerned than students in intervention schools).

Table 28. Baseline measures for the student sample in schools with 3 years of participation: High-risk subgroup (Perpetrators)

| Measure | Baseline event rates or scale means (SE) | | Difference | t-statistic | p-value[1] |
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	5.49 (0.12)	4.393 (0.11)	0.56	3.44	0.00*
Violence: Weapons-related	0.31 (0.03)	0.27 (0.03)	0.04	0.95	0.35
Violence: Not weapons-related	5.18 (0.11)	4.66 (0.09)	0.52	3.72	0.00*
Victimization (All items)	6.17 (0.12)	6.21 (0.11)	−0.04	−0.23	0.82
Victimization: Overt	3.78 (0.07)	3.79 (0.06)	0.00	−0.04	0.97
Victimization: Relational	2.41 (0.06)	2.42 (0.05)	−0.02	−0.24	0.81
Other outcomes[3]					
Self-reported coping strategies: Positive	2.61 (0.03)	2.68 (0.02)	−0.08	−2.05	0.05*
Self-reported coping strategies: Negative[2]	1.13 (0.04)	1.07 (0.03)	0.06	10.9	0.28
Attitudes toward violence	2.59 (0.03)	2.64 (0.03)	−0.05	−1.26	0.22
Prosocial behaviors: Extended to others	2.73 (0.03)	2.77 (0.03)	−0.04	−0.92	0.36
Prosocial behaviors: Received from others	2.70 (0.03)	2.69 (0.03)	0.01	0.31	0.76
Behavioral expectations	3.03 (0.02)	3.01 (0.02)	0.02	0.52	0.61
Safety concerns	1.98 (0.04)	2.21 (0.04)	0.14	−2.47	0.02*
Sample size (Schools)	18	18			
Sample size (Students)[4]	790	922			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to.05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and SEs, difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and SEs, difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level. Based on continuous scale measures (unless otherwise indicated). General linear mixed model used to estimate group-specific baseline scale means and standard errors (SEs), difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

4.2   Interpreting Program Impacts

As described in chapter 2, the constructed indices of violence and victimization that constitute our primary measures for impact outcomes represent a count of the number of violent behaviors or victimization events occurring in the past 30 days. Additional information on items and scale construction, including response category options and ranges, are presented in section 2.8 and appendix C.

To account for the nature of the data, indices are assumed to follow a Poisson distribution, and the results are presented in terms of event rates (ERs) and event rate ratios (ERRs). ERs indicate the incidence density; this refers to the number of events among a particular group for a given period of time. For the measures of violent behavior and victimization in this study, all items assessed occurrences in the past 30 days. Accordingly, an ER of 2.5 among students in intervention schools indicates that students in these schools reported an average of 2.5 incidences in the past 30 days. ERRs *compare* the incidence density among a group of interest (intervention schools) with a group used as a reference (control schools). Where ERRs are greater than 1.00, the indicated group reports a higher frequency of occurrences than the reference group; where ERRs are less than 1.00, the indicated group reports a lower frequency of occurrence than the reference group. An ERR of 2.00 would indicate that, on average, students in the intervention schools reported twice as many incidents in the past 30 days as students in the control schools; similarly, an ERR of 0.50 would indicate that, on average, students in control schools reported twice as many incidents in the past 30 days as students in intervention schools.

To account for multiple comparisons and reduce the risk of spurious rejection of null hypotheses for the main impact estimates, we provide critical value adjustments based on the approach detailed by Benjamini and Hochberg (1995; Thissen, Steinberg, and Kuang 2002). The method provides a sequential approach to control the false discovery rate in multiple comparisons that is less stringent than methods aimed at controlling the experiment-wise error rate. The Benjamini-Hochberg (B-H) correction adjusts the critical *p*-values to which each observed *p*-value is compared to determine significance; the correction is applied to a family of hypotheses examining the utility of a program or intervention. Specifically, the B-H correction assumes the set of *p*-values within a given family of hypotheses represents an ordered distribution of independent observations of a random variable. The correction adjusts the critical value according to the size of the family and the position of the hypothesis test among the ordered set of observations. Based on a priori decision rules, statistical tests for the two main indices—violence and victimization—are treated as independent assessments of program impacts; accordingly, each set of subindices is evaluated as a separate family of hypotheses. The composite index is not included in the family of hypotheses, as this would violate the assumption of independence required for the use of the B-H correction. Where the reported Wald *p*-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

Hypothesis tests related to secondary outcomes and teacher outcomes assess various aspects of the program but are not explicitly targeted by the intervention. Accordingly, the results of these tests do not lead to inferences of program success or failure and so are not corrected for multiplicity.

4.3   Main Program Impacts

This section reports findings that address the main impact research question: "Are there differences in the degree of violence in schools that implement the violence prevention program, relative to schools that do not implement it?" Constructed indices of student violence and victimization constitute the primary measures used to address this research question.

In addition to main indices, we also derived subscales that reflect the underlying dimensions of the main impact outcomes. The overall measure of student violence is disaggregated into (1) student violence with a weapon, and (2) student violence without a weapon. The overall measure of student victimization is disaggregated into (1) overt victimization, and (2) relational victimization. Based on the theory of action, the combined intervention is theorized to impact the levels of student violence and victimization over the 3 years of the study; these findings are based on the 3 years of program implementation.

Analyses examining these outcomes are presented in table 29. In each case, the model predicts average response at follow-up, adjusting for the following covariates: baseline school mean of the response, school size, and individual demographic variables (gender, race/ethnicity, and number of parents in the household). The estimated impacts in the fourth column are ERRs comparing ERs in intervention schools with those reported in control schools. Overall, there were no statistically significant program impacts on student behaviors for either violence (all items, without a weapon, with a weapon) or victimization (all items, overt, relational) after 3 years of implementation. In each case, using the Wald statistic and after controlling for multiple comparisons, results indicated that the rates of behavior reported by students in intervention schools were not statistically different from the rates of behavior reported by students in schools not receiving the intervention. Unadjusted means are presented in appendix G. In summary, intervention students reported an average of 2.88 violent behaviors in the past 30 days (2.72 non-weapon related events and 0.16 weapon-related events), while students in control schools reported 2.69 violent behaviors in the past 30 days (2.58 non-weapon related events and 0.11 weapon-related events). Intervention students also reported being victimized an average of 4.09 times in the past 30 days (2.39 overt victimizations and 1.71 relational victimizations), while students in control schools reported 4.27 violent behaviors in the past 30 days (2.47 overt victimization events and 1.80 relational victimization events).

Table 29. **Main program impacts on self-reported violence and victimization after 3 years of program delivery**

Self-reported student outcome	Model-adjusted follow-up event rates (SE)		Estimated impact (95% CI)[1]	Wald *p*-value[2]	B-H critical *p*-value[3]
	Intervention group	Control group			
Violence (All items)[4]	2.86 (0.11)	2.70 (0.10)	1.06 (0.97, 1.16)	0.1961	†
Not weapons-related	2.72 (0.10)	2.59 (0.09)	1.05 (0.96, 1.15)	0.2699	0.0250
Weapons-related	0.13 (0.01)	0.10 (0.01)	1.27 (1.01, 1.61)	0.0449	0.0125
Victimization (All items)[4]	4.14 (0.11)	4.18 (0.11)	0.99 (0.93, 1.06)	0.7677	†
Overt	2.39 (0.08)	2.41 (0.08)	0.99 (0.91, 1.07)	0.7511	0.0250
Relational	1.76 (0.04)	1.76 (0.04)	1.00 (0.94, 1.07)	0.9938	0.0125
Sample size (Schools)	18	18			
Sample size (5,854 students clustered within schools)[5]	2,784	3,070			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald *p*-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald *p*-value ≤ B-H critical *p*-value, two-tailed test.

[1] Program impact estimated as a model-adjusted event rate ratio (ERR) for intervention versus controls at follow-up, with 95 percent confidence limits. Impact estimates of 1.00 indicate no difference between intervention and control groups.

[2] The Wald *p*-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical *p*-value.

[3] Critical *p*-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional *p*-value when determining statistical significance. Where the Wald *p*-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[4] Based on count data.

[5] Student sample sizes used in the analysis vary due to item nonresponse at follow-up and/or covariate nonresponse. Missing data ranged from 2 percent to 4 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, intervention condition (intervention vs. control), race/ethnicity, sex, number of parents in household, and school size. CI = confidence interval. SE = standard error. SOURCE: Student survey, fall 2006 (baseline) and spring 2009 (third follow-up).

4.4 Other Program Impacts

Besides the main outcomes, there are several secondary outcomes of interest to schools that, while not explicitly targeted by the intervention, may also be affected (i.e., spillover effects). Therefore, impacts were also estimated for the following: student safety concerns, student prosocial behaviors, teacher safety concerns, and teacher victimization. In addition, the intervention theory of action is that the program fosters a number of intermediate changes. These outcomes were measured for students (e.g., attitudes and perceptions toward violence) and teachers (e.g., classroom management techniques).

To assess the influence of the intervention on other outcomes for students, we analyzed the following measures:

- student safety concerns,

- prosocial behavior extended to others,

- prosocial behavior received from others,

- perceptions of behavioral expectations,

- attitudes toward violence,

- positive (appropriate) self-reported coping strategies when faced with violence, and

- negative (inappropriate) self-reported coping strategies when faced with violence.

Analyses examining these student outcomes are presented in table 30. In each case, the covariate models predict the average response at follow-up, adjusting for the following covariates: baseline school mean of the response, school size, and individual demographic variables (gender, race/ethnicity, and number of parents in the household).

Results for secondary student outcomes indicated that there were no significant impacts after 3 years of implementation. With the exception of negative self-reported coping strategies, these outcomes are based on linear scales. The negative self-reported coping strategies scale did not meet the assumptions of linearity and was treated as a count variable. In each case, using the Wald statistic, results indicated that the rates of behaviors, attitudes, and perceptions reported by students in intervention schools were not significantly different from the rates reported by students in control schools.

To assess the influence of the intervention on outcomes for teachers, we analyzed the following measures:

- teacher safety concerns,

- teacher victimization,

- consistency of enforcing school rules,

- classroom management techniques,

- interactions with victims, and

- interactions with aggressors.

Findings for teacher outcomes are presented in table 31. The models predict the average response at follow-up, adjusting for school size. Across all teacher measures, there were no significant impacts. Using the Wald statistic, results indicated that there were no statistically significant differences in the rates of these responses between teachers in intervention schools and those in control schools.

Table 30. **Secondary program impacts on student-level outcomes after 3 years of program delivery**

Self-reported student outcome	Model-adjusted follow-up event rates or scale means (SE)		Estimated impact (95% CI)	Wald *p*-value
	Intervention group	Control group		
Safety concerns[1,2]	1.64 (0.05)	1.65 (0.05)	−0.00 (−0.07, 0.07)	0.9494
Prosocial behaviors extended[1,3]	2.93 (0.05)	2.88 (0.05)	0.05 (−0.03, 0.12)	0.2213
Prosocial behaviors received[1,3]	2.85 (0.05)	2.83 (0.05)	0.02 (−0.04, 0.08)	0.4734
Perceived behavioral expectations[1,3]	2.84 (0.03)	2.81 (0.03)	0.03 (−0.05, 0.11)	0.4327
Attitudes toward violence[1,3]	2.73 (0.02)	2.75 (0.02)	−0.02 (−0.09, 0.05)	0.5696
Coping strategies (Negative)[2,4]	1.08 (0.04)	1.04 (0.04)	1.04 (0.95, 1.13)	0.3868
Coping strategies (Positive)[1,3]	2.70 (0.02)	2.70 (0.02)	−0.01 (−0.06, 0.05)	0.8282
Sample size (Schools)	18	18		
Sample size (5,554 students clustered within schools)[5]	2,784	3,070		

* Statistically significant at Wald *p*-value (unadjusted) ≤ .05, two-tailed test.

[1] Scales based on continuous measures of the identified construct, unless otherwise indicated. Results presented include the estimated group-specific scale means and standard errors (SEs) at follow-up, estimated program impact for intervention versus control (difference in scale means, with 95 percent confidence limits), and the Wald *p*-value indicating statistical significance of the program impact. Impact estimates of 0.00 indicate no difference between intervention and control conditions.

[2] Lower scores indicate better outcomes.

[3] Higher scores indicate better outcomes.

[4] Based on count data. Results include the estimated group-specific event rates and SEs at follow-up, estimated program impact for intervention versus control (event rate ratio [ERR], with 95 percent confidence limits), and Wald *p*-value indicating statistical significance of the program impact. Impact estimates of 1.00 indicate no difference between intervention and control conditions.

[5] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 2 percent to 6 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function for count data) and linear mixed models (PROC MIXED, for continuous data) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, race/ethnicity, gender, number of parents in household, and school size. CI = confidence interval.

SOURCE: Student survey, fall 2006 (baseline) and spring 2009 (third follow-up).

Table 31.   Secondary program impacts on teacher outcomes after 3 years of program delivery

Teacher self-reported outcome[1]	Model-adjusted odds or scale means (SE)		Estimated impact (95% CI)	Wald Chi-Square *p*-value
	Intervention group	Control group		
Teacher self-reported victimization[2,3]	0.57 (0.05)	0.57 (0.05)	0.50 (0.40, 0.61)	0.9397
Teacher safety concerns[3]	1.36 (0.04)	1.35 (0.04)	0.01 (−0.13, 0.14)	0.9109
School consistency of enforcing behavioral rules[4]	3.00 (0.06)	2.94 (0.06)	0.06 (−0.13, 0.26)	0.4904
Interactions with victims[4]	3.42 (0.04)	3.33 (0.04)	0.09 (0.00, 0.19)	0.0588
Interactions with aggressors[4]	3.53 (0.03)	3.53 (0.03)	0.00 (−0.09, 0.09)	0.9852
Classroom management techniques[4]	3.65 (0.02)	3.62 (0.02)	0.02 (−0.04, 0.09)	0.4596
Sample size (Schools)	18	18		
Sample size (917 teachers nested within schools)[5]	429	428		

* Statistically significant at Wald *p*-value (unadjusted) ≤ .05, two-tailed test.

[1] Reported as scale scores, unless otherwise indicated, based on continuous measures of the identified construct. Results presented include the estimated group-specific scale means and standard errors (SEs), estimated program impact (difference in scale means for intervention versus control, with 95 percent confidence limits), and the Wald *p*-value indicating statistical significance of the program impact. Impact estimates of 0.00 indicate no difference between intervention and control conditions.

[2] Teacher victimization is based on a dichotomous indicator. Results presented include the estimated group-specific odds and SEs, estimated program impact for intervention versus control (odds ratio, with 95 percent confidence limits), and the Wald *p*-value indicating statistical significance of the program impact. Impact estimates of 1.00 indicate no difference between intervention and control conditions.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

[5] Teacher sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0 percent to 1 percent.

NOTE: Program impacts are estimated using restricted maximum likelihood (MIXED procedure) or pseudo-likelihood (GLIMMIX procedure), controlling for the random assignment of schools to program condition from pairs matched within district on the level of free or reduced-price lunches received by students. CI = confidence interval.

SOURCE: Teacher survey, spring 2009.

4.5   High-Risk Student Analyses

This section reports the results of analyses conducted to assess the capacity of the Responding in Peaceful and Positive Ways (RiPP) and Best Behavior programs to promote individual change in a group of students identified as being at higher risk for perpetration of violent acts. These students constitute a cohort that was tracked longitudinally from within the overall sample of students. Analyses for the high-risk group addressed the second primary research question: "What is the impact of the violence prevention program over time on students who are at elevated risk for violence?"

Data are presented separately for two groups of high-risk youth. Table 32 contains results of analyses based on a repeated measures model of students identified at baseline as high-risk based on their responses to a number of attitude and belief questions regarding the appropriateness of using violence while not reporting any violent behaviors in the past 30 days (i.e., nonperpetrator group). Table 33 contains results of analyses based on a repeated measures model of students identified at baseline as high-risk based on their acknowledgment of engaging in violent behaviors in the past 30 days (i.e., perpetrator group).

The repeated measures model contains the student's treatment condition (intervention vs. control), data collection wave, wave-by-group interaction effect, gender, race/ethnicity, number of parents in household, and school size. The repeated measures approach is also known as the difference-of-differences approach because the interaction term tests the effect of baseline versus follow-up for intervention versus controls. Tables 32 and 33 include the model-adjusted group-specific average rates of violent behaviors and victimization at baseline and third follow-up. Estimated program impacts reflect the net difference of the within-group change from pretest to third follow-up for intervention versus controls.

Overall, after controlling for multiple comparisons, there were no statistically significant program impacts on high-risk student behaviors for either violence (all items, without a weapon, with a weapon) or victimization (all items, overt, relational) after 3 years of implementation, among high-risk, nonperpetrator students (table 32) and high-risk, perpetrator students (table 33). These analyses indicate that the observed changes from baseline to third follow-up in the reported rates of violent behaviors and victimization among high-risk students are similar among schools receiving the RiPP and Best Behavior programs, compared with those in the control condition.

Table 32. **Main program impacts on self-reported violence and victimization, after 3 years of program delivery: High-risk, nonperpetrator subgroup (Via repeated measures)**

Student self-reported outcome[1]	Model-adjusted baseline event rates (SE)[2]		Model-adjusted follow-up event rates (SE)[2]		Estimated impact (95% CI)[3]	Wald p-value[4]	B-H critical p-value[5]
	Intervention group[2]	Control group[2]	Intervention group[2]	Control group[2]			
Violence (All items)	1.28 (0.27)	1.32 (0.27)	3.30 (0.54)	3.12 (0.51)	0.92 (0.53, 1.60)	0.7557	†
Not weapons-related	1.32 (0.25)	1.31 (0.24)	3.18 (0.46)	2.86 (0.42)	0.90 (0.49, 1.67)	0.7409	0.0250
Weapons-related[6]	0.00 (0.00)	0.00 (0.00)	0.08 (0.18)	0.14 (0.34)	0.54 (0.13, 2.29)	0.3905	0.0125
Victimization (All items)	3.61 (0.47)	3.41 (0.43)	4.30 (0.52)	3.79 (0.47)	0.93 (0.60, 1.46)	0.7600	†
Overt	2.13 (0.32)	1.90 (0.28)	2.64 (0.37)	2.26 (0.32)	0.96 (0.56, 1.65)	0.8794	0.0250
Relational	1.48 (0.18)	1.48 (0.17)	1.63 (0.19)	1.50 (0.17)	0.93 (0.62, 1.38)	0.6969	0.0125
Sample size (Schools)	18	18	18	18			
Sample size (Students within schools)[7]	70	74	70	74			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald p-value ≤ B-H critical p-value, two-tailed test.

[1] Based on count data.

[2] Group by time-specific event rates.

[3] Program impact (with 95 percent confidence limits) estimated via difference-in-difference models comparing change across time in the intervention versus control group. Ratios of impact estimates of 1.00 indicate no interaction between time and program group (i.e., no program impact).

[4] The Wald p-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical p-value.

[5] Critical p-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional p-value when determining statistical significance. Where the Wald p-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[6] Parameter estimates cannot be obtained due to convergence issues. This may be due to lack of variability of the response variable and/or model complexity.

[7] Student sample sizes used in the analysis vary due to item nonresponse at baseline, follow-up, and/or covariate nonresponse. Missing data ranged from 3 percent to 5 percent, with 240 students missing at follow-up.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included gender, race/ethnicity, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student surveys limited to a high risk-subgroup of 6th-graders surveyed in fall 2006 and followed up in spring 2009.

Table 33.   Main program impacts on self-reported violence and victimization, after 3 years of program delivery: High-risk, perpetrator subgroup (Via repeated measures)

Student self-reported outcome[1]	Model-adjusted baseline event rates (SE)[2]		Model-adjusted follow-up event rates (SE)[2]		Estimated impact (95% CI)[3]	Wald p-value[4]	B-H critical p-value[5]
	Intervention group[2]	Control group[2]	Intervention group[2]	Control group[2]			
Violence (All items)	5.36 (0.20)	4.92 (0.18)	3.95 (0.17)	3.90 (0.15)	1.08 (0.93, 1.25)	0.3244	†
Not weapons-related	5.08 (0.19)	4.65 (0.16)	3.73 (0.15)	3.71 (0.14)	1.09 (0.94, 1.26)	0.2577	0.0125
Weapons-related	0.27 (0.03)	0.26 (0.03)	0.21 (0.03)	0.18 (0.02)	0.91 (0.56, 1.47)	0.6830	0.0250
Victimization (All items)	6.29 (0.19)	6.10 (0.17)	4.90 (0.16)	5.05 (0.15)	1.06 (0.95, 1.19)	0.2703	†
Overt	3.80 (0.13)	3.75 (0.12)	3.00 (0.11)	3.02 (0.11)	1.02 (0.90, 1.17)	0.7247	0.0250
Relational	2.49 (0.07)	2.35 (0.06)	1.91 (0.06)	2.02 (0.06)	1.12 (1.01, 1.24)	0.0274*	0.0125
Sample size (Schools)	18	18	18	18			
Sample size (Students within schools)[6]	552	661	552	661			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald p-value ≤ B-H critical p-value, two-tailed test.

[1] Based on count data.

[2] Group by time-specific event rates.

[3] Program impact (with 95 percent confidence limits) estimated via difference-in-difference models comparing change across time in the intervention versus control group. Ratios of impact estimates of 1.00 indicate no interaction between time and program group (i.e., no program impact).

[4] The Wald p-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical p-value.

[5] Critical p-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional p-value when determining statistical significance. Where the Wald p-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[6] Student sample sizes used in the analysis vary due to item nonresponse at baseline, follow-up, and/or covariate nonresponse. Missing data ranged from 4 percent to 5 percent, with 240 students missing at follow-up.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included gender, race/ethnicity, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student surveys limited to a high-risk subgroup of 6th-graders surveyed in fall 2006 and followed up in spring 2009.

4.6   Conclusions

The purpose of this study was to evaluate the impact of two selected programs, RiPP (Meyer and Northup 2002a, 2002b, 2006) and Best Behavior (Sprague and Golly 2005), which formed a combined intervention. Implementation of the intervention was planned over 3 years in a purposive sample of 40 schools, of which one-half were randomly selected to receive the intervention and one-half served as controls. By the end of the third year, 36 schools remained in the study. This report presents the findings after 3 years of program implementation.

While the ultimate goal of the combined intervention is to decrease student levels of violence and victimization over the 3 years of the study, changes in student and teacher attitudes, perceptions, and self-reported coping skills are expected to emerge prior to any changes in the main outcomes. Thus, in addition to measuring student levels of violence and victimization, we also measured a number of other intermediate outcomes for both students and teachers, through self-report surveys.

After 3 years of implementation, the combined curriculum and whole-school intervention did not show impact effects on any of the student or teacher outcome measures. As the implementation results document, the programs being evaluated as part of the study were not fully implemented with complete fidelity during these 3 years. This has the potential to limit the ability to find statistically significant differences between intervention and control schools.

After 3 years of exposure to the RiPP and Best Behavior intervention, the key impact findings are summarized as follows:

- After controlling for multiple comparisons, there were no significant differences between the students in intervention and control schools on the main outcome measures of violence (overall, with a weapon, without a weapon) or victimization (overall, overt, relational).

- Student measures for other outcomes—including safety concerns, prosocial behaviors, perception of behavioral expectations, attitudes toward violence, and self-reported strategies for coping with violence—did not differ between students in intervention schools and students in control schools.

- Additional impact outcomes for teachers indicated that teacher reports of victimization, safety concerns, enforcement of school rules, and student behavior management were not statistically different between intervention and control schools.

- There were no statistically significant impacts of the RiPP and Best Behavior programs among the subpopulations of high-risk youth, as measured by student violence and victimization after controlling for multiple comparisons.

4.7   Study Limitations

This evaluation study has several limitations:

- The study utilized a random assignment design, the most rigorous approach for estimating program impacts. The sampling frame was constructed to meet certain eligibility requirements of the study design (e.g., schools including at least grades 6 through 8 with a 6th-grade enrollment of at least 250 students; districts with three or more middle schools; oversampling of urban, high-poverty, and high-minority schools).

Schools within this frame were excluded if they were already implementing or had plans to implement a competing violence prevention program similar to the treatment programs. Among the remaining schools, we solicited participation among schools that had available time during the school day to implement the curriculum and then randomized half of the self-selected sample to receive the intervention. Accordingly, results of this study are only generalizable to districts and schools that are similar to those included in the sampling frame.

- In practice, schools or districts seek out prevention programs to address an identified need or problem. As is the case with random assignment studies, schools were recruited into the evaluation and, if assigned to the intervention group, were also asked to implement the two programs over 3 years. Motivation for implementing the programs and staying the course over 3 years varied, as did the priority placed on the intervention by administrators; for example, four schools were lost over the course of the study due to the intervention schools dropping out of the study. Further, the intervention was not uniformly implemented with fidelity, as indicated by challenges faced by teachers with delivery of the curriculum and inconsistent delivery of the whole-school portion of the intervention in year three.

- The current study is a test of the combination of the two interventions as a hybrid model. Thus, it does not provide information on whether separate tests of the two programs would yield different results. The assumption, based on prior research and expert advice, was that the combination of the two programs had a greater likelihood of impacts than a single program. The results of this study are thus limited to the combined effects of the two programs.

- The study uses a serial cross-sectional design for the main impact analysis and a cohort design among the subsample of high-risk students. For the main impact analysis, data collection was conducted in middle schools among the 6th-grade class in year one of the study, among the 7th-grade class in year two of the study, and among the 8th-grade class in year three of the study. This choice was based on the combined intervention program's theory of change, which espoused a social-environmental emphasis on altering school contexts that support violent and aggressive responding. For the main impact analyses, this means that program impacts are modeled as population average change rather than measures of individual difference.

- The self-report student survey was the single source of outcome data for student impacts on violence and victimization. However, student self-report provided data that may not have been captured through other means. In particular, these types of behaviors more commonly occur outside of the classroom and thus may not be witnessed by staff and reported to the school administration. In general, students also have the best knowledge about their own behaviors. While there is a potential for students to underreport violence and victimization behaviors when using self-report, especially if there are concerns about data confidentiality, underreporting would be expected to occur to the same extent in treatment and control schools.

This page intentionally left blank.

Chapter 5.
Exploratory Analyses

This chapter presents three sets of analyses that seek to provide the reader with a richer context for understanding the implementation and outcomes of the Responding in Peaceful and Positive Ways (RiPP) and Best Behavior programs. The first analysis examines the potential for differential impacts between boys and girls in the intervention and control groups, given that girls and boys have been found to differ with regard to type and expression of aggressive behavior (Orpinas and Horne 2006; Crick and Grotpeter 1995). The model used to explore differences between girls and boys permits estimation of gender-specific impacts. The second set of analyses uses the data that were collected about the extent to which execution of program activities adheres to program guidelines(fidelity). These analyses examine how these factors relate to program impacts. Finally, in order to assess the association of "full dosage" on the outcomes of interest, the third analysis examines program outcomes only for students who received the maximum intervention exposure (i.e., in a treatment school for all 3 years of program implementation). This analysis is limited to students who attended the same study schools over the 3 years of program implementation, representing approximately 55 percent of the full sample. The analysis does not account for number of lessons received over the study period as data at that level were not available. Both the second and third sets of analyses are correlational.

5.1 Gender Subgroup Analysis

This section presents results of gender subgroup analyses. The youth aggression literature suggests that girls differ from boys in the types of aggression displayed, with relational aggression being more prevalent among girls than among boys (Orpinas and Horne 2006; Crick and Grotpeter 1995). Furthermore, research indicates that prevention programs may impact girls and boys differently (Farrell and Meyer 1997; Simon et al. 2002). These analyses were undertaken to explore whether or not the intervention affects boys and girls differently.

The outcome measures at baseline for gender subgroups in schools with 3 years of participation are shown in tables 34 and 35, for boys and girls, respectively. There were no differences at baseline between boys in intervention and control groups or between girls in intervention and control groups, with the exception of one measure. For both boys and girls, students in the control group reported greater concerns for their safety at school than did students in the intervention group.

Table 36 reports on impact analyses for gender subgroups. The table includes the model-adjusted mean rates by intervention condition separately for boys and for girls. Impacts for boys and impacts for girls are reported as the within-gender event rate ratios (ERRs) comparing students receiving the RiPP and Best Behavior programs with students in control schools. The within-gender ERRs associated with both violent behaviors and victimization were not statistically significant, indicating that the program did not have impacts on either boys or girls.

The overall net impacts between boys and girls were obtained from the interaction of gender with intervention condition and estimated as the ratio of the two within-gender comparisons, with boys as the indicated group and girls as the reference group. In this case, the ERRs associated with violent behaviors were not statistically significant, indicating that the program did not have differential effects on boys relative to girls. However, the ERRs associated with victimization

indicate a statistically significant effect. Model-adjusted means based on follow-up data demonstrate that boys in intervention schools reported higher levels of victimization relative to boys in control schools at follow-up, while girls in intervention schools reported lower levels of victimization than girls in control schools, producing a disordinal interaction. Examining victimization closely reveals that the effect is driven by items measuring overt victimization. This interaction should be interpreted with care. Note that neither of the gender-specific intervention-control comparisons is statistically significant. Rather, it is the disordinal nature of the interaction (i.e., effects moving in different directions) that provides an overall mean difference that reaches a level of significance.

Table 34.   Baseline measures for the student sample in schools with 3 years of participation: Boys

Measure	Baseline event rates or scale means (SE)		Difference	*t*-statistic	*p*-value[1]
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	2.07 (0.15)	2.08 (0.15)	−0.01	−0.05	0.96
Violence: Weapons-related	0.10 (0.01)	0.10 (0.01)	−0.00	−0.21	0.83
Violence: Not weapons-related	1.97 (0.14)	1.97 (0.14)	−0.00	−0.01	0.99
Victimization (All items)	4.12 (0.16)	4.50 (0.17)	−0.39	−1.68	0.10
Victimization: Overt	2.46 (0.10)	2.69 (0.11)	−0.22	−1.47	0.15
Victimization: Relational	1.65 (0.06)	1.82 (0.07)	−0.17	−1.82	0.08
Other outcomes[3]					
Self-reported coping strategies: Positive	2.77 (0.03)	2.82 (0.02)	−0.05	−1.34	0.19
Self-reported coping strategies: Negative[2]	0.63 (0.04)	0.70 (0.04)	−0.06	−1.23	0.23
Attitudes toward violence	2.90 (0.03)	2.90 (0.03)	0.01	0.13	0.90
Prosocial behaviors: Extended to others	2.84 (0.03)	2.85 (0.03)	−0.01	−0.21	0.83
Prosocial behaviors: Received from others	2.69 (0.03)	2.68 (0.03)	0.01	0.25	0.80
Behavioral expectations	3.13 (0.02)	3.12 (0.02)	0.00	0.14	0.89
Safety concerns	1.79 (0.03)	1.88 (0.03)	−0.09	−2.28	0.03*
Sample size (Schools)	18	18			
Sample size (Students)[4]	1,560	1,670			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and standard errors (SEs), difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.2 percent to 4.6 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table 35. **Baseline measures for the student sample in schools with 3 years of participation: Girls**

Measure	Intervention group	Control group	Difference	*t*-statistic	*p*-value[1]
	Baseline event rates or scale means (SE)				
Main outcomes[2]					
Violence (All items)	1.83 (0.16)	1.60 (0.14)	0.23	1.09	0.28
Violence: Weapons-related	0.06 (0.01)	0.04 (0.01)	0.02	1.12	0.27
Violence: Not weapons-related	1.77 (0.15)	1.56 (0.13)	0.22	1.08	0.29
Victimization (All items)	3.63 (0.18)	3.67 (0.19)	−0.04	−0.15	0.88
Victimization: Overt	1.91 (0.11)	1.92 (0.11)	−0.01	−0.06	0.96
Victimization: Relational	1.72 (0.08)	1.76 (0.08)	−0.04	−0.36	0.72
Other outcomes[3]					
Self-reported coping strategies: Positive	2.98 (0.02)	3.00 (0.02)	−0.02	−0.70	0.49
Self-reported coping strategies: Negative[2]	0.52 (0.04)	0.46 (0.04)	0.06	0.98	0.33
Attitudes toward violence	3.09 (0.04)	3.14 (0.04)	−0.05	−0.80	0.43
Prosocial behaviors: Extended to others	3.02 (0.03)	3.08 (0.03)	−0.06	−1.26	0.22
Prosocial behaviors: Received from others	2.86 (0.03)	2.89 (0.03)	−0.03	−0.74	0.47
Behavioral expectations	3.20 (0.03)	3.19 (0.02)	0.01	0.33	0.74
Safety concerns	1.92 (0.04)	2.04 (0.04)	−0.12	−2.16	0.04*
Sample size (Schools)	18	18			
Sample size (Students)[4]	1,625	1,739			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and standard errors (SEs), difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.1 percent to 2.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table 36. **Main program impacts on self-reported violence and victimization—Year three: Gender subgroups**

Student self-reported outcome[1]	Model-adjusted follow-up event rates (SE)				Impact on boys (95% CI)[2]	Impact on girls (95% CI)[2]	Ratio of impacts (95% CI)[3]	Wald Chi-Square p-value[4]	B-H critical p-values[5]
	Boys		Girls						
	Intervention group	Control group	Intervention group	Control group					
Violent and aggressive behaviors (All items)	3.00 (0.13)	2.75 (0.11)	2.73 (0.12)	2.65 (0.11)	1.09 (0.97, 1.22)	1.03 (0.92, 1.15)	1.06 (0.93, 1.21)	0.3831	†
Not weapons-related	2.81 (0.12)	2.60 (0.11)	2.63 (0.11)	2.57 (0.10)	1.08 (0.96, 1.21)	1.02 (0.92, 1.15)	1.05 (0.93, 1.20)	0.3988	0.0125
Weapons-related	0.18 (0.02)	0.14 (0.02)	0.10 (0.01)	0.08 (0.01)	1.27 (0.95, 1.68)	1.28 (0.89, 1.84)	0.99 (0.64, 1.54)	0.9623	0.0250
Victimization (All items)	4.40 (0.14)	4.23 (0.13)	3.90 (0.12)	4.14 (0.12)	1.04 (0.96, 1.13)	0.94 (0.87, 1.02)	1.10 (1.00, 1.22)	0.0469	†
Overt	2.71 (0.10)	2.55 (0.09)	2.10 (0.08)	2.30 (0.08)	1.06 (0.96, 1.17)	0.91 (0.82, 1.01)	1.16 (1.04, 1.30)	0.0115*‡	0.0125
Relational	1.71 (0.05)	1.68 (0.04)	1.81 (0.05)	1.83 (0.05)	1.02 (0.93, 1.10)	0.99 (0.91, 1.07)	1.03 (0.94, 1.13)	0.5336	0.0250
Sample size (Schools)	18	18	18	18					
Sample size (Students nested within schools)[6]	1,372	1,501	1,399	1,547					

† Not applicable. The Benjamini-Hochberg (BH) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald p-value ≤ B-H critical p-value, two-tailed test.

[1] Based on count data.

[2] Model-adjusted event rate ratios (ERRs) for intervention versus controls at follow-up, separately for boys versus girls. Impact estimates of 1.00 indicate no difference between intervention and control conditions.

[3] ERRs of program impacts for boys versus girls, with 95 percent confidence limits. Ratios of impact estimates of 1.00 indicate no interaction between gender and program group.

[4] The Wald p-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical p-value.

[5] Critical p-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional p-value when determining statistical significance. Where the Wald p-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[6] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 2 percent to 4 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, gender, race/ethnicity, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student survey, fall 2006 (baseline) and spring 2009 (third follow-up).

5.2   Level of Implementation

This section presents results of analyses that assess the association between fidelity indicators of the two programs, as described in section 2.8.3, and then compares implementation fidelity to program impacts. For the RiPP curriculum, these include program adherence and student responsiveness, also described in appendix D, and program exposure. For the Best Behavior program, these include program progress and saturation. In section 5.2.1, we present the associations among the five implementation indices. In section 5.2.2, we present associations between each implementation index and the main program impacts for self-reported violence and victimization.

5.2.1   Association Among Fidelity Indices

The goal of these analyses was to explore the association between implementation of the RiPP curriculum and the Best Behavior program. Using the Pearson Correlation Coefficient (r), we examined the associations between the RiPP and Best Behavior fidelity scores at the school level. For this analysis, there are 18 observations representing the 18 schools in the intervention condition. Information on the five implementation indices is presented in chapter 2.

The correlations among the five fidelity indices for RiPP and Best Behavior are reported in table 37. The two fidelity indicators for Best Behavior, measuring program progress and saturation, were positively correlated. Among the RiPP measures, program adherence was positively correlated with both student responsiveness and program exposure; however, the latter two measures were not correlated with each other. Results for the correlations between RiPP and Best Behavior fidelity measures indicate that the Best Behavior measure of program progress is positively associated with both RiPP adherence and RiPP exposure.

Table 37.   Pearson's Correlation Coefficient between RiPP and Best Behavior fidelity indices—Year three

Fidelity index	Best Behavior: Saturation	Best Behavior: Progress	RiPP: Adherence	RiPP: Student responsiveness	RiPP: Exposure
Best Behavior: Saturation	1.00				
Best Behavior: Progress	0.607 (0.01*)	1.00			
RiPP: Adherence	0.391 (0.11)	0.486 (0.04*)	1.00		
RiPP: Student responsiveness	−0.133 (0.60)	0.088 (0.73)	0.605 (0.01*)	1.00	
RiPP: Exposure	0.461 (0.54)	0.501 (0.03*)	0.677 (0.00*)	0.309 (0.21)	1.00

NOTE: N = 18 intervention schools. *p*-values appear in parentheses; statistical significance is indicated by * if the *p*-value is less than or equal to .05.
SOURCE: Classroom observations, curriculum implementation records, year-end reports, and teacher survey (spring 2009).

5.2.2 Association of Fidelity Indices and Program Impacts

Table 38 examines the association between fidelity indices for RiPP and Best Behavior and primary program impacts for violence and victimization. Results indicate that the RiPP fidelity measure for student responsiveness is significantly correlated with each of the impact variables, except for weapons-related violence. These correlations are positive, implying that higher scores of the RiPP fidelity measure for student responsiveness are associated with larger impacts for violence and victimization. However, these correlations do not imply a cause and effect; therefore, other plausible explanations must be considered. The Best Behavior fidelity measures did not show any significant levels of correlation with the impact variables.

Table 38. Pearson's Correlation Coefficient between RiPP and Best Behavior fidelity indices, and program impacts—Year three

	Violence			Victimization		
	Overall	Weapons-related	Not weapons-related	Overall	Relational	Overt
RiPP: Adherence	0.215 (0.39)	−0.002 (0.99)	0.248 (0.32)	0.266 (0.29)	0.344 (0.16)	0.233 (0.35)
RiPP: Student responsiveness	0.624 (0.006*)	0.405 (0.10)	0.638 (0.004*)	0.638 (0.004*)	0.603 (0.008*)	0.627 (0.005*)
RiPP: Exposure	0.186 (0.46)	0.038 (0.88)	0.209 (0.41)	0.224 (0.37)	0.274 (0.27)	0.204 (0.42)
Best Behavior: Saturation	−0.030 (0.91)	−0.287 (0.25)	0.023 (0.93)	0.061 (0.81)	0.229 (0.36)	0.010 (0.97)
Best Behavior: Progress	−0.034 (0.89)	−0.131 (0.61)	−0.015 (0.95)	0.001 (1.00)	0.061 (0.81)	−0.017 (0.95)

NOTE: N = 18 intervention schools and 2,784 students. *p*-values appear in parentheses; statistical significance is indicated by * if the *p*-value is less than or equal to .05.

SOURCE: Outcome data are from the student survey, spring 2009 (third follow-up); implementation fidelity data were obtained through classroom observations, curriculum implementation records, year-end reports, and teacher survey (spring 2009).

5.3 Examining Maximum Program Exposure and Outcomes

This section presents an analysis designed to examine program outcomes among a subgroup of students who received "full dosage" via exposure to the RiPP curriculum and Best Behavior program across the 3-year study period. The model used to conduct this analysis is the same as the main impact model presented in chapter 4; here, however, we have altered the characteristics of the selected sample to examine whether program impacts are affected. At the time of the year three analysis, approximately 55 percent of the student sample represented youth who had remained in the same school across the study period. For students in the intervention group, the 3-year attendance at the same school thus represents the maximum exposure possible for the intervention. *Exposure,* in this sense, is not related to quality of program implementation. Maximum exposure refers only to the fact that students were present at an intervention school and, therefore, exposed to elements of the intervention across the entire study period.

Results of this analysis are presented in table 39. Indices assessing both violence and victimization show patterns that are similar to the main impact tables presented in chapter 4. None of the ERRs reported fall beyond the 95 percent confidence interval, suggesting that none of the differences between intervention and control groups occurred beyond chance.

Table 39.   Main program outcomes on self-reported violence and victimization in year three for students who attended either intervention or control schools for all 3 years

Self-reported student outcome	Model-adjusted follow-up event rates (SE)		Estimated outcomes (95% CI)[1]	Wald Chi-Square p-value[2]	B-H critical p-values[3]
	Intervention group	Control group			
Violence (All items)[4]	2.68 (0.15)	2.54 (0.13)	1.05 (0.92, 1.21)	0.4075	†
Not weapons-related	2.57 (0.14)	2.43 (0.12)	1.06 (0.93, 1.21)	0.3739	0.0125
Weapons-related	0.10 (0.02)	0.10 (0.02)	1.09 (0.64, 1.84)	0.7395	0.0250
Victimization (All items)[4]	4.25 (0.14)	4.08 (0.13)	1.04 (0.95, 1.14)	0.3583	†
Overt	2.42 (0.10)	2.33 (0.09)	1.04 (0.94, 1.15)	0.4269	0.0250
Relational	1.81 (0.05)	1.75 (0.05)	1.04 (0.95, 1.13)	0.3818	0.0125
Sample size (Schools)	18	18			
Sample size (3,240 students clustered within schools)[5]	1,526	1,714			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald *p*-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald *p*-value ≤ B-H critical *p*-value, two-tailed test.

[1] Program impact estimated as a model-adjusted event rate ratio (ERR) for intervention versus controls at follow-up, with 95 percent confidence limits. Impact estimates of 1.00 indicate no difference between intervention and control groups.

[2] The Wald *p*-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical *p*-value.

[3] Critical *p*-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional *p*-value when determining statistical significance. Where the Wald *p*-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[4] Based on count data.

[5] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 2 percent to 4 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, race/ethnicity, gender, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student survey, fall 2006 (baseline) and spring 2009 (third follow-up).

This page intentionally left blank.

References

Allred, C.G. (1984). *Positive Action.* Twin Falls, ID: Positive Action, Inc. Retrieved December 7, 2009, from http://www.positiveaction.net.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Barrett, S., Bradshaw, C., and Lewis-Palmer, T. (2008). Maryland Statewide PBIS Initiative: Systems, Evaluation, and Next Steps. *Journal of Positive Behavior Interventions, 10:* 105–114.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1): 289–300.

Bos, H., Weinstock, P., and Frankenberg, A. (2004). *Design Considerations for a National Evaluation of a School-Based Violence Prevention Program.* Oakland, CA: Berkeley Policy Associates for U.S. Department of Education.

Bosworth, K., and Sailes, J. (1993). Content and Teaching Strategies in 10 Selected Drug Abuse Prevention Curricula. *Journal of School Health*, *63*(6): 247–253.

Bradshaw, C., Koth, C., Bevans, K., Ialongo, N., and Leaf, P. (2008). The Impact of School-wide Positive Behavioral Interventions and Supports (PBIS) on the Organizational Health of Elementary Schools. *School Psychology Quarterly, 23*: 462–473.

Bradshaw, C., Mitchell, M., and Leaf, P. (2010). Examining the Effects of Schoolwide Positive Behavioral Interventions and Supports on Student Outcomes. *Journal of Positive Behavior Interventions, 12*: 133–148.

Brener, N.D., Kann, L., McManus, T., Kinchen, S., Sundberg, E., and Ross, J. (2002). Reliability of the 1999 Youth Risk Behavior Survey Questionnaire. *Journal of Adolescent Health, 31:* 336–342.

Collins, L.M., and Lanza, S.T. (2009). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences.* Hoboken, NJ: John Wiley & Sons.

Colvin, G. (2004). *Defusing Anger and Aggression: Safe Strategies for Secondary School Educators* [Media Program]. Eugene, OR: IRIS Media.

Committee for Children. (1990) *Second Step: A Violence Prevention Curriculum. Grades 6–8*. Seattle, WA: Committee for Children.

Crick, N., and Dodge, K. (1994). A Review and Reformulation of Social-Information Processing Mechanisms in Children's Social Adjustment. *Psychological Bulletin, 115*: 74–101.

Crick, N.R., and Grotpeter, J.K. (1995). Relational Aggression, Gender, and Social-Psychological Adjustment. *Child Development*, *66*: 710–722.

References

DeVoe, J.F., Bauer, L., and Hill, M.R. (2010). *Student Victimization in U.S. Schools: Results From the 2007 School Crime Supplement to the National Crime Victimization Survey* (NCES 2010-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Dinkes, R., Kemp, J., and Baum, K. (2009). *Indicators of School Crime and Safety: 2009* (NCES 2010-012/NCJ 228478). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

DuRant, R.H., Cadenhead, C., Pendergrast, R.A., Slavens, G., and Linder, C.W. (1994). Factors Associated With the Use of Violence Among Urban Black Adolescents. *American Journal of Public Health, 84*(4): 612–617.

DuRant, R.H., Kahn, J., Beckford, P.H., and Woods, E.R. (1997). The Association of Weapon Carrying and Fighting on School Property and Other Health Risk and Problem Behaviors Among High School Students. *Archives of Pediatrics and Adolescent Medicine, 151*: 360–366.

DuRant, R.H., Krowchuk, D.P., Kreiter, S., Sinal, S.H., and Woods, C.R. (1999). Weapon Carrying on School Property Among Middle School Students. *Archives of Pediatrics and Adolescent Medicine, 153*(1): 21–26.

Engelmann, S.E.(1968). Relating Operant Techniques to Programming and Teaching. *Journal of School Psychology, 6,* 89–96.

Farrell, A.D., Kung, E.M., White, K.S., and Valois, R. (2000). The Structure of Self-Reported Aggression, Drug Use, and Delinquent Behaviors During Early Adolescence. *Journal of Clinical Child Psychology*, *29*: 282–292.

Farrell, A.D., and Meyer, A.L. (1997). The Effectiveness of a School-Based Curriculum for Reducing Violence Among Urban Sixth-Grade Students. *American Journal of Public Health*, *87*(6): 979–984.

Farrell, A.D., Meyer, A.L., Sullivan, T.N., and Kung, E.M. (2003). Evaluation of the Responding in Peaceful and Positive Ways (RiPP) Seventh Grade Violence Prevention Curriculum. *Journal of Child and Family Studies, 12*: 101–120.

Farrell, A.D., Meyer, A.L., and White, K.S. (2001). Evaluation of Responding in Peaceful and Positive Ways (RiPP): A School-Based Prevention Program for Reducing Violence Among Urban Adolescents. *Journal of Clinical Child Psychology, 30*: 451–463.

Farrell, A., Valois, R., and Meyer, A. (2002). Evaluation of the RiPP-6 Violence Prevention Program at a Rural Middle School. *American Journal of Health Education*, *33*(3): 167–172.

Farrell, A.D., Valois, R.F., Meyer, A.L., and Tidwell, R.P. (2003). Impact of the RiPP Violence Prevention Program on Rural Middle School Students. *Journal of Primary Prevention, 24*(2): 143–167.

Flay, B., and Allred, C. (2003). Long-term Effects of the Positive Action Program. *American Journal of Health Behavior, 27*: S6–S21.

Friedlander, B. (1993). Community Violence, Children's Development, and Mass Media: In Pursuit of New Insights, New Goals, and New Strategies. *Psychiatry, 56*: 66–81.

Hantman, I., and Crosse, S. (2000). *Progress in Prevention: National Study of Local Education Activities Under the Safe and Drug-Free Schools and Communities Act.* Washington, DC: U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service. Retrieved June 7, 2008, from http://www2.ed.gov/rschstat/eval/other/progressinprevention.pdf.

Horner, R., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A., and Esperanza, J. (2009). A Randomized, Wait-List Controlled Effectiveness Trial Assessing School-Wide Positive Behavior Support in Elementary Schools. *Journal of Positive Behavior Interventions, 11*: 133–144.

Huessmann, L. (1988). An Information Processing Model for the Development of Aggression. *Aggressive Behavior, 14*: 13–24.

Janega, J.B., Murray, D.M., Varnell, S.P., Blitstein, J.L., Birnbaum, A.S., and Lytle, L.A. (2004). Assessing the Most Powerful Analysis Method for Schools Intervention Studies With Alcohol, Tobacco, and Other Drug Outcomes. *Addictive Behaviors*, *29*(3): 595–606.

Jessor, R. (1985). Bridging Etiology and Prevention in Drug Abuse Research (DHHS Publication No. [ADM]85-1335). *NIDA Research Monograph, 56*: 257–268.

Kam, C.M., Greenberg, M.T., and Walls, C.T. (2003). Examining the Role of Implementation Quality in School-Based Prevention Using the PATHS Curriculum. *Prevention Science, 4*(1): 55–63.

Kelder, S.H., Orpinas, P., McAlister, A., Frankowski, R., Parcel, G.S., and Friday, J. (1996). The Students for Peace Project: A Comprehensive Violence-Prevention Program for Middle School Students. *American Journal of Preventive Medicine*, *12*: 22–30.

May, S., Ard, W., Todd, A., Horner, R., Glasgow, A., Sugai, G., and Sprague, J. (2010). *School-wide Information System* (version 4.3). Eugene, OR: University of Oregon.

McCulloch, C.E., and Searle, S.R. (2001). *Generalized, Linear and Mixed Models.* New York: Wiley-Interscience.

McCutcheon, A.C. (1987). *Latent Class Analysis.* Beverly Hills, CA: Sage.

Mendez Foundation. (1995). *Too Good for Violence K–8.* Tampa. FL: Author. Retrieved December 7, 2009, from http://www.mendezfoundation.org/home.php.

Metzler, C., Biglan, A., Rusby, J., and Sprague, J. (2001). Evaluation of a Comprehensive Behavior Management Program to Improve School-Wide Positive Behavior Support. *Education and Treatment of Children, 24*: 448–479.

Meyer, A.L., and Northup, W. (2002a). *RiPP: A Violence Prevention Curriculum for the Seventh Grade.* Ashland, VA: Authors.

Meyer, A.L., and Northup, W. (2002b). *RiPP: A Violence Prevention Curriculum for the Sixth Grade.* Ashland, VA: Authors.

References

Meyer, A.L., and Northup, W. (2006). *RiPP: A Violence Prevention Curriculum for the Eighth Grade.* Ashland, VA: Authors.

Miller, A.M. (2003). *Violence in U.S. Public Schools: 2000 School Survey on Crime and Safety* (NCES 2004–314 REVISED). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved June 7, 2008, from http://nces.ed.gov/pubs2004/2004314.pdf.

Muhr, T., and Friese, S. (2009). *ATLAS.ti 6.0.* Berlin: Scientific Software Development.

Multisite Violence Prevention Project. (2006). *Description of Measures: Cohort-Wide Student Survey.* Available from the Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Atlanta, GA. (Unpublished)

Murray, D.M. (1998). *Design and Analysis of Group-Randomized Trials.* New York: Oxford University Press.

Muthén, L.K., and Muthén, B.O. (1998–2007). *Mplus User's Guide.* Fifth Edition. Los Angeles, CA: Muthén and Muthén.

National Institute on Drug Abuse. (2003). *Preventing Drug Use Among Children and Adolescents: A Research-Based Guide for Parents, Educators, and Community Leaders* (NIH 04-4212(A)). Bethesda, MD: U.S. Department of Health and Human Services. Retrieved June 7, 2008, from http://www.drugabuse.gov/pdf/prevention/redbook.pdf.

Nunnally, J.C., and Bernstein, I.H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill, Inc.

Orpinas, P. (2009). *Measurement Manual: Description of Aggression, Victimization, and Social Skills Scales.* Retrieved April 14, 2011, from http://www.publichealth.uga.edu/hpb/sites/default/files/ORPINAS2009.pdf.

Orpinas, P., and Horne, A.M. (2006). *Bullying Prevention: Creating a Positive School Climate and Developing Social Competence.* Washington, DC: American Psychological Association.

Orpinas, P., Kelder, S., Frankowski, R., Murray, N., Zhang, Q., and McAlister, A. (2000). Outcome Evaluation of a Multi-Component Violence-Prevention Program for Middle Schools: The Students for Peace Project. *Health Education Research, 15*: 45–58.

Patterson, G.R. (1982). *Coercive Family Process.* Eugene, OR: Castalia.

Raudenbush, S.W., and Bryk, A.S. (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Remboldt, C. (1998). Making Violence Unacceptable. *Educational Leadership, 56*(1): 32–38.

Ringwalt, C.L., Pankratz, M.M., Hansen, W.B., Dusenbury, L., Jackson-Newsom, J., Giles, S.M., and Brodish, P.H. (2009). The Potential of Coaching as a Strategy to Improve the Effectiveness of School-Based Substance Use Prevention Curricula. *Health Education & Behavior, 36*(4): 696–710.

SAS Institute. (2004). *SAS Procedures Guide, On-line Version 9*. Cary, NC: SAS Institute, Inc.

SAS Institute. (2006). *The GLIMMIX Procedure*. Cary, NC: SAS Institute, Inc.

Shadish, W.R., Cook, T.D., and Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.

Silvia, S., Blitstein, J., Williams, Ringwalt, C., Dusenbury, L., and Hansen, W.B. (2010). *Impacts of a Violence Prevention Program for Middle Schools: Findings From the First Year of Implementation* (NCEE 2010-4007). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Simon, T.R., Sussman, S., Dahlberg, L.L., and Dent, C.W. (2002). Influence of a Substance-Abuse-Prevention Curriculum on Violence-Related Behavior. *American Journal of Health Behavior*, *26*(2): 103–110.

Slaby, R.G. (1989). *An Evaluation of a Violence Prevention Program. Health Program for Urban Youth*. Newton, MA: Education Development Center, Inc. (Unpublished)

Smith, S., and Sprague, J.R. (2006). *Systematic Supervision for High School: A Positive Way to Monitor Common Areas* [Media Program]. Eugene, OR: IRIS Media.

Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., and Li, K. (2010). Impact of a Social-Emotional and Character Development Program on School-Level Indicators of Academic Achievement, Absenteeism, and Disciplinary Outcomes: A Matched-Pair, Cluster Randomized, Controlled Trial. *Journal of Research on Educational Effectiveness, 3,* 26–55.

Sprague, J.R., and Golly, A. (2005). *Best Behavior: Building Positive Behavior Support in Schools*. Frederick, CO: Sopris West Educational Services.

Sprague, J.R., Sugai, G., and Walker, H. (1998). Antisocial Behavior in Schools. In T.S. Watson and F.M. Gresham (Eds.), *Handbook of Child Behavior Therapy* (pp. 451–474). New York: Plenum.

Sprague, J., Walker, H., Golly, A., Myers, D., and Shannon, T. (2001). Translating Research Into Effective Practice: The Effects of a Universal Staff and Student Intervention on Indicators of Discipline and School Safety. *Education and Treatment of Children, 24*: 495–511.

Sugai, G., and Horner, R. (1994). Including Students With Severe Behavior Problems in General Education Settings: Assumptions, Challenges, and Solutions. In J. Marr, G. Sugai, and G. Tindal (Eds.), *Oregon Conference Monograph*, *Vol. 6* (pp. 102–120). Eugene, OR: University of Oregon.

Thissen, D., Steinberg, L., and Kuang, D. (2002). Quick and Easy Implementation of the Benjamini-Hochberg Procedure for Controlling the False-Positive Rate in Multiple Comparisons. *Journal of Educational and Behavioral Statistics*, 27(1): 77–83.

U.S. Department of Education. (1998). Notice of Final Principles of Effectiveness. *Federal Register*, *63*(104): 29902–29906.

References

U.S. Department of Education, National Center for Education Statistics. (2004). Schools and Staffing Survey (SASS).

U.S. Department of Health and Human Services. (2001). *Youth Violence: A Report of the Surgeon General.* Rockville, MD: Author, Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, and National Institute of Mental Health. Retrieved June 7, 2008, from http://www.surgeongeneral.gov/library/youthviolence/default.htm.

U.S. Department of Justice, Bureau of Justice Statistics. (2003). School Crime Supplement (SCS) to the National Crime Victimization Survey.

Weiner, B., Graham, S., & Chandler, C. (1982). Pity, Anger, and Guilt: An Attributional Analysis. *Personality and Social Psychology Bulletin, 8*: 226–232.

Wilson, S.J., and Lipsey, M.W. (2005). *The Effectiveness of School-Based Violence Prevention Programs for Reducing Disruptive and Aggressive Behavior* (NCJ 211376). Rockville, MD: National Criminal Justice Reference Service. Retrieved June 7, 2008, from http://www.ncjrs.gov/pdffiles1/nij/grants/211376.pdf.

Appendix A:
Survey Instruments

This page intentionally left blank.

Student Survey

ID Label

General Instructions

- Use a #2 pencil only to complete the survey. Do NOT use ink or ballpoint pens.

- Make heavy black marks that fill out the circle completely.

CORRECT   **INCORRECT**

- If you are not sure, give the answer you think is best.

- If you make a mistake, please erase cleanly any wrong answer and completely black out the circle beside your correct answer choice.

- Do not make any stray marks of any kind anywhere in this booklet.

- DO NOT write your name anywhere on this booklet.

Part A. Your Background

These first questions ask some background information about you.

1. Are you a
   ○ Boy?   or   ○ Girl?

2. Are you Hispanic or Latino?
   ○ No   ○ Yes

3. What is your race? **Choose one or more**.

   ○ American Indian or Alaska Native

   ○ Asian

   ○ Black or African American

   ○ Native Hawaiian or other Pacific Islander

   ○ White

4. When is your birthday?

Month	Day	Year

   Jan ○
   Feb ○   0○ 0○   **19**   8○ 0○
   Mar ○   1○ 1○   9○ 1○
   Apr ○   2○ 2○   2○
   May ○   3○ 3○   3○
   Jun ○   4○   4○
   Jul ○   5○   5○
   Aug ○   6○   6○
   Sep ○   7○   7○
   Oct ○   8○   8○
   Nov ○   9○   9○
   Dec ○

5. Which of these people do you live with <u>most</u> of the time? **Choose all that apply**.

   ○ Mother

   ○ Father

   ○ Stepmother

   ○ Stepfather

   ○ Foster mother, female guardian

   ○ Foster father, male guardian

   ○ Sisters

   ○ Brothers

   ○ Children other than brothers or sisters

   ○ Grandparents

   ○ Other adults

Page 1 of 6

Part B. Your School

The following statements could describe a school. Think about your school over the PAST 30 DAYS when answering the next few questions. If you are not sure, give the answer you think is best.

	Strongly agree ▼	Agree ▼	Disagree ▼	Strongly disagree ▼
6. Everyone knows what the school rules are.	O	O	O	O
7. The school rules are strictly enforced.	O	O	O	O
8. If a school rule is broken, students know what kind of consequence will follow.	O	O	O	O
9. The punishment for breaking school rules is the same no matter who you are.	O	O	O	O
10. Students are complimented or rewarded when they follow the rules.	O	O	O	O
11. The rewards for following the rules are the same no matter who you are.	O	O	O	O
12. We get taught at school about getting along with others and about respecting them.	O	O	O	O
13. We get taught at school about avoiding and dealing with violent situations.	O	O	O	O
14. Teachers or other adults at my school try to prevent or stop bullying.	O	O	O	O
15. Teachers treat students with respect.	O	O	O	O

Part C. Getting Along with People at Your School

The next questions are about things other students from your school may have done to you in the PAST 30 DAYS.

In the PAST 30 DAYS, how often did OTHER STUDENTS FROM YOUR SCHOOL:	Never ▼	Once or twice ▼	Several times ▼	Often ▼
16. Say or do something nice to you	O	O	O	O
17. Say "thanks" or "you're welcome" to you	O	O	O	O
18. Say or do something that made you feel good	O	O	O	O
19. Invite you to participate in a game, group conversation, or a class activity	O	O	O	O
20. Say a compliment (praise, kind word) to you	O	O	O	O
21. Offer to help you	O	O	O	O
22. Share something with you	O	O	O	O
23. Act friendly with you	O	O	O	O
24. Show interest in your ideas or activities	O	O	O	O

In the PAST 30 DAYS, how often did OTHER STUDENTS FROM YOUR SCHOOL:	Never ▼	Once or twice ▼	Several times ▼	Often ▼
25. Threaten to hurt you by hitting, pushing, slapping, or shoving you	O	O	O	O
26. Actually hurt you by hitting, pushing, slapping, or shoving you	O	O	O	O
27. Threaten you (but not actually injure you) with a weapon such as a gun, knife, or club	O	O	O	O
28. Actually injure you with a weapon such as a gun, knife, or club	O	O	O	O
29. Yell at you when they were angry	O	O	O	O
30. Throw something at you to hurt you	O	O	O	O
31. Pick a fight with you	O	O	O	O
32. Take, damage, or destroy on purpose something that belonged to you	O	O	O	O
33. Try to force you to do something that you didn't want to do	O	O	O	O
34. Leave you out from a group or activity on purpose	O	O	O	O
35. Tell lies, spread rumors, or say mean things about you to other students	O	O	O	O
36. Call you an insulting name or word	O	O	O	O
37. Make fun of you in front of other people just to be mean	O	O	O	O

Page 3 of 6

The next questions ask about YOUR behaviors in the PAST 30 DAYS.

In the PAST 30 DAYS, how often did YOU:	Never ▼	Once or twice ▼	Several times ▼	Often ▼
38. Say or do something nice to a kid from your school	O	O	O	O
39. Say "thanks" or "you're welcome" to a kid from your school	O	O	O	O
40. Say or do something that made a kid from your school feel good	O	O	O	O
41. Invite a kid from your school to participate in a game, group conversation, or activity	O	O	O	O
42. Compliment (praise, say a kind word) a kid from your school	O	O	O	O
43. Offer to help a kid from your school	O	O	O	O
44. Share something with a kid from your school	O	O	O	O
45. Act friendly with a kid from your school	O	O	O	O
46. Show interest in the ideas or activities of a kid from your school	O	O	O	O

In the PAST 30 DAYS, how often did YOU:	Did not have the chance ▼	Never ▼	Once or twice ▼	Several times ▼	Often ▼
47. Stop someone from getting in a fight	O	O	O	O	O
48. Stand up for someone who was being "bullied"	O	O	O	O	O

Page 4 of 6

In the PAST 30 DAYS, how often did YOU do each of these things at school:	Never ▼	Once or twice ▼	Several times ▼	Often ▼
49. Threaten to hurt another student by hitting, pushing, slapping, or shoving them	○	○	○	○
50. Actually hurt another student by hitting, pushing, slapping, or shoving them	○	○	○	○
51. Bring a weapon such as a gun, knife, or club to school	○	○	○	○
52. Threaten (but not actually injure) another student with a weapon such as a knife, gun, or club	○	○	○	○
53. Actually injure another student with a weapon such as a knife, gun, or club	○	○	○	○
54. Get angry and yell at another student	○	○	○	○
55. Throw something at another student to hurt him or her	○	○	○	○
56. Pick a fight with another student	○	○	○	○
57. Take, damage, or destroy on purpose something that belonged to another student	○	○	○	○
58. Try to force another student to do something they didn't want to do	○	○	○	○
59. Leave out another student on purpose from your group or activity	○	○	○	○
60. Tell lies, spread rumors, or say mean things about someone	○	○	○	○
61. Call another student an insulting name or word to be mean	○	○	○	○
62. Make fun of another student in front of him or her just to be mean	○	○	○	○

The next questions ask about how safe you feel at school.

	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
63. How often do you worry that someone from your school will attack or hurt you?	○	○	○	○
64. How often do you worry that someone from your school will bully you?	○	○	○	○

Part D. Your Feelings and Attitudes

The next few questions ask about YOUR feelings and how you get along in general.

Please choose the answer that best describes how much you agree or disagree with each of the following statements.

	Strongly agree ▼	Agree ▼	Disagree ▼	Strongly disagree ▼
65. If I walked away from a fight, I'd be a coward ("chicken")	◯	◯	◯	◯
66. Anyone who won't fight is going to be "picked on" even more	◯	◯	◯	◯
67. I don't need to fight because there are other ways to deal with being mad	◯	◯	◯	◯
68. It's OK to hit someone who hits you first	◯	◯	◯	◯
69. If a kid teases me or "disses" me, I usually cannot get them to stop unless I hit them	◯	◯	◯	◯
70. Sometimes you have to physically fight to get what you want	◯	◯	◯	◯
71. Some kids deserve to be picked on or bullied	◯	◯	◯	◯
72. It's OK to spread gossip about someone to get even with them	◯	◯	◯	◯

The next time you find yourself really angry at someone or about something, how likely is it that YOU would…	Very likely ▼	Likely ▼	Unlikely ▼	Very unlikely ▼
73. Walk away or ignore the situation or person	◯	◯	◯	◯
74. Try to talk it out with the other person	◯	◯	◯	◯
75. Do something else to get your mind off of it	◯	◯	◯	◯
76. Laugh it off	◯	◯	◯	◯
77. Try to see the other person's point of view	◯	◯	◯	◯
78. Yell at the person	◯	◯	◯	◯
79. Break something	◯	◯	◯	◯
80. Hit or threaten to hurt the person	◯	◯	◯	◯
81. Try calming yourself down	◯	◯	◯	◯
82. Go talk with a friend	◯	◯	◯	◯
83. Apologize to the other person	◯	◯	◯	◯
84. Get help from a teacher or other adult	◯	◯	◯	◯

Page 6 of 6

Teacher Survey

WHO IS CONDUCTING THIS SURVEY?
The U.S. Department of Education requests your participation in this survey. RTI International is conducting this survey for the Department of Education.

WHAT IS THE PURPOSE OF THIS SURVEY?
The purpose of this survey is to gather information about teachers' experiences concerning the learning environment in the classroom and school. The survey also asks about your perceptions of safety in and around the school and experiences with student misconduct. Your answers on this survey will help describe the school environment and climate.

IS THIS SURVEY CONFIDENTIAL?
Please do not put your name on this survey. Any information you provide will be kept confidential by RTI and other project research staff. School staff will not see your responses nor will they know if you have or have not agreed to complete the survey. Surveys will be labeled with a bar code, and your name will not appear on the completed survey.

WHO SHOULD COMPLETE THIS SURVEY?
This survey should be completed by teachers in grades 6 to 8.

TO WHOM SHOULD YOU GIVE YOUR COMPLETED SURVEY?
The school contact will collect the completed surveys during the week of your school's student survey administration. Please put your completed survey in the accompanying envelope.

Thank you for completing this survey.

General Instructions

- Use a #2 pencil only to complete the survey. Do NOT use ink or ballpoint pens.

- Make heavy black marks that fill out the circle completely.

CORRECT	**INCORRECT**

- If you make a mistake, please erase cleanly any wrong answer and completely black out the circle beside your correct answer choice.

- Do not make any stray marks of any kind anywhere in this booklet.

- DO NOT write your name anywhere on this booklet.

During the PAST 30 DAYS at the school where you work, how often did YOU FEEL UNSAFE in any of the following areas?

	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
1. The entrance into the school	O	O	O	O
2. Any hallways or stairs in the school	O	O	O	O
3. Any part of the school cafeteria	O	O	O	O
4. Any school restroom	O	O	O	O
5. In any classroom	O	O	O	O
6. School parking lot, athletic fields, or other places outside school buildings	O	O	O	O

During the PAST 30 DAYS, how often did YOU WITNESS the following events at your school?

	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
7. A student threaten to hit, push, slap, or shove another student to hurt him or her	O	O	O	O
8. A student actually hit, slap, shove, or push another student to hurt him or her	O	O	O	O
9. A student threaten (but not actually injure) another student with a weapon such as a knife, gun, or club	O	O	O	O
10. A student actually injure another student with a weapon such as a knife, gun, or club	O	O	O	O
11. A student get angry and yell at another student	O	O	O	O
12. A student throw something at another student to hurt him or her	O	O	O	O
13. A student pick a fight with another student	O	O	O	O
14. A student take, damage, or destroy on purpose something that belonged to another student	O	O	O	O
15. A student try to force another student to do something he or she didn't want to do	O	O	O	O
16. A student leave out another student on purpose from a group or activity	O	O	O	O

	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
17. A student tell lies, spread rumors, or say mean things about someone to other students	O	O	O	O
18. A student call another student an insulting name or word	O	O	O	O
19. A student make fun of another student in front of him or her just to be mean	O	O	O	O
20. A student sexually harass another student	O	O	O	O
21. A student disrupt my class due to misbehavior	O	O	O	O

How often in the PAST SIX MONTHS have YOU been…?

	Never ▼	Once ▼	2 to 5 times ▼	More than 5 times ▼
22. Verbally abused by a student from your school	O	O	O	O
23. Threatened with physical harm by a student from your school	O	O	O	O
24. Physically attacked or injured by a student from your school	O	O	O	O

Think about what happened in your school during the PAST 30 DAYS, when you answer these questions.

In the PAST 30 DAYS, how often did you SEE OR HEAR A STUDENT:	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
25. Say or do something nice to another student	O	O	O	O
26. Say "thanks" or "you're welcome" to another student	O	O	O	O
27. Say or do something that made another student feel good	O	O	O	O
28. Invite another student to participate in a game, group conversation, or a class activity	O	O	O	O
29. Say a compliment (praise, kind word) to another student	O	O	O	O
30. Offer to help another student	O	O	O	O

Impacts of a Violence Prevention Program for Middle Schools

	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
31. Share something with another student	O	O	O	O
32. Act friendly with another student	O	O	O	O
33. Show positive interest in another student's ideas or activities	O	O	O	O
34. Stop someone from getting in a fight	O	O	O	O
35. Stand up for someone who was being "bullied"	O	O	O	O
36. Voluntarily apologize to another student	O	O	O	O

The following questions are intended to indicate how often you use certain techniques and resources with <u>aggressors</u> (students who display aggressive behaviors) and <u>victims</u> or <u>targets</u> (students who are the recipients of that aggressive behavior). Please complete every item by choosing the response that most closely reflects YOUR USE of the intervention or approach during the past 30 days.

In the PAST 30 DAYS, how often did you…?	Did not have the opportunity ▼	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
37. Model strategies for solving conflicts	O	O	O	O	O
38. Create an "open door" policy for students who are the target of aggression	O	O	O	O	O
39. Include victimized or isolated children in group projects	O	O	O	O	O
40. Use classroom routines that reduce the opportunity for acting out behaviors	O	O	O	O	O
41. Reward small improvements toward desired behavior	O	O	O	O	O
42. Use a behavior plan that provides students choices and consequences for their choices	O	O	O	O	O
43. Provide opportunities for students to confidentially report aggressive acts	O	O	O	O	O
44. Assist students victimized by aggressive peers in identifying skills and behaviors they can use in these situations	O	O	O	O	O
45. Address aggressive situations in the classroom immediately	O	O	O	O	O

Page 3 of 5

In the PAST 30 DAYS, how often did you…?	Did not have the opportunity ▼	Never ▼	Almost never ▼	Sometimes ▼	Often ▼
46. Maintain calmness when faced with an aggressive or disruptive student	O	O	O	O	O
47. Confront students who make inappropriate comments	O	O	O	O	O
48. Consult with school administrators for support	O	O	O	O	O
49. Use self-calming techniques during the school day	O	O	O	O	O
50. Model dignity and respect at school	O	O	O	O	O
51. Provide positive reinforcement for prosocial behavior	O	O	O	O	O

The following is a list of statements that could describe a school. Think about the current situation in your school when responding to the following statements.

	Strongly agree ▼	Agree ▼	Disagree ▼	Strongly disagree ▼
52. The school rules for student behavior are clearly defined	O	O	O	O
53. The school makes sure that students know the rules for student behavior	O	O	O	O
54. The school rules emphasize reinforcing desired behavior	O	O	O	O
55. The school rules emphasize consequences for undesired behavior	O	O	O	O
56. When a school rule is broken, it is clear to the school staff what consequences should follow	O	O	O	O
57. Teachers at my school consistently enforce the rules	O	O	O	O
58. Teachers at my school punish students the same way for breaking the same rule no matter who the students are	O	O	O	O
59. Teachers compliment or reward students when they follow the rules	O	O	O	O
60. Teachers at my school reward students the same way for following the same rule no matter who the students are	O	O	O	O

Page 4 of 5

	Strongly agree ▼	Agree ▼	Disagree ▼	Strongly disagree ▼
61. Teachers or other adults at the school try to prevent or stop bullying	○	○	○	○
62. Teachers treat students with respect	○	○	○	○
63. Teachers know the procedure for reporting bullying and violence	○	○	○	○
64. Administrators and/or teachers periodically review school rules to determine if they need modification	○	○	○	○
65. Administrators at my school are supportive of teachers in creating a safe school	○	○	○	○
66. Teachers receive adequate training in classroom management/discipline strategies	○	○	○	○
67. Administrators at my school consistently enforce the rules	○	○	○	○

This last set of questions asks about your background.

68. Are you:

    ○ Male   ○ Female

69. How long have you been teaching (at this school or any other school)?

    ○ less than 5 years

    ○ 5 to 10 years

    ○ 11 to 20 years

    ○ more than 20 years

70. What is the highest level of education you completed?

    ○ High school graduate

    ○ Some college, no degree

    ○ Associates or 2-year degree

    ○ Bachelor's or 4-year degree

    ○ Master's degree

    ○ Doctoral degree

71. Did you teach the RiPP curriculum this year?

    a. Yes

    b. No/not applicable

72. Were you a member of the Best Behavior team this year?

    a. Yes

    b. No/not applicable

Appendix B:
Defining the High-Risk Student Subgroup

Identification of high-risk students used two complementary strategies. First, risk was estimated as a function of actual perpetration of violence. Second, students were classified as high risk based on profiles of attitudes and intentions toward violence.

B.1   High-Risk Classification Based on Overt Behaviors

Identification of high-risk students using overt acts of violence and aggression was based on 8 items from the full set of 14 perpetration items in the baseline student survey. These 8 items were determined to be the most severe items, based on analyses using item response theory (IRT). The IRT rank column in table B-1 displays these items' severity rank, with higher ranking equaling greater severity. In the IRT modeling context, severity is analogous to item difficulty, or the value of the latent factor or dimension (here, a measure of violence perpetration identified by all 14 survey items) where there is a 50 percent probability of a positive response to that item.

Table B-1.   Item response theory (IRT) severity rankings for the eight most severe perpetration items

Item	Description	IRT severity rank
Q56	Pick a fight with another student	7
Q50	Actually hurt by hitting, pushing, slapping, or shoving	8
Q55	Throw something at another student to hurt them	9
Q57	Take, damage, or destroy on purpose something that belonged to another student	10
Q58	Try to force another student to do something they did not want to do	11
Q52	Threaten another student with a weapon such as a knife, gun, or club	12
Q51	Bring a weapon such as a knife, gun, or club to school	13
Q53	Actually injure another student with a weapon such as a knife, gun, or club	14

SOURCE: Student survey, fall 2006 (baseline).

Students were placed in the high-risk group if they reported committing any of these eight most severe acts one or more times in the past 30 days. Table B-2 shows the frequency of responding to one or more of these items. In total, approximately 28 percent (2,135) of students were deemed to be high risk by this criterion.

Table B-2.   Frequency of student responses to eight most severe perpetration items

Number of perpetration items	Number of students reporting	Percent of students reporting
0	5,466	71.91
1	1,002	13.18
2	457	6.01
3	268	3.53
4	148	1.95
5	111	1.46
6	77	1.01
7	42	0.55
8	30	0.39

SOURCE: Student survey, fall 2006 (baseline).

B.1   High-Risk Classification Based on Attitudes and Intentions

The second strategy employed was to examine more proximal student measures that, while not directly assessing violence or perpetration, were believed to be related to such acts. These constructs theoretically serve as mediators of program effects and are thus relevant for identifying risk—successful changes to these outcomes are believed to then change behavioral outcomes for the better. Theoretically derived mediators believed to be important for risk identification were the following:

- attitudes endorsing violence (items Q65–Q72);

- reactions to anger—positive (items Q73–Q77, Q81–Q84); and

- reactions to anger—negative (items Q78–Q80).

As detailed previously, factor analyses indicated unidimensionality for attitudes endorsing violence items. Behavioral intentions or reactions to anger indicated a two-factor structure with separate dimensions for positive and negative reactions.

To estimate risk using the three dimensions above, scale scores were first computed (as the mean of all component items), and these scale scores were then entered into latent class analysis (LCA) to determine unobserved heterogeneity of response patterns within the sample. LCA is a latent variable technique similar to factor analysis. However, whereas factor analysis attempts to find common dimensions that group items, LCA attempts to find profiles of responses that group respondents.

A five-class solution emerged, indicating that students typically had one of the five response patterns shown in figure B-1.

Figure B-1. Latent class–based profiles of student responses to violence-related constructs
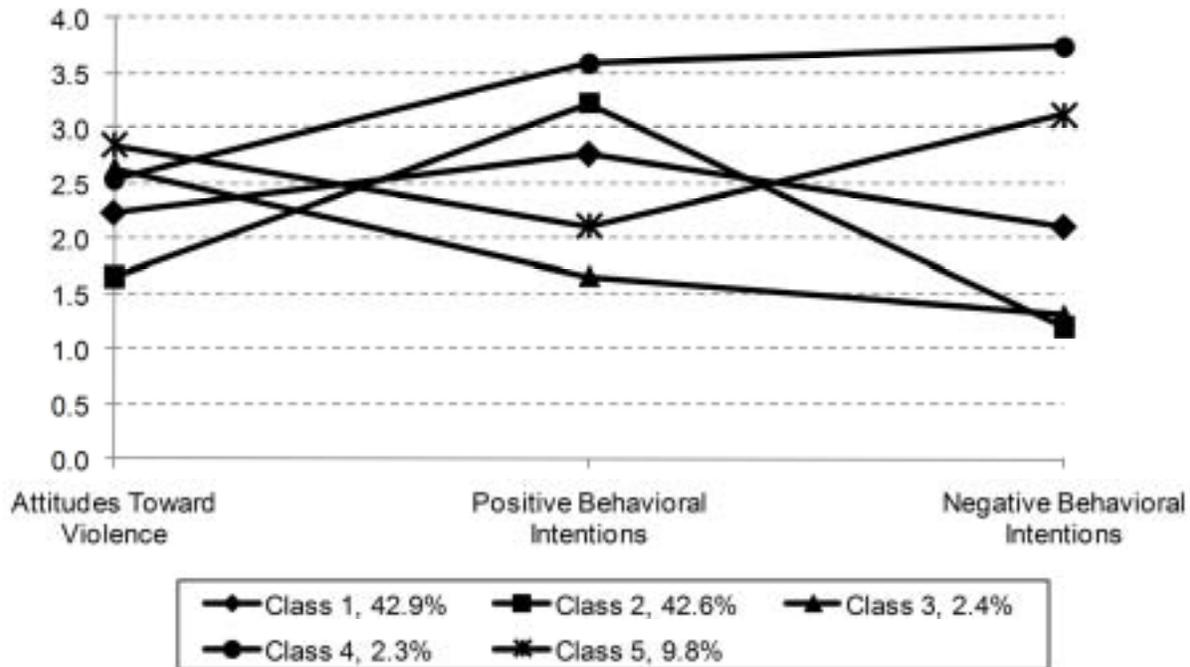
Class 5 represented the most clearly high-risk profile, with the greatest level of attitudes endorsing violence, low positive behavioral intentions when angered, and high negative behavioral intentions. This class placed approximately 10 percent of the total sample in the high-risk group.

B.3 Total High-Risk Sample

Combining criteria (high-risk based on perpetration, or high-risk based on attitudes and intentions) yields the breakdown in table B-3. Among the 7,601 students in the total sample at baseline, 5,212 (68.55 percent) were classified as not high risk, and 2,391 (31.45 percent) were classified as high-risk students. The high-risk total included 2,135 students (28 percent of the total sample) identified as high risk based on perpetration and 256 students (3 percent of the total sample) identified based on profiles of attitudes and intentions toward violence. Only about 20 percent of those classified as high risk were classified according to both criteria (466 out of 2,391 students). These students were counted in the perpetration subclassification for analysis purposes.

Table B-3.   High-risk classification using joint criteria

Classification	Number of students	Percent of total sample
Total not high risk	5,212	68.55
Total high risk	2,391	31.45
High risk, based on perpetration	2,135	28.09
High risk, perpetration only	1,669	21.95
High risk, both perpetration and attitudes/intentions	466	6.13
High risk, based on attitudes and intentions only	256	3.37

SOURCE: Student survey, fall 2006 (baseline).

The goal at baseline was to identify an average of 54 high-risk students per school and establish a cohort of 2,160 high-risk youth that would be tracked across the intervention period. Allowing for nonparticipation and loss to follow-up, this would provide a cohort of approximately 1,440 high-risk students at the 3-year follow-up data collection. Not all high-risk students identified at baseline could be followed for the longitudinal high-risk sample, based on budget considerations. Accordingly, the researchers randomly selected 60 students in nine of the larger schools, in which more than 60 students were classified as high risk at baseline. This resulted in a final sample of 2,153 high-risk students (28 percent of the total sample), of which 1,923 students (25 percent of the total sample) were classified as high-risk based on perpetration and 230 students (3 percent of the total sample) were classified as high-risk based on profiles of attitudes and intentions toward violence but not perpetration.

Table B-4 provides self-reported demographic data for the high-risk subgroup and the total sample at baseline. Differences between the two groups were not tested for statistical significance. Over 30 percent of students in the high-risk sample were Hispanic, 29 percent were Black, and 26 percent were White. Less than one-half of the students (42 percent) were male, and just over 54 percent lived in single-adult households. Among students in the total sample, 37 percent were Hispanic, 32 percent were White, and 22 percent were Black. Male students made up 51 percent of the total sample, and 59 percent of students lived in single-adult households.

Table B-4. **Baseline demographic characteristics of the high-risk subgroup and the total sample in schools that were part of the study in the first year**

Characteristic	Total sample	High-risk subgroup
Sample size	7,601	2,153
Race/ethnicity (%)		
Hispanic	36.86	34.82
Black, non-Hispanic	22.08	28.90
White, non-Hispanic	31.62	25.99
Other, non-Hispanic, or mixed[1]	9.44	10.28
Gender (%)		
Male	50.62	41.80
Single-adult household (%)	59.48	54.39

[1] This category includes American Indian or Alaskan Native, Asian, Native Hawaiian/other Pacific Islander, and multiracial.
SOURCE: Student survey, fall 2006 (baseline).

Appendix C:
Construction of Outcome Measures

As discussed in chapter 2, the two sources of data for outcome measures were the student and teacher surveys. In this appendix, we provide additional details on the construction of the surveys and the construction of the measures. Copies of the survey instruments are included in appendix A.

C.1   Construction of the Student and Teacher Surveys

As described in chapter 2, the student and teacher surveys were both constructed to reflect the domains and outcome measures outlined in the logic model. As shown in table C-1, the scales were, for the most part, constructed or adapted from existing instruments.

Table C-1.   Sources of outcome measures for the student and teacher surveys

Outcome measure	Description	Instrument type	Source	# items in scale
Violence: Overall	Past 30 days: Threats or actual violence, at school	Student survey	Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000)	14
Violence: Weapons-related	Past 30 days: Threats or actual violence, at school, with a weapon (e.g., *actually injure another student with a weapon such as a gun, knife, or club*)	Student survey	Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000)	3
Violence: Not weapons-related	Past 30 days: Threats or actual violence, at school, without a weapon (e.g., *threaten to hurt another student by hitting, pushing, slapping, or shoving them*)	Student survey	Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000)	11
Student victimization: Overall	Past 30 days: Threats or actual victimization, at school	Student survey	Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000)	13
Student victimization: Overt	Past 30 days: Threats or actual victimization, at school, by direct means (e.g., *pick a fight with you*)	Student survey	Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000)	9

continued

Table C-1. Sources of outcome measures for the student and teacher surveys—Continued

Outcome measure	Description	Instrument type	Source	# items in scale
Student victimization: Relational	Past 30 days: Threats or actual victimization, at school, by indirect or social means (e.g., *leave you out from a group or activity on purpose*)	Student survey	Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000)	4
Student safety concerns	Student concerns about attacks or bullying at school	Student survey	Adapted from the School Crime Supplement to the National Crime Victimization Survey (U.S. Department of Justice [DOJ] 2003)	2
Student behavioral expectations	Student perception of the enforcement of school rules and consequences for misbehavior	Student survey	Adapted from the School Crime Supplement to the National Crime Victimization Survey (DOJ 2003)	10
Prosocial behaviors—Extended to others	Past 30 days: Acts of prosocial behavior extended to others (e.g., *say or do something nice to a kid from your school*)	Student survey	Adapted from the Positive Behavior Scale (Orpinas 2009)	9
Prosocial behaviors—Received from others	Past 30 days: Acts of prosocial behavior received from others (e.g., *say or do something nice to you*)	Student survey	Adapted from the Positive Behavior Scale (Orpinas 2009)	9
Prosocial behaviors—Active intervention	Past 30 days: Acts of prosocial behavior requiring active intervention on the part of the student (e.g., *stand up for someone who was being "bullied"*)	Student survey	Developed for this study	2

continued

Table C-1. Sources of outcome measures for the student and teacher surveys—Continued

Outcome measure	Description	Instrument type	Source	# items in scale
Student self-reported coping strategies	Student positive or negative reactions when angered	Student survey	Adapted from the Self-Efficacy for Alternatives to Aggression Scale (Multisite Violence Prevention Project 2006).	12
Student attitudes toward violence	Student positive or negative attitudes toward resolving conflict	Student survey	Adapted from Attitude Toward Interpersonal Violence Scale (Slaby 1989)	8
Teacher victimization	Past 6 months: Verbal or physical threats or attacks (e.g., *been threatened with physical harm by a student from your school*)	Teacher survey	Adapted from the Schools and Staffing Survey (SASS): 2003–2004 (U.S. Department of Education 2004)	3
Teacher safety concerns	Past 30 days: Safety concerns about specific areas in and around the school	Teacher survey	Adapted from the School Crime Supplement to the National Crime Victimization Survey (DOJ 2003)	6
School consistency of enforcing rules	Understanding and clarity of school policies and rules	Teacher survey	Adapted from the Teacher Survey (Multisite Violence Prevention Project 2006); additional items were developed for this study	16
Interactions with victims	Past 30 days: Interactions with victims in various ways	Teacher survey	Adapted from the Teacher Survey (Multisite Violence Prevention Project 2006)	4
Interactions with aggressors	Past 30 days: Interactions with aggressors in various ways	Teacher survey	Adapted from the Teacher Survey (Multisite Violence Prevention Project 2006)	4
Classroom management techniques	Past 30 days: Use of various classroom management techniques	Teacher survey	Adapted from the Teacher Survey (Multisite Violence Prevention Project 2006)	5

C.2   Psychometrics for Outcome Measures

The student survey contained indices of violence and victimization and measures of a number of secondary outcomes, such as beliefs about violence. The teacher survey included indices of teacher experiences of violence and safety at school. Prior to estimation of the intervention's effects, it was critical to ascertain the relevant dimensions for each measure, the quality of measurement or reliability of each dimension, and the optimal scoring approach for composites.

C.2.1 Factor Analysis

The first psychometric task was to identify the dimensionality of the data for each domain specified in the student survey instrument. Subject domains were specified by subsections of the survey, and therefore, we conducted exploratory factor analysis on the full set of items for each section on the survey. For example, "Part D. Your Feelings and Attitudes" contains eight items in a subsection that assesses attitudes regarding violence (e.g., "Anyone who won't fight is going to be picked on even more").

Exploratory factor analyses were conducted in Mplus version 4 (Muthén and Muthén 1998–2007) using oblique factor rotation. We attempted to estimate models with 1 to f factors for each set of items, where f was the number of items. In practice, however, the actual number of factors that could be extracted from the data without problems in estimation was usually less than f. In choosing the optimal number of dimensions or factors, we were guided by three considerations. The most weight in determining the optimal number of factors was given to inspection of the screen plot of the eigenvalues. The point at which there is a bend or elbow in the plot where all subsequent eigenvalues array in a fairly level plain indicates how many factors are present in the data. This criterion was compared to the common "eigenvalues greater than 1" rule of thumb, which can miss important dimensions in the data or suggest too many factors when several smaller factors load on a more general factor. In cases where these methods diverged, the third consideration—substantive clarity—was used to choose the most interpretively clear array of items to dimensions or factors. In addition to examining the factor structure, this stage of psychometric analysis was used to identify items that did not appear to load cleanly or well on any dimension and that potentially could be discarded from composite measures. Only two items appeared to be poor indicators of the latent factor representing the other items in their survey subheading group. These were two items measuring attitudes about violence: "Anyone who won't fight is going to be picked on even more," and "I don't need to fight because there are other ways to deal with being mad." These items were dropped from the attitudes about violence measure.

The teacher survey data were examined using the same methods. Table C-2 shows the items, mean or event rate, range and Cronbach α for the primary student outcomes. Secondary student measures are shown in table C-3, while the teacher survey measures are shown in table C-4.

C.2.2 Measurement Scale

Items measuring violence, especially the assessment of violence perpetration and the experience of aggressive acts by peers, often show distributional properties that do not match the assumptions of normality on which most analytic methods and scaling procedures (e.g., means, sum scores) depend. Creating scales from noncontinuous or nonnormal constituent items can severely limit power, increase type 1 error, or have other unintended consequences.

Table C-2.   Student survey: Outcomes addressing key research questions

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
Student violence: Overall	Past 30 days: Threats or actual violence, at school	Count	1.98 (0–14)	14	0.84
Student violence: Weapons-related	In the PAST 30 DAYS, how often did YOU: -Bring a weapon to school -Threaten with a weapon (knife, club, or gun) -Actually injure with a weapon (knife, club, or gun)	Count	0.09 (0–3)	3	0.72
Student violence: Not weapons-related	In the PAST 30 DAYS, how often did YOU: -Threaten to hurt others by hitting, punching, slapping, or shoving -Actually hurt others by hitting, punching, slapping, or shoving -Get angry and yell at others -Throw something at others -Pick a fight -Take, damage, or destroy someone's property -Try to force others to do something against their will -Leave others out from a group or activity on purpose -Tell lies, spread rumors, or say mean things -Call others an insulting name or word -Make fun of others	Count	1.90 (0–11)	11	0.90
Student victimization: Overall	Past 30 days: Threats or actual victimization, at school	Count	4.00 (0–13)	13	0.89
Student victimization: Overt	In the PAST 30 DAYS, how often did OTHER STUDENTS AT YOUR SCHOOL: -Threaten to hurt by hitting, punching, slapping, or shoving -Actually hurt by hitting, punching, slapping, or shoving -Threaten with a weapon (knife, club, or gun) -Actually injure with a weapon (knife, club, or gun) -Yell at you when angry -Throw something at you to hurt you -Pick a fight -Take, damage, or destroy your property -Try to force you to do something against your will	Count	2.26 (0–9)	9	0.84

See notes at end of table.

Table C-2. Student survey: Outcomes addressing key research questions—Continued

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
Student victimization: Relational	In the past 30 days, how often did OTHER STUDENTS AT YOUR SCHOOL:	Count	1.74 (0–4)	4	0.84
	-Leave you out from a group or activity on purpose				
	-Tell lies, spread rumors, or say mean things				
	-Call you an insulting name or word				
	-Make fun of you in front of others				

[1]Alpha coefficients for student outcomes were calculated based on a sample of 7,351 students in the 6th grade.

SOURCE: Student survey, fall 2006.

Table C-3. Student survey: Other outcome measures

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
Student safety concerns	The next questions ask about how safe you feel at school: -How often do you worry that someone from your school will attack or hurt you -How often do you worry that someone from your school will bully you	Continuous: 4-point scale, range "never" to "often"	1.92 (1–4)	2	†
Prosocial behaviors: Extended to others	In the PAST 30 DAYS, how often did YOU: -Say/do something nice to a student at school -Say "thanks" or "you're welcome" to a student at school -Say/do something that made a student feel good -Invite a student to participate in a game, conversation, activity -Compliment another student -Offer to help a student -Share something -Act friendly with another student -Show interest in the ideas/activities of another student	Continuous: 4-point scale, range "never" to "often"	2.93 (1–4)	9	0.90
Prosocial behaviors: Received from others	In the PAST 30 DAYS, how often did OTHER STUDENTS AT YOUR SCHOOL: -Say or do something nice to you -Say "thanks" or "you're welcome" to you -Say or do something that made you feel good -Invite you to participate in a game, group conversation, or class activity -Say a compliment (praise, kind word) to you -Offer to help you -Act friendly with you -Show interest in your ideas or activities	Continuous: 4-point scale, range "never" to "often"	2.50 (1–5)	9	0.86

See notes at end of table.

Table C-3. Student survey: Other outcome measures—Continued

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
Perceived behavioral expectations	The following statements could describe a school. Think about your school over the PAST 30 DAYS: -Everyone knows the rules -The school rules are strictly enforced -If a rule is broken, students know the consequences -Punishments are the same regardless of who you are -Students are complimented or rewarded for following rules	Continuous: 4-point scale; range "strongly disagree" to "strongly agree"	3.16 (1–4)	10	0.74
Perceived behavioral expectations— Continued	-Rewards are the same for following the rules, regardless of who you are -We are taught at school about getting along and respecting others -We are taught at school about avoiding and dealing with violent situations -Teachers/other adults at school try to prevent/stop bullying -Teachers treat students with respect				
Student self-reported coping strategies: Positive	The next time you find yourself really angry at someone or about something, how likely is it that YOU would: -Walk away or ignore the situation/person -Try to talk it out with another person -Do something else to get your mind off it -Laugh it off -Try to see the other person's point of view -Try calming yourself down -Go talk with a friend -Apologize to the other person -Get help from a teacher or other adult	Continuous: 4-point scale; range "very unlikely" to "very likely"	2.89 (1–4)	9	0.84

See notes at end of table.

Table C-3. Student survey: Other outcome measures—Continued

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
Student self-reported coping strategies: Negative	The next time you find yourself really angry at someone or about something, how likely is it that YOU would: -Yell at the person -Break something -Hit or threaten to hurt the person	Count	0.60 (0–3)	3	0.77
Student attitudes toward violence[2]	The next few questions ask about your feelings and how you get along in general: -If I walked away from a fight, I would be called "chicken" -It's OK to hit someone if that person hits you first -If a kid teases/picks on me, I cannot get that person to stop unless I hit that person -Sometimes you have to physically fight to get what you want -Some kids deserve to be picked on or bullied -It's OK to spread gossip to get even	Continuous: 4-point scale; range "strongly agree" to "strongly disagree"	3.0 (1–4)	6	0.78

† Not applicable.

[1] Alpha coefficients for student outcomes were calculated based on a sample of 7,351 students in the 6th grade.

[2] Two items, 66 and 67, were dropped from the scale because they were found to be unrelated to the other items in the scale.

SOURCE: Student survey, fall 2006.

Table C-4.   Teacher survey: Other outcome measures

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
Teacher victimization	How often in the PAST 6 MONTHS have YOU been: -Verbally abused by a student at school -Threatened with physical harm by a student -Physically attacked or injured by a student	Dichotomous	0.65 (0–1)	3	0.68
Teacher safety concerns	During the PAST 30 days at the school where you work, how often did YOU FEEL UNSAFE in the following areas: -The entrance to the school -Any hallway or stairs -Any part of the school cafeteria -Any school restroom -Any classroom -The school parking lot, fields, or other places outside the school	Continuous: 4-point scale; range "never" to "often"	1.44 (1–4)	6	0.89
Teacher interactions with victims	In the PAST 30 DAYS, how often did you: -Model strategies for solving conflicts -Create an open-door policy for students who are targets of violence -Include victimized or isolated children in group activities -Assist victimized students in identifying skills and behaviors they can use in these situations	Continuous: 4-point scale; range "never" to "often"	3.95 (1–5)	4	0.75
Teacher interactions with aggressors	In the PAST 30 DAYS, how often did you: -Address aggressive situations in the classroom immediately -Maintain calmness when faced with an aggressive or disruptive student -Confront students who make inappropriate comments -Consult with school administrators for support	Continuous: 4-point scale; range "never" to "often"	4.42 (1–5)	4	0.71
Classroom management techniques	In the PAST 30 DAYS, how often did you: -Use classroom routines that reduce the opportunity for acting out behaviors -Reward small improvements toward desired behavior -Use a behavior plan that provides students with choices and consequences for their choices -Model dignity and respect at school -Provide positive reinforcement for prosocial behavior	Continuous: 4-point scale; range "never" to "often"	4.57 (1–5)	5	0.68

See notes at end of table.

Table C-4. Teacher survey: Other outcome measures—Continued

Outcome measure	Items	Variable type	Mean or event rate (range)	# items in scale	Cronbach α[1]
School consistency of enforcing behavioral rules	Statements that could describe your school: -The school rules for student behavior are clearly defined -The school makes sure that students know the rules for student behavior -The school rules emphasize reinforcing desired behavior -The school rules emphasize consequences for undesired behavior -When a school rule is broken, it is clear to the school staff what consequences should follow -Teachers at my school consistently enforce the rules -Teachers at my school punish students the same way for breaking the same rule, no matter who the students are -Teachers compliment or reward students when they follow the rules -Teachers at my school reward students the same way for following the same rule, no matter who the students are -Teachers or other adults at the school try to prevent or stop bullying -Teachers treat students with respect -Teachers know the procedure for reporting bullying and violence -Administrators and/or teachers periodically review school rules to determine if they need modification -Administrators at my school are supportive of teachers in creating a safe school -Teachers receive adequate training in classroom management/discipline strategies -Administrators at my school consistently enforce the rules	Continuous: 4-point scale; range "strongly disagree" to "strongly agree"	2.85 (1.07–4)	16	0.92

[1] Alpha coefficients for teacher outcomes were calculated based on a sample of 917 middle school teachers.

SOURCE: Teacher survey, spring 2007.

To avoid these problems, a series of latent variable measurement models were compared to determine the optimal method for obtaining satisfactory composites of relevant dimensions in the data. Confirmatory factor analysis (CFA) was used to estimate three measurement models for each dimension identified above. The first model represented the CFA version of the exploratory models and served as a general test of the stability of each factor or dimension. For some of the secondary outcomes in which responses were set to a Likert scale, this model was also expected to be the most appropriate. However, many of the survey items are count type items, or suggestive of a count, and thus, the distributions are quite skewed. For items that were not believed to be continuous, the second and third models represented opposing ends of the measurement spectrum. The second model treated responses as ordered categorical, or polytomous, items, and model estimation took this response type into account by using a weighted least squares estimator for categorical data. This model represented the most statistically rigorous measurement model, in that it accounted for variation in frequency of response and did not weight items identically (i.e., factor loadings were unconstrained). The third measurement model was estimated in a fashion that would maximize conceptual clarity (as opposed to statistical rigor, as in the second model) and enable estimation of outcome models in which the dependent variable could be treated as a count variable (e.g., Poisson regression). In this model, items were recoded into binary indicators of absence or presence of the aggressive or victimization behavior, and factor loadings were constrained to be equivalent for all items. This model was analogous to a simple sum score created by adding all binary items for a dimension.

Using these binary and polytomous (i.e., composed of more than two categories) items, confirmatory factor models were used to obtain continuous factor scores, which served as the continuous underlying measure of the dimension being explored (e.g., prosocial behavior or violence). For the primary victimization and violence outcomes, only the binary and polytomous models were compared. The secondary and intermediate outcomes included several dimensions that could feasibly use continuous items (those that used Likert-type responses), and so CFA models with continuous, binary, and polytomous items were compared. Table C-5 shows the model results for student violence and victimization measures. Three fit indices, comparative fix index, root mean square error of approximation, and weighted root mean square residual were used to evaluate overall model fit of each model type, but comparison of fit across model types is equivocal, given the different structure of items used in each model and the added constraints on the factor loadings in the binary model. As shown in table C-5, the correspondence in the underlying factor scores for each scoring method was very high, with correlations ranging from 0.963 to 0.995. In the interest of conceptual clarity, it was decided that the violence and victimization items could be used as sum scores derived from dichotomized items.

For the secondary outcomes, all measures were found to be best represented by models using continuous items, indicating that composite measures could be formed as simple means. The only exception to this was the "negative reactions to anger" dimension, which was best represented using dichotomized items. This composite was formed as a sum score, just like the violence and victimization measures. Table C-6 shows the model fit and factor score correlations for the secondary outcomes. Note that the models for "negative reactions to anger" were saturated, and so the fit is perfect for models 1 and 3. Model 2 for this outcome variable has an added degree of freedom due to the factor loading constraints and so is not saturated.

Table C-5.   Correspondence of measurement models for primary outcomes

	Correlation of CFA factor scores	Polytomous CFA			Binary CFA		
		CFI	RMSEA	WRMR	CFI	RMSEA	WRMR
Violence							
General	0.989	0.944	0.063	3.020	0.970	0.053	4.016
Weapon	0.963	1.000	0.000	0.005	0.995	0.043	1.255
Non-weapon	0.969	0.963	0.057	2.416	0.973	0.055	3.830
Victimization							
General	0.967	0.931	0.078	3.724	0.958	0.064	4.844
Social	0.986	1.000	0.050	1.283	0.977	0.097	4.439
Physical	0.995	0.971	0.061	2.407	0.962	0.067	3.907

NOTE: CFA = confirmatory factor analysis. CFI = comparative fix index. RMSEA = root mean square error of approximation. WRMR = weighted root mean square residual.
SOURCE: Student survey, fall 2006 (baseline).

Table C-6. **Correspondence of measurement models for secondary and intermediate outcomes**

Outcome	Correlation of models 1 and 2	Correlation of models 1 and 3	Correlation of models 2 and 3	Model 1: Polytomous CFA			Model 2: Binary CFA			Model 3: Continuous CFA		
				CFI	RMSEA	WRMR	CFI	RMSEA	WRMR	CFI	RMSEA	WRMR
Prosocial behaviors												
Extended	0.664	0.992	0.688	0.977	0.062	2.288	0.980	0.029	2.071	0.982	0.052	0.019
Received	0.757	0.995	0.766	0.979	0.048	2.008	0.988	0.024	1.853	0.981	0.043	0.019
Student self-reported coping strategies												
Positive	0.879	0.990	0.911	0.916	0.101	4.017	0.909	0.083	5.876	0.928	0.082	0.039
Negative	0.798	0.986	0.845	1.000	0.000	0.003	0.998	0.024	1.128	1.000	0.000	0.000
Attitudes toward violence	0.874	0.988	0.815	0.981	0.071	2.261	0.990	0.051	2.213	0.977	0.059	0.022
Perceived behavioral expectations	0.790	0.989	0.833	0.587	0.095	4.633	0.878	0.055	3.428	0.845	0.080	0.046

NOTE: CFA = confirmatory factor analysis. CFI = comparative fix index. RMSEA = root mean square error of approximation. WRMR = weighted root mean square residual.
SOURCE: Student survey, fall 2006 (baseline).

Appendix D:
Structured Protocols for Collecting Implementation Data

Project staff used structured protocols to conduct observations and interviews and to record other program implementation data. In this appendix, we highlight the topics and questions covered in each protocol, describe the procedures for coding the data, and describe the methods used to maximize consistency across data collectors.

D.1 Staff Interviews

Structured interviews were conducted during annual site visits with: (1) violence prevention coordinators, (2) violence prevention teachers, and (3) school management team members.

D.1.1 Violence Prevention Coordinator Interview Protocol

The Violence Prevention Coordinator Interview provided information on existing violence prevention strategies in both intervention and comparison schools. The protocol covered the following topics: specific approaches, strategies, or programs used to prevent and reduce violence; role of law enforcement and task force teams; staff professional development for violence prevention; future plans for staff training, new programs or strategies, or changes in policy; and perceptions of the Responding in Peaceful and Positive Ways (RiPP) and Best Behavior implementation (intervention schools only).

Questions about the following specific strategies and programs provided data for this report:

- environmental approaches, such as security cameras, metal detectors, area monitoring, random locker searches, visitor sign-in, and restrictions (e.g., use of IDs, dress code, clear book bags);

- any policy-based or whole-school strategies;

- names of classroom-based curricula and other programs implemented outside the classroom (e.g., conflict resolution, peer mediation, student court); and

- other approaches or strategies (e.g., interventions for high-risk students, classroom management programs, coordinated crisis intervention plans).

D.1.2 Violence Prevention Staff Interview Protocol

The Violence Prevention Staff Interview provided information on implementation of RiPP and was conducted with RiPP teachers in intervention schools. The protocol covered the following topics: background and responsibilities of the respondent; RiPP implementation experiences; fidelity and adaptation; training and technical assistance received; and awareness of and involvement with Best Behavior.

The following questions provided data for this report:

- How often do you teach the RiPP materials (number of sessions per week)? How long does each session generally last (minutes per session)? Is this amount of time usually enough to cover all of the lesson or not enough?

- What are the major challenges that you face in teaching the RiPP material and working with the students?

- Are there specific activities or topics that students have trouble with?

- Please tell me about any challenges you encountered in using the following approaches or techniques and how you handled those challenges:

  - making RiPP real—tying it to students' daily life;

  - having students role play;

  - having students work in small groups;

  - emphasizing the importance of "self-talk"; and

  - using Review to begin sessions and Closure to end them.

- How often did you leave out any of the approaches or techniques? Which ones? What were some of the reasons for not including these elements?

- Did you change or adapt the content of any of the RiPP lessons you have taught so far? Did you leave anything out or add anything? If so, why? What? How?

- Did you change or adapt how you taught the lessons? If so, why? What? How?

- Are you aware of the school's involvement with implementing the Best Behavior program?

- How have you been made aware of activities for this program (e.g., school-wide communications, in-service training, banners)?

- Have you been directly involved with any of the Best Behavior activities? How?

- Have you received any in-service training for activities related to Best Behavior, such as classroom management, school rules, or reward system?

- To what extent do the RiPP and Best Behavior programs work together?

D.1.3 School Management Team Interview Protocol

Interviews were conducted with members of the school management team to obtain information about implementation of Best Behavior. Topics addressed during the interview were as follows: background and role of respondent; Best Behavior implementation experiences; fidelity and adaptation of Best Behavior; training and technical assistance received; and familiarity with RiPP.

The following specific questions provided data for this report:

- How long have you been involved in implementing the Best Behavior program in your school?

- How involved have you been with the school management team? Have you attended most meetings?

- How comfortable are you with the Best Behavior approach?

- What are the major challenges in implementing the Best Behavior program?

- How familiar are you with the RiPP program being taught in the classroom?

- If you are familiar with them, to what extent do the Best Behavior and the RiPP programs work together?

D.1.4 Procedures for Collecting and Coding Interview Data

Interview data were gathered by RTI staff with experience conducting interviews with school staff. A 1-day training was held at RTI for these staff during which the structured protocols were discussed in detail and staff role-played portions of the interview. During the interview, the researcher took notes and also recorded the session. The recordings facilitated the production of verbatim transcriptions of each interview following the site visit.

To analyze the interview data, the transcribed interviews were imported into the qualitative software program ATLAS.ti 6.0 (Muhr and Friese 2009). A coding scheme was developed for each question, based initially on a review of 10 interviews of each type and subsequently updated and revised based on a review of all interviews. For example, the question, "What are the major challenges in implementing the Best Behavior program?" included such codes as the following: 1 = time constraints; 2 = difficulty using program; 3 = funding; 4 = lack of leadership; 5 = low teacher buy-in; 6 = desirable rewards. Two experienced analysts coded the interviews (one analyst coded one protocol, and the other analyst coded the other two protocols) and then discussed each other's coding schemes to ensure consistency with their approach. Once coding was completed, the data were analyzed and summarized for each of the three interviews.

D.2   Classroom Observations

RTI staff used a standard protocol to record observations of RiPP sessions in intervention schools. In preparation for the classroom observations, staff were instructed to review the curriculum lesson plan for the particular lesson to be observed and become familiar with the activities, teaching techniques, and sequencing of lesson elements. Observers were also instructed to have a copy of the lesson plan available for reference during the observation and to observe an entire lesson or class.

Topics covered in the protocol were the following: session (grade, date, start/end time, lesson name/number); general observations (any special circumstances during the session, interruptions); whether the RiPP poster was visible in the classroom; use of teaching techniques; fidelity to the lesson; student engagement; quality of the lesson delivery; and completeness of the lesson delivery.

Three specific measures provided data for this report:

- The fidelity indicator for "adherence to the prescribed teaching techniques" was measured by the extent to which the teacher followed the same teaching techniques outlined in the lesson plan (e.g., lecture, class discussion, small group, brainstorming, game, role play, worksheet). Observers listed the techniques outlined in the lesson plan and then recorded the techniques used by the teacher during the session. Adherence to

the prescribed teaching techniques was demonstrated if all techniques outlined in the lesson plan were used. Fidelity alignment for adherence to prescribed teaching methods was coded as follows: *well aligned* was indicated by all three classes observed in the intervention school displaying the techniques; *moderately aligned* was indicated by two of the three classes displaying the techniques; and *poorly aligned* was indicated by one of the three classes displaying the techniques.

- The fidelity indicator for "adherence to the scripted lesson plan" was measured by an item on the observation protocol that asked the observer to rate the extent to which the teacher followed the written instructions for activities. Observers followed along with the curriculum lesson plan and at the end of the session rated the adherence to the curriculum as follows: 1 = not at all; 2 = for 1 or 2 activities; 3 = for most but not all activities; 4 = always. For each intervention school, scores were averaged across the three classroom observations. Average scores greater than or equal to three were coded as well aligned; scores between two and three were coded as moderately aligned; and scores below two were coded as poorly aligned.

- The fidelity indicator for "student responsiveness" was measured by an item on the observation protocol that asked the observer to rate the extent to which students were engaged during the session. The observer rated the level of overall class engagement at the end of the session, using the following scale: 1 = not at all; 2 = a little; 3 = some; 4 = very much. The observer was instructed to base the rating on the level of student participation in the various activities, such as large-group discussions, role playing, and small group work, and the extent to which students stayed on task. For example, a rating of "1" was used when students did not ask questions, make comments, discuss, or stay on task to complete the group work or assignment. A score of "4" would indicate a high level of participation in discussions, group participants that contributed to the group activity, and tasks that were completed by the end of the period. For each intervention school, scores were averaged across the three classroom observations. Average scores greater than or equal to three were coded as well aligned; scores between two and three were coded as moderately aligned; and scores below two were coded as poorly aligned.

Inter-rater reliability could not be performed due to the fact that only one person observed each class and that each class could be on a different lesson. To increase the consistency in scoring among observers, we took the following steps:

- to become familiar with the RiPP curriculum, observers attended a 2-day RiPP training for teachers held by the RiPP developers during which lessons were discussed and concepts demonstrated;

- observers were trained to use the classroom observation protocol during a separate training held at RTI for staff who conducted the observations;

- each observer was provided a training guide with instructions and a copy of the curriculum;

- observers used a structured protocol with checkboxes or rating scales to be marked for each item so as to facilitate completion of the form; and

- upon returning from a site visit, during weekly project meetings, staff discussed any issues they encountered with the use of the protocols, which served to clarify items for all site visitors.

D.3 Liaison Implementation Reports

Site liaisons noted the schools' progress with RiPP and Best Behavior using a standard protocol which they discussed with the implementation task leaders at Tanglewood, Inc. (subcontractor that provided implementation oversight), during monthly phone calls. At the end of the school year, the same protocol was used to summarize the implementation progress for each intervention school for the year. Again, the site liaisons discussed their responses with Tanglewood researchers during individual calls.

The monthly and year-end implementation reports focused on the following topics: overall school status (staff turnover, school probation status), principal characteristics (change in principal, principal support and leadership); status of Best Behavior implementation; Best Behavior training; major accomplishments and challenges; plans and concerns for coming year; status of RiPP delivery; RiPP training; major accomplishments and major challenges for RiPP; and plans and concerns for next year.

The following questions from the protocol provided data for this report:

- Is this school on probation of any kind?

- Was the principal new this year? Was the principal the one recruited into the project?

- Was the principal supportive of Best Behavior and RiPP? Did the principal use leadership to promote the programs?

- Assessment of progress on Best Behavior (10 dimensions representing key program elements, rated on 4-point scale).

- What staff attended the Best Behavior workshop (name, position, date, RiPP teacher or not)?

- From your perspective as the school liaison, what were the major accomplishments (for Best Behavior) this year? What were the major challenges (with Best Behavior) the school faced this year?

- RiPP delivery status, by teacher (teacher name, subject area, attendance at training, number of classes responsible for teaching, lesson numbers completed to date, quality of delivery rating: 1 = poor to 5 = outstanding, rating comment).

- From your perspective as the school liaison, what were the major accomplishments (with RiPP) this year? What were the major challenges (with RiPP) the school faced this year?

Program implementation data from the site liaison reports were analyzed and used, as appropriate, to provide additional information to that collected by the RTI staff. Count data (e.g., number of schools on probation) provided summary information across schools. Other data, such as the ratings of Best Behavior progress, were used as fidelity measures. Responses to open-ended questions provided contextual information about the schools, for example, to describe the types of challenges faced by schools.

Appendix E:
Statistical Precision

Statistical power indicates the likelihood that an intervention effect will be judged to be statistically significant, given the assumptions of the specified model. In other words, power is the likelihood of observing a statistically significant difference, where such a difference exists. An underpowered study, then, is one in which the investigators risk failing to notice a significant intervention effect. Previous small-scale, less-rigorous evaluations of the Responding in Peaceful and Positive Ways (RiPP) intervention have reported significant effects for suspensions, violent behaviors, self-reported frequency of physical violence, drug use, peer provocation, nonphysical violence among boys, attitudes for violence in boys, delinquent behaviors, and victimization (Farrell, Meyer et al. 2003; Farrell, Valois et al. 2003; Farrell, Meyer, and White 2001; Farrell, Valois, and Meyer 2002). Significant effects from these studies have been in the range of 0.11 to 0.45 (Cohen's *d*). The current study is designed to detect an effect size of 0.20, given the effect sizes found in previous studies of RiPP and the fact that the inclusion of the Best Behavior program is anticipated to increase program effects above what would be found with a curriculum alone. In addition, the U.S. Department of Education commissioned a design paper (Bos, Weinstock, and Frankenberg 2004), and the evaluation team for the current study convened technical working groups of experts in the fields of school-based violence prevention, research design, and statistics. Through consultation with these sources, the researchers for the current study concluded that an effect size (ES) of 0.20 was reasonable, given the estimated monetary and opportunity costs of a school-based intervention combining curricular and whole-school approaches.

To determine the sample size required for an ES of 0.20, the study team assumed an intraclass correlation coefficient (ICC) of 0.045, based on a review of research measuring self-reported violent behavior in students (Janega et al. 2004). The type 1 error rate was set at 0.05, and a two-tailed significance test was employed. The type 2 error rate was set at 0.20 (yielding 80 percent statistical power). Also, a modest reduction in variation for the inclusion of covariates and matching of schools within districts was assumed; matching was based on the percentage of students receiving free and reduced-price lunches. Given these assumptions, randomly assigning 40 schools from within matched pairs to either intervention or control conditions (i.e., 20 schools in each condition) and surveying 243 students at each school was assumed to provide sufficient power for testing the null hypothesis regarding treatment effectiveness on schools (table E-1).[24]

These estimates are based on the assumptions described above as implemented through the mixed-effects model, as is appropriate for an analysis that includes random effects at the individual and school levels. The realized level of statistical power for tests of program effects may vary where the above assumptions are not met. For example, if the observed ICC is higher than the value used here, power will be reduced. Similarly, higher than anticipated levels of attrition could adversely affect power. On the other hand, statistical power could be improved even further due to the inclusion of covariates in the models or if the reduction in variation associated with the matching factor exceeds minimum expectations.

---

[24] Estimations of statistical precision are based on assumptions appropriate for a normally distributed variable and are considered approximate for nonnormally distributed variables.

Table E-1.   Eighty percent statistical power as a function of minimal detectable difference (Number of students per school = 243)

Effect size estimate	Number of schools required for each study condition
0.24	13
0.22	16
0.20	20
0.18	24
0.16	30

NOTE: Calculations assume α = 0.05 and a two-tailed test. Calculations are based on intraclass correlation coefficient (ICC) = 0.045.
SOURCE: RTI calculations.

Table E-2 provides a power table for the nested cohort analysis that was used to assess the subsample of high-risk youth.[25] Using assumptions noted in section 2.6 regarding attrition rates across the 3 years of the program, the study team anticipated that approximately 36 students from each school will provide data for this analysis. Although fewer students provide data for the cohort analysis than the cross-sectional analysis, this potential limitation is offset by the anticipated gain in precision associated with taking replicate measures on individuals. In this case, a sample size of 40 schools should allow us to detect effects of 0.21 or larger. This ES is slightly larger than the 0.20 ES for the primary outcome but is not unreasonable, given the potential for the greater impact that a violence prevention program may have on high-risk youth.

Table E-2.   Eighty percent statistical power as a function of minimal detectable difference (Number of high-risk students per school = 36)

Effect size estimate	Number of schools required for each study condition
0.22	15
0.21	17
0.21	20
0.19	21
0.17	26
0.15	33

NOTE: Calculations assume α = 0.05 and a two-tailed test. Calculations are based on intraclass correlation coefficient (ICC) = 0.019.
SOURCE: RTI calculations.

In table E-2, statistical power could be improved even further by including covariates (such as baseline data) in the models. In summary, tables E-1 and E-2 demonstrate that a total of 40 schools are required to detect the target minimal detectable effects.

---

[25] Estimations of statistical precision are based on assumptions appropriate for a normally distributed variable and are considered approximate for nonnormally distributed variables.

Appendix F:
Sample Characteristics and Baseline Measures for Year Two Participants

In this appendix, we present the demographic characteristics of the school and student samples and the baseline outcome measures for the 38 schools that participated in the first 2 years of the study. These inform the impacts findings for year two that are presented in appendix H.

F.1 Sample Characteristics for Participants in Year Two

Table F-1. Baseline demographic characteristics in schools with 2 years of participation

Characteristic	All schools	Intervention	Control	Difference	*p*-value
Enrollment (38 schools)					
Mean	852	822	881	−59.00	0.37
Range	462–1,404	462–1,404	634–1,209		
Race/ethnicity minority (%)					
Mean percentage	65.94	68.47	63.42	5.04	0.62
Range	10.27–100.00	10.27–100.00	14.79–100.00		
Students eligible for free or reduced-price lunches (%)					
Mean percentage	54.52	54.73	54.30	0.42	0.95
Range	16.44–96.88	27.48–83.55	16.44–96.88		

NOTE: A two-tailed *t*-test was used to test differences between intervention schools and control schools. Statistical significance is indicated by * if the *p*-value is less than or equal to .05.
SOURCE: *Search for Public School Districts: School Year 2005–06*, Common Core of Data (CCD), U.S. Department of Education, Institute of Education Sciences, National Center for Educational Statistics. Retrieved April 29, 2009, from http://nces.ed.gov/ccd/districtsearch/index.asp.

Table F-2. **Baseline demographic characteristics of the student sample in schools with 2 years of participation**

Characteristic	Intervention group	Control group	Difference	t-test[1]	p-value
Sample size, grade 6 (N = 38 schools)	3,390	3,639			
Race/ethnicity (%)					
Hispanic	39.93	31.89	8.04	0.28	0.78
Black, non-Hispanic	23.31	21.26	2.05	0.26	0.80
White, non-Hispanic	27.30	37.48	−10.18	−0.67	0.51
Other or mixed[2]	9.46	9.36	0.10	0.34	0.74
Gender (%)					
Male	49.14	48.99	0.15	0.07	0.95
Single-adult household (%)	60.12	59.66	0.46	0.06	0.96

[1] Adjusted for intraclass correlations.

[2] This category includes American Indian or Alaskan Native, Asian, Native Hawaiian/other Pacific Islander, and multiracial.

NOTE: Statistical significance is indicated by * if the p-value is less than or equal to .05.

SOURCE: Student survey, fall 2006 (baseline).

Table F-3. Difference in baseline demographic characteristics between the student sample and population in schools with 2 years of participation

Characteristic	6th-grade population	6th-grade study participants	Difference	*t*-test	*p*-value
Total sample (N = 38 schools)					
Race/ethnicity (Mean %)					
Hispanic	32.07	32.24	−0.17	−0.41	0.69
Black, non-Hispanic	27.61	24.36	3.25	8.50	0.00*
White, non-Hispanic	34.91	33.94	0.98	2.16	0.04*
Gender (Mean %)					
Male	50.56	48.90	1.66	2.78	0.01*
Intervention group (*N* = 18 schools)					
Race/ethnicity (Mean %)					
Hispanic	33.58	33.45	0.12	0.20	0.84
Black, non-Hispanic	28.97	25.72	3.26	6.16	0.00*
White, non-Hispanic	32.08	31.04	1.04	1.65	0.12
Gender (Mean %)					
Male	51.24	48.91	2.34	2.73	0.01*
Control group (*N* = 18 schools)					
Race/ethnicity (Mean %)					
Hispanic	30.57	31.01	−0.44	−0.76	0.46
Black, non-Hispanic	26.25	23.01	3.24	5.90	0.00*
White, non-Hispanic	37.74	36.82	0.92	1.41	0.18
Gender (Mean %)					
Male	49.90	48.88	1.02	1.24	0.23

NOTE: A two-tailed test adjusted for intraclass correlations was used to test differences between population estimates and study sample estimates. Statistical significance is indicated by * if the *p*-value is less than or equal to .05. Comparisons for race/ethnicity may not be valid because of differences in the way the data were collected.
SOURCE: Calculations for the "6th-grade study participants" are based on a student survey administered at baseline in fall 2006. Calculations for the "6th-grade population" were based on *Search for Public School Districts: School Year 2005–06*, Common Core of Data (CCD), U.S. Department of Education, Institute of Education Sciences, National Center for Educational Statistics. Retrieved April 29, 2009, from http://nces.ed.gov/ccd/districtsearch/index.asp.

Table F-4.   Student response rates, total sample—Year two

	Total 7th-grade enrollment (2007–08)	Percent consented of total 7th-grade enrollment	Percent surveyed of consented	Percent surveyed of total 7th-grade enrollment
Second follow-up (N = 38 schools)				
Total sample	10,257	64	96	62
School median	254	65	99	62
School range	151–416	35–88	84–100	34–88

SOURCE: Data collection records maintained by the research team.

Table F-5.   Student response rates, high-risk subsample—Year two

	Number tracked	Percent surveyed	Percent not surveyed—absence, refusal	Percent not surveyed—withdrawal, transfer, not located
Second follow-up (N = 38 schools)				
High-risk subsample	2,033	78	2	20

SOURCE: Data collection records maintained by the research team.

Table F-6.   Student response and mobility rates, by group—Year two

	Total	Intervention	Control	Difference	t-test	p-value
(Mean rate)						
Full sample response rate at second follow-up	62.44	62.57	62.31	0.26	0.06	0.95
Perpetrator, high-risk subsample response rate at second follow-up	78.25	78.12	78.36	−0.24	−0.12	0.90
Nonperpetrator, high-risk subsample response rate at second follow-up	78.34	78.85	77.88	0.97	0.17	0.86
Full sample rate of exiting students[1]	11.26	11.12	11.40	−0.28	−0.25	0.81
Full sample rate of entering students[2]	21.73	21.55	21.90	−0.35	−0.06	0.95

[1] Exiting students are defined as those who consented at prior time points but left the school by the second follow-up.
[2] Entering students are defined as those who were identified as new to the school at the second follow-up and for whom consent was attempted.
NOTE: A two-tailed test was used to test differences between intervention schools and control schools. Statistical significance is indicated by * if the *p*-value is less than or equal to .05.
SOURCE: Data collection records maintained by the research team.

F.2 Baseline Measures in Schools With 2 Years of Study Participation

In this section, we present the results of analyses conducted to assess whether or not there were any differences between intervention and control groups on the student measures at baseline, for the total sample and for high-risk and gender subgroups, in schools with 2 years of participation.

Table F-7. Baseline measures for the student sample in schools with 2 years of participation: Full sample

| Measure | Baseline event rates or scale means (SE) | | Difference | *t*-statistic | *p*-value[1] |
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	1.94 (0.13)	1.86 (0.13)	0.08	0.44	0.66
Violence: Weapons-related	0.08 (0.01)	0.08 (0.01)	0.00	0.17	0.87
Violence: Not weapons-related	1.86 (0.13)	1.78 (0.12)	0.08	0.47	0.64
Victimization (All items)	3.85 (0.14)	4.07 (0.15)	−0.22	−1.09	0.28
Victimization: Overt	2.17 (0.09)	2.30 (0.09)	−0.13	−1.02	0.31
Victimization: Relational	1.68 (0.06)	1.78 (0.06)	−0.10	−1.13	0.27
Secondary outcomes[3]					
Self-reported coping strategies: Positive	2.88 (0.02)	2.91 (0.02)	−0.03	−1.50	0.14
Self-reported coping strategies: Negative[2]	0.58 (0.03)	0.58 (0.03)	0.00	−0.02	0.98
Attitudes toward violence	3.00 (0.03)	3.02 (0.03)	−0.02	−0.50	0.62
Prosocial behaviors: Extended to others	2.93 (0.03)	2.96 (0.03)	−0.03	−0.68	0.50
Prosocial behaviors: Received from others	2.77 (0.02)	2.78 (0.02)	−0.01	−0.23	0.82
Behavioral expectations	3.16 (0.02)	3.16 (0.02)	0.00	0.10	0.92
Safety concerns	1.87 (0.03)	1.95 (0.03)	−0.08	−2.16	0.04*
Sample size (Schools)	19	19			
Sample size (Students)[4]	3,390	3,639			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data unless otherwise indicated. Generalized linear mixed model used to estimate group-specific baseline event rates and standard errors (SEs); *t*-statistic adjusted for clustering of students within schools used to test the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.2 percent to 3.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table F-8. Baseline measures for the student sample in schools with 2 years of participation: High-risk, perpetrator subgroup

Measure	Baseline event rates or scale means (SE)		Difference	t-statistic	p-value[1]
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	5.46 (0.12)	4.97 (0.11)	0.49	3.01	0.00*
Violence: Weapons-related	0.30 (0.03)	0.28 (0.03)	0.02	0.45	0.65
Violence: Not weapons-related	5.16 (0.10)	4.69 (0.09)	0.47	3.42	0.00*
Victimization (All items)	6.10 (0.13)	6.26 (0.12)	−0.16	−0.92	0.37
Victimization: Overt	3.72 (0.08)	3.81 (0.07)	−0.09	−0.87	0.39
Victimization: Relational	2.38 (0.06)	2.44 (0.05)	−0.06	−0.74	0.46
Secondary outcomes[3]					
Self-reported coping strategies: Positive	2.60 (0.03)	2.69 (0.02)	−0.08	−2.38	0.02*
Self-reported coping strategies: Negative[2]	1.11 (0.04)	1.06 (0.03)	0.05	1.00	0.33
Attitudes toward violence	2.59 (0.03)	2.64 (0.02)	−0.05	−1.49	0.15
Prosocial behaviors: Extended to others	2.72 (0.03)	2.77 (0.03)	−0.05	−1.06	0.30
Prosocial behaviors: Received from others	2.70 (0.03)	2.69 (0.02)	0.01	0.35	0.73
Behavioral expectations	3.03 (0.02)	3.02 (0.02)	0.01	0.35	0.73
Safety concerns	1.96 (0.04)	2.11 (0.04)	−0.15	−2.71	0.01*
Sample size (Schools)	19	19			
Sample size (Students)[4]	841	975			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and standard errors (SEs), difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table F-9. **Baseline measures for the student sample in schools with 2 years of participation: High-risk, nonperpetrator subgroup**

Measure	Baseline event rates or scale means (SE)		Difference	*t*-statistic	*p*-value[1]
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	1.22 (0.16)	1.22 (0.16)	−0.01	−0.03	0.98
Violence: Weapons-related	0.00 (0.00)	0.00 (0.00)	—	—	—
Violence: Not weapons-related	1.22 (0.16)	1.22 (0.16)	−0.01	−0.03	0.98
Victimization (All items)	3.45 (0.28)	3.77 (0.30)	−0.33	−0.80	0.43
Victimization: Overt	1.99 (0.18)	2.14 (0.19)	−0.15	−0.57	0.57
Victimization: Relational	1.50 (0.14)	1.66 (0.14)	−0.16	−0.82	0.42
Secondary outcomes[3]					
Self-reported coping strategies: Positive	2.04 (0.07)	2.11 (0.07)	−0.08	−0.78	0.44
Self-reported coping strategies: Negative[2]	2.29 (0.15)	2.35 (0.15)	−0.06	−0.28	0.78
Attitudes toward violence	2.19 (0.07)	2.25 (0.07)	−0.06	−0.59	0.56
Prosocial behaviors: Extended to others	2.62 (0.11)	2.61 (0.11)	0.00	0.03	0.98
Prosocial behaviors: Received from others	2.63 (0.09)	2.50 (0.09)	0.13	1.02	0.32
Behavioral expectations	2.84 (0.05)	3.03 (0.05)	−0.19	−2.87	0.01*
Safety concerns	1.71 (0.10)	1.71 (0.10)	0.01	0.04	0.97
Sample size (Schools)	19	19			
Sample size (Students)[4]	104	113			

— Not available. Nonperpetrator subgroup defined as those youth who reported no violent behaviors at baseline.

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and standard errors (SEs), difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table F-10. Baseline measures for the student sample in schools with 2 years of participation: Boys

Measure	Baseline event rates or scale means (SE)		Difference	*t*-statistic	*p*-value[1]
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	2.07 (0.14)	2.09 (0.14)	−0.02	−0.11	0.92
Violence: Weapons-related	0.10 (0.01)	0.11 (0.01)	−0.01	−0.49	0.62
Violence: Not weapons-related	1.97 (0.13)	1.98 (0.13)	−0.01	−0.04	0.97
Victimization (All items)	4.09 (0.15)	4.48 (0.16)	−0.40	−1.80	0.08
Victimization: Overt	2.45 (0.10)	2.68 (0.11)	−0.24	−1.63	0.11
Victimization: Relational	1.64 (0.06)	1.80 (0.07)	−0.16	−1.84	0.07
Secondary outcomes[3]					
Self-reported coping strategies: Positive	2.77 (0.02)	2.82 (0.02)	−0.05	−1.45	0.16
Self-reported coping strategies: Negative[2]	0.63 (0.03)	0.69 (0.04)	−0.06	−1.21	0.23
Attitudes toward violence	2.90 (0.03)	2.90 (0.03)	0.00	0.04	0.97
Prosocial behaviors: Extended to others	2.84 (0.03)	2.84 (0.03)	0.00	−0.01	1.00
Prosocial behaviors: Received from others	2.69 (0.03)	2.68 (0.03)	0.01	0.24	0.81
Behavioral expectations	3.13 (0.02)	3.13 (0.02)	0.00	−0.11	0.91
Safety concerns	1.80 (0.03)	1.87 (0.03)	−0.07	−2.02	0.05*
Sample size (Schools)	19	19			
Sample size (Students)[4]	1,659	1,778			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and standard errors (SEs), difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.2 percent to 4.6 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table F-11. Baseline measures for the student sample in schools with 2 years of participation: Girls

Measure	Baseline event rates or scale means (SE)		Difference	*t*-statistic	*p*-value[1]
	Intervention group	Control group			
Main outcomes[2]					
Violence (All items)	1.81 (0.15)	1.61 (0.14)	0.20	0.98	0.33
Violence: Weapons-related	0.06 (0.01)	0.04 (0.01)	0.01	0.83	0.41
Violence: Not weapons-related	1.75 (0.14)	1.546 (0.13)	0.19	0.99	0.33
Victimization (All items)	3.60 (0.17)	3.65 (0.18)	−0.05	−0.19	0.85
Victimization: Overt	1.89 (0.11)	1.90 (0.11)	−0.01	−0.10	0.92
Victimization: Relational	1.71 (0.07)	1.75 (0.07)	−0.04	−0.41	0.69
Secondary outcomes[3]					
Self-reported coping strategies: Positive	2.98 (0.02)	3.00 (0.02)	−0.02	−0.65	0.52
Self-reported coping strategies: Negative[2]	0.52 (0.04)	0.46 (0.04)	0.06	1.01	0.32
Attitudes toward violence	3.09 (0.04)	3.14 (0.04)	−0.05	−0.80	0.43
Prosocial behaviors: Extended to others	3.02 (0.03)	3.07 (0.03)	−0.05	−1.20	0.24
Prosocial behaviors: Received from others	2.85 (0.03)	2.88 (0.03)	−0.03	−0.67	0.51
Behavioral expectations	3.20 (0.02)	3.19 (0.02)	0.01	0.17	0.87
Safety concerns	1.94 (0.04)	2.03 (0.04)	−0.10	−1.82	0.08
Sample size (Schools)	19	19			
Sample size (Students)[4]	1,717	1,851			

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to .05.

[2] Based on count data. Generalized linear mixed model used to estimate group-specific baseline event rates (ERs) and standard errors (SEs), difference in ERs, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[3] Based on continuous scale measures (unless otherwise indicated). Generalized linear mixed model used to estimate group-specific baseline scale means and SEs, difference in means, *t*-statistic for testing the null hypothesis of no difference, and significance level.

[4] Missing data ranged from 0.1 percent to 2.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Appendix G:
Unadjusted Means and Standard Deviations for Impact Variables—Years Two and Three

G.1 Unadjusted Means and Standard Deviations—Year Two

Table G-1. Unadjusted means and standard deviations for student outcomes in schools with 2 years of participation—Baseline

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	1.97 (2.82)	1.89 (2.68)
Violence: Weapons-related	0.08 (0.39)	0.08 (0.38)
Violence: Not weapons-related	1.89 (2.63)	1.81 (2.50)
Victimization (All items)	3.83 (3.39)	4.08 (3.41)
Victimization: Overt	2.17 (2.22)	2.30 (2.25)
Victimization: Relational	1.67 (1.47)	1.79 (1.49)
Secondary outcomes[2]		
Safety concerns[3]	1.88 (0.92)	1.96 (0.94)
Prosocial behaviors extended[4]	2.93 (0.68)	2.96 (0.68)
Prosocial behaviors received[4]	2.78 (0.66)	2.78 (0.65)
Perceived behavioral expectations[4]	3.16 (0.43)	3.16 (0.42)
Attitudes toward violence[4]	3.01 (0.67)	3.03 (0.66)
Self-reported coping strategies (Positive)[3]	2.88 (0.64)	2.91 (0.63)
Self-reported coping strategies (Negative)[1,4]	0.58 (0.88)	0.58 (0.90)
Sample size (Schools)	19	19
Sample size (Students)	3,390	3,639

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0.2 percent to 3.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-2. **Unadjusted means and standard deviations for high-risk student (Perpetrator) outcomes in schools with 2 years of participation—Baseline**

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	5.47 (3.07)	4.99 (3.02)
Violence: Weapons-related	0.31 (0.70)	0.29 (0.69)
Violence: Not weapons-related	5.17 (2.77)	4.70 (2.72)
Victimization (All items)	6.10 (3.30)	6.25 (3.21)
Victimization: Overt	3.73 (2.27)	3.81 (2.26)
Victimization: Relational	2.38 (1.37)	2.44 (1.34)
Secondary outcomes[2]		
Safety concerns[3]	1.96 (0.92)	2.11 (0.96)
Prosocial behaviors extended[4]	2.72 (0.67)	2.77 (0.66)
Prosocial behaviors received[4]	2.70 (0.63)	2.69 (0.63)
Perceived behavioral expectations[4]	3.02 (0.45)	3.02 (0.44)
Attitudes toward violence[4]	2.59 (0.64)	2.64 (0.64)
Self-reported coping strategies (Positive)[3]	2.61 (0.67)	2.69 (0.65)
Self-reported coping strategies (Negative)[1,4]	1.11 (1.05)	1.06 (1.02)
Sample size: Students within 38 schools	841	975

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-3. Unadjusted means and standard deviations for high-risk (Nonperpetrator) outcomes in schools with 2 years of participation—Baseline

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	1.27 (1.39)	1.29 (1.51)
Violence: Weapons-related	0.00 (0.00)	0.00 (0.00)
Violence: Not weapons-related	1.27 (1.39)	1.29 (1.51)
Victimization (All items)	3.49 (3.09)	3.85 (3.18)
Victimization: Overt	2.01 (2.05)	2.19 (2.09)
Victimization: Relational	1.50 (1.37)	1.65 (1.46)
Secondary outcomes[2]		
Safety concerns[3]	1.71 (0.90)	1.70 (0.88)
Prosocial behaviors extended[4]	2.62 (0.87)	2.58 (0.81)
Prosocial behaviors received[4]	2.62 (0.77)	2.48 (0.78)
Perceived behavioral expectations[4]	2.84 (0.56)	3.03 (0.41)
Attitudes toward violence[4]	2.19 (0.59)	2.25 (0.60)
Self-reported coping strategies (Positive)[3]	2.03 (0.60)	2.12 (0.63)
Self-reported coping strategies (Negative)[1,4]	2.29 (0.57)	2.35 (0.64)
Sample size: Students within 38 schools	104	113

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

**Table G-4.   Unadjusted means and standard deviations for student outcomes, by gender
subgroup, in schools with 2 years of participation—Baseline**

Measure	Event rates or scale means (standard deviations)			
	Boys		Girls	
	Intervention group	Control group	Intervention group	Control group
Main outcomes[1]				
Violence (All items)	2.09 (2.88)	2.11 (2.84)	1.85 (2.76)	1.67 (2.50)
Violence: Weapons-related	0.10 (0.44)	0.11 (0.45)	0.06 (0.32)	0.05 (0.30)
Violence: Not weapons-related	1.98 (2.65)	2.00 (2.61)	1.79 (2.61)	1.62 (2.37)
Victimization (All items)	4.07 (3.53)	4.47 (3.57)	3.60 (3.22)	3.70 (3.19)
Victimization: Overt	2.44 (2.35)	2.67 (2.38)	1.90 (2.06)	1.94 (2.05)
Victimization: Relational	1.63 (1.47)	1.80 (1.50)	1.70 (1.48)	1.77 (1.48)
Secondary outcomes[2]				
Safety concerns[3]	1.80 (0.89)	1.87 (0.92)	1.95 (0.93)	2.04 (0.95)
Prosocial behaviors extended[4]	2.83 (0.70)	2.84 (0.70)	3.02 (0.66)	3.07 (0.63)
Prosocial behaviors received[4]	2.69 (0.66)	2.68 (0.65)	2.86 (0.64)	2.88 (0.62)
Perceived behavioral expectations[4]	3.12 (0.44)	3.13 (0.43)	3.20 (0.42)	3.19 (0.41)
Attitudes toward violence[4]	2.91 (0.66)	2.91 (0.68)	3.10 (0.66)	3.14 (0.62)
Self-reported coping strategies (Positive)[3]	2.77 (0.66)	2.81 (0.67)	2.98 (0.60)	3.00 (0.58)
Self-reported coping strategies (Negative)[1,4]	0.63 (0.92)	0.69 (0.98)	0.53 (0.84)	0.47 (0.80)
Sample size (Schools)	19	19	19	19
Sample size (Students nested within schools)	1,659	1,778	1,717	1,851

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from
0.2 percent to 3.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-5.   Unadjusted means and standard deviations for student outcomes—Year two

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	3.05 (3.36)	2.93 (3.28)
Violence: Weapons-related	0.13 (0.47)	0.11 (0.45)
Violence: Not weapons-related	2.93 (3.15)	2.82 (3.07)
Victimization (All items)	4.69 (3.64)	4.89 (3.63)
Victimization: Overt	2.77 (2.48)	2.87 (2.47)
Victimization: Relational	1.92 (1.49)	2.03 (1.49)
Secondary outcomes[2]		
Safety concerns[3]	1.76 (0.89)	1.82 (0.92)
Prosocial behaviors extended[4]	2.89 (0.71)	2.94 (0.72)
Prosocial behaviors received[4]	2.85 (0.69)	2.85 (0.71)
Perceived behavioral expectations[4]	2.91 (0.53)	2.86 (0.56)
Attitudes toward violence[4]	2.73 (0.71)	2.79 (0.71)
Self-reported coping strategies (Positive)[3]	2.71 (0.68)	2.72 (0.65)
Self-reported coping strategies (Negative)[1,4]	1.02 (1.07)	0.97 (1.04)
Sample size (Schools)	19	19
Sample size (Students)	3,171	3,329

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 1 percent to 5 percent.

SOURCE: Student survey, spring 2008 (second follow-up).

Table G-6. Unadjusted means and standard deviations for high-risk student (Perpetrator)—Year two

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	4.77 (3.70)	4.47 (3.66)
Violence: Weapons-related	0.23 (0.64)	0.22 (0.61)
Violence: Not weapons-related	4.53 (3.42)	4.24 (3.37)
Victimization (All items)	5.70 (3.68)	6.04 (3.61)
Victimization: Overt	3.54 (2.54)	3.71 (2.51)
Victimization: Relational	2.16 (1.45)	2.34 (1.43)
Secondary outcomes[2]		
Safety concerns[3]	1.77 (0.89)	1.89 (0.93)
Prosocial behaviors extended[4]	2.76 (0.73)	2.79 (0.73)
Prosocial behaviors received[4]	2.85 (0.69)	2.77 (0.72)
Perceived behavioral expectations[4]	2.80 (0.55)	2.74 (0.56)
Attitudes toward violence[4]	2.47 (0.65)	2.55 (0.68)
Self-reported coping strategies (Positive)[3]	2.50 (0.68)	2.55 (0.66)
Self-reported coping strategies (Negative)[1,4]	1.40 (1.11)	1.28 (1.10)
Sample size: Students within 38 schools	657	764

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 4 percent to 5 percent.

SOURCE: Student survey, spring 2008 (second follow-up).

Table G-7. **Unadjusted means and standard deviations for high-risk student (Nonperpetrator) outcomes—Year two**

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	4.29 (3.79)	3.46 (3.43)
Violence: Weapons-related	0.30 (0.79)	0.19 (0.57)
Violence: Not weapons-related	3.99 (3.34)	3.27 (3.16)
Victimization (All items)	4.73 (3.35)	4.55 (4.02)
Victimization: Overt	2.80 (2.47)	2.79 (2.73)
Victimization: Relational	1.95 (1.30)	1.76 (1.57)
Secondary outcomes[2]		
Safety concerns[3]	1.59 (0.81)	1.78 (0.94)
Prosocial behaviors extended[4]	2.69 (0.76)	2.63 (0.82)
Prosocial behaviors received[4]	2.76 (0.80)	2.74 (0.82)
Perceived behavioral expectations[4]	2.67 (0.51)	2.68 (0.59)
Attitudes toward violence[4]	2.22 (0.62)	2.31 (0.74)
Self-reported coping strategies (Positive)[3]	2.26 (0.70)	2.21 (0.72)
Self-reported coping strategies (Negative)[1,4]	1.63 (1.12)	1.66 (1.18)
Sample size: Students within 38 schools	82	88

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 4 percent to 5 percent.

SOURCE: Student survey, spring 2008 (second follow-up).

Table G-8.   Unadjusted means and standard deviations for student outcomes, by gender subgroup—Year two

| Measure | Event rates or scale means (standard deviations) | | | |
| | Boys | | Girls | |
	Intervention group	Control group	Intervention group	Control group
Main outcomes[1]				
Violence (All items)	3.12 (3.48)	2.99 (3.39)	2.98 (3.23)	2.87 (3.15)
Violence: Weapons-related	0.16 (0.54)	0.15 (0.52)	0.09 (0.39)	0.08 (0.36)
Violence: Not weapons-related	2.96 (3.21)	2.84 (3.13)	2.89 (3.07)	2.79 (3.01)
Victimization (All items)	4.89 (3.78)	5.02 (3.79)	4.49 (3.49)	4.75 (3.44)
Victimization: Overt	3.02 (2.57)	3.11 (2.57)	2.52 (2.35)	2.62 (2.34)
Victimization: Relational	1.88 (1.49)	1.92 (1.52)	1.97 (1.49)	2.13 (1.46)
Secondary outcomes[2]				
Safety concerns[3]	1.74 (0.89)	1.77 (0.93)	1.78 (0.89)	1.85 (0.90)
Prosocial behaviors extended[4]	2.77 (0.72)	2.82 (0.72)	3.01 (0.69)	3.06 (0.69)
Prosocial behaviors received[4]	2.73 (0.69)	2.70 (0.71)	2.97 (0.67)	3.00 (0.68)
Perceived behavioral expectations[4]	2.88 (0.55)	2.82 (0.56)	2.95 (0.51)	2.91 (0.55)
Attitudes toward violence[4]	2.61 (0.70)	2.68 (0.70)	2.84 (0.69)	2.90 (0.70)
Self-reported coping strategies (Positive)[3]	2.57 (0.69)	2.59 (0.67)	2.83 (0.64)	2.83 (0.61)
Self-reported coping strategies (Negative)[1,4]	1.05 (1.09)	1.04 (1.08)	0.99 (1.04)	0.89 (0.99)
Sample size (Schools)	19	19	19	19
Sample size (Students nested within schools)	1,557	1,634	1,602	1,683

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 3 percent to 4 percent.

SOURCE: Student survey, spring 2008 (second follow-up).

Table G-9. **Unadjusted means and standard deviations for impacts on teacher outcomes—Year two**

Measure[1]	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Teacher self-reported victimization[2,3]	0.64 (0.48)	0.60 (0.49)
Teacher safety concerns[3]	1.43 (0.64)	1.41 (0.59)
School consistency of enforcing behavioral rules[4]	2.89 (0.56)	2.90 (0.58)
Interactions with victims[4]	3.38 (0.53)	3.34 (0.56)
Interactions with aggressors[4]	3.51 (0.49)	3.52 (0.50)
Classroom management techniques[4]	3.68 (0.35)	3.65 (0.38)
Sample size (Schools)	19	19
Sample size (917 teachers nested within schools)[5]	435	443

[1] Reported as scale scores based on continuous measures of the identified construct, unless otherwise indicated.

[2] Teacher victimization is based on a dichotomous indicator. Results presented indicate the proportion of teachers who reported victimization by a student.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

[5] Teacher sample sizes vary due to item nonresponse. Missing data ranged from 0 percent to 1 percent.

SOURCE: Teacher survey, spring 2008.

G.2 Unadjusted Means and Standard Deviations—Year Three

**Table G-10. Unadjusted means and standard deviations for student outcomes in schools with
3 years of participation—Baseline**

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	1.98 (2.84)	1.88 (2.66)
Violence: Weapons-related	0.08 (0.39)	0.08 (0.37)
Violence: Not weapons-related	1.90 (2.64)	1.80 (2.48)
Victimization (All items)	3.86 (3.39)	4.10 (3.40)
Victimization: Overt	2.18 (2.22)	2.31 (2.25)
Victimization: Relational	1.68 (1.47)	1.80 (1.49)
Secondary outcomes[2]		
Safety concerns[3]	1.87 (0.91)	1.96 (0.94)
Prosocial behaviors extended[4]	2.93 (0.68)	2.97 (0.67)
Prosocial behaviors received[4]	2.78 (0.65)	2.79 (0.65)
Perceived behavioral expectations[4]	3.16 (0.43)	3.16 (0.42)
Attitudes toward violence[4]	3.01 (0.67)	3.02 (0.66)
Self-reported coping strategies (Positive)[3]	2.88 (0.64)	2.91 (0.63)
Self-reported coping strategies (Negative)[1,4]	0.58 (0.88)	0.58 (0.90)
Sample size (Schools)	18	18
Sample size (Students)	3,198	3,418

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from
0.2 percent to 3.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-11. Unadjusted means and standard deviations for high-risk student (Perpetrator) outcomes in schools with 3 years of participation—Baseline

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	5.50 (3.09)	4.95 (2.99)
Violence: Weapons-related	0.32 (0.71)	0.28 (0.68)
Violence: Not weapons-related	5.19 (2.77)	4.67 (2.70)
Victimization (All items)	6.18 (3.28)	6.21 (3.21)
Victimization: Overt	3.78 (2.26)	3.79 (2.25)
Victimization: Relational	2.40 (1.36)	2.42 (1.35)
Secondary outcomes[2]		
Safety concerns[3]	1.97 (0.92)	2.11 (0.96)
Prosocial behaviors extended[4]	2.72 (0.66)	2.77 (0.66)
Prosocial behaviors received[4]	2.70 (0.63)	2.69 (0.63)
Perceived behavioral expectations[4]	3.02 (0.45)	3.01 (0.44)
Attitudes toward violence[4]	2.59 (0.64)	2.64 (0.65)
Self-reported coping strategies (Positive)[3]	2.61 (0.67)	2.68 (0.65)
Self-reported coping strategies (Negative)[1,4]	1.13 (1.04)	1.07 (1.01)
Sample size: Students within 36 schools	790	922

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-12. Unadjusted means and standard deviations for high-risk student (Nonperpetrator) outcomes in schools with 3 years of participation—Baseline

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	1.29 (1.36)	1.29 (1.49)
Violence: Weapons-related	0.00 (0.00)	0.00 (0.00)
Violence: Not weapons-related	1.29 (1.36)	1.29 (1.49)
Victimization (All items)	3.60 (3.14)	3.94 (3.18)
Victimization: Overt	2.08 (2.09)	2.25 (2.11)
Victimization: Relational	1.53 (1.38)	1.69 (1.46)
Secondary outcomes[2]		
Safety concerns[3]	1.67 (0.88)	1.73 (0.89)
Prosocial behaviors extended[4]	2.64 (0.87)	2.60 (0.80)
Prosocial behaviors received[4]	2.65 (0.75)	2.48 (0.78)
Perceived behavioral expectations[4]	2.83 (0.56)	3.03 (0.40)
Attitudes toward violence[4]	2.18 (0.60)	2.21 (0.57)
Self-reported coping strategies (Positive)[3]	2.04 (0.61)	2.11 (0.64)
Self-reported coping strategies (Negative)[1,4]	2.27 (0.55)	2.32 (0.65)
Sample size: Students within 36 schools	97	106

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0.1 percent to 3.1 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-13. Unadjusted means and standard deviations for student outcomes, by gender subgroup, in schools with 3 years of participation—Baseline

| | Event rates or scale means (standard deviations) | | | |
| | Boys | | Girls | |
Measure	Intervention group	Control group	Intervention group	Control group
Main outcomes[1]				
Violence (All items)	2.09 (2.89)	2.10 (2.83)	1.87 (2.78)	1.66 (2.47)
Violence: Weapons-related	0.11 (0.45)	0.11 (0.44)	0.06 (0.32)	0.05 (0.30)
Violence: Not weapons-related	1.99 (2.66)	2.00 (2.60)	1.81 (2.62)	1.61 (2.35)
Victimization (All items)	4.10 (3.53)	4.49 (3.57)	3.62 (3.22)	3.72 (3.19)
Victimization: Overt	2.46 (2.35)	2.68 (2.38)	1.92 (2.06)	1.95 (2.04)
Victimization: Relational	1.64 (1.47)	1.82 (1.50)	1.71 (1.47)	1.77 (1.48)
Secondary outcomes[2]				
Safety concerns[3]	1.79 (0.89)	1.88 (0.92)	1.94 (0.93)	2.05 (0.95)
Prosocial behaviors extended[4]	2.83 (0.69)	2.85 (0.70)	3.02 (0.66)	3.08 (0.63)
Prosocial behaviors received[4]	2.69 (0.66)	2.69 (0.66)	2.86 (0.64)	2.88 (0.62)
Perceived behavioral expectations[4]	3.12 (0.44)	3.12 (0.43)	3.20 (0.42)	3.19 (0.41)
Attitudes toward violence[4]	2.91 (0.67)	2.91 (0.68)	3.10 (0.66)	3.14 (0.63)
Self-reported coping strategies (Positive)[3]	2.77 (0.66)	2.82 (0.67)	2.98 (0.60)	3.00 (0.58)
Self-reported coping strategies (Negative)[1,4]	0.63 (0.92)	0.70 (0.98)	0.53 (0.84)	0.47 (0.80)
Sample size (Schools)	18	18	18	18
Sample size (Students nested within schools)	1,560	1,670	1,625	1,739

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0.2 percent to 3.5 percent.

SOURCE: Student survey, fall 2006 (baseline).

Table G-14. Unadjusted means and standard deviations for student outcomes—Year three

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes1		
Violence (All items)	2.88 (3.39)	2.69 (3.14)
Violence: Weapons-related	0.16 (0.56)	0.11 (0.46)
Violence: Not weapons-related	2.72 (3.09)	2.58 (2.92)
Victimization (All items)	4.09 (3.59)	4.27 (3.61)
Victimization: Overt	2.39 (2.42)	2.47 (2.44)
Victimization: Relational	1.71 (1.47)	1.80 (1.49)
Secondary outcomes[2]		
Safety concerns[3]	1.61 (0.84)	1.65 (0.84)
Prosocial behaviors extended[4]	2.94 (0.72)	2.94 (0.73)
Prosocial behaviors received[4]	2.92 (0.70)	2.91 (0.71)
Perceived behavioral expectations[4]	2.85 (0.53)	2.81 (0.56)
Attitudes toward violence[4]	2.72 (0.68)	2.76 (0.69)
Self-reported coping strategies (Positive)[3]	2.69 (0.64)	2.71 (0.62)
Self-reported coping strategies (Negative)[1,4]	1.09 (1.06)	1.04 (1.06)
Sample size (Schools)	18	18
Sample size (Students)	2,842	3,086

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 1 percent to 5 percent.

SOURCE: Student survey, spring 2009 (third follow-up).

Table G-15. Unadjusted means and standard deviations for high-risk student (Perpetrator) outcomes—Year three

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	3.97 (3.67)	3.90 (3.56)
Violence: Weapons-related	0.23 (0.65)	0.19 (0.57)
Violence: Not weapons-related	3.74 (3.36)	3.72 (3.29)
Victimization (All items)	4.83 (3.67)	5.16 (3.78)
Victimization: Overt	2.96 (2.54)	3.09 (2.58)
Victimization: Relational	1.87 (1.43)	2.07 (1.52)
Secondary outcomes[2]		
Safety concerns[3]	1.65 (0.88)	1.75 (0.91)
Prosocial behaviors extended[4]	2.85 (0.73)	2.78 (0.73)
Prosocial behaviors received[4]	2.88 (0.72)	2.81 (0.71)
Perceived behavioral expectations[4]	2.73 (0.57)	2.70 (0.57)
Attitudes toward violence[4]	2.50 (0.66)	2.56 (0.66)
Self-reported coping strategies (Positive)[3]	2.54 (0.63)	2.60 (0.61)
Self-reported coping strategies (Negative)[1,4]	1.42 (1.10)	1.35 (1.12)
Sample size: Students within 36 schools	552	661

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 4 percent to 5 percent.

SOURCE: Student survey, spring 2009 (third follow-up).

Table G-16. Unadjusted means and standard deviations for high-risk student (Nonperpetrator) outcomes—Year three

Measure	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Main outcomes[1]		
Violence (All items)	3.51 (3.94)	3.10 (3.80)
Violence: Weapons-related	0.22 (0.68)	0.26 (0.73)
Violence: Not weapons-related	3.29 (3.53)	2.84 (3.25)
Victimization (All items)	4.01 (3.82)	4.14 (4.08)
Victimization: Overt	2.47 (2.64)	2.47 (2.78)
Victimization: Relational	1.54 (1.40)	1.67 (1.60)
Secondary outcomes[2]		
Safety concerns[3]	1.48 (0.87)	1.68 (0.95)
Prosocial behaviors extended[4]	2.80 (0.81)	2.71 (0.78)
Prosocial behaviors received[4]	2.82 (0.66)	2.76 (0.84)
Perceived behavioral expectations[4]	2.58 (0.52)	2.57 (0.65)
Attitudes toward violence[4]	2.30 (0.66)	2.42 (0.66)
Self-reported coping strategies (Positive)[3]	2.35 (0.70)	2.29 (0.65)
Self-reported coping strategies (Negative)[1,4]	1.73 (1.02)	1.50 (1.10)
Sample size: Students within 36 schools	70	74

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 4 percent to 5 percent.

SOURCE: Student survey, spring 2009 (third follow-up).

Table G-17. Unadjusted means and standard deviations for student outcomes, by gender subgroup—Year three

| | Event rates or scale means (standard deviations) | | | |
| | Boys | | Girls | |
Measure	Intervention group	Control group	Intervention group	Control group
Main outcomes[1]				
Violence (All items)	2.98 (3.60)	2.75 (3.32)	2.77 (3.17)	2.64 (2.978)
Violence: Weapons-related	0.20 (0.65)	0.15 (0.54)	0.11 (0.45)	0.08 (0.35)
Violence: Not weapons-related	2.78 (3.22)	2.61 (3.03)	2.66 (2.96)	2.56 (2.83)
Victimization (All items)	4.31 (3.80)	4.32 (3.78)	3.87 (3.35)	4.21 (3.45)
Victimization: Overt	2.67 (2.57)	2.60 (2.56)	2.10 (2.21)	2.34 (2.32)
Victimization: Relational	1.65 (1.48)	1.72 (1.52)	1.77 (1.46)	1.87 (1.47)
Secondary outcomes[2]				
Safety concerns[3]	1.60 (0.85)	1.62 (0.84)	1.63 (0.83)	1.68 (0.84)
Prosocial behaviors extended[4]	2.83 (0.73)	2.79 (0.76)	3.05 (0.69)	3.08 (0.67)
Prosocial behaviors received[4]	2.80 (0.71)	2.77 (0.73)	3.04 (0.67)	3.04 (0.67)
Perceived behavioral expectations[4]	2.81 (0.54)	2.78 (0.58)	2.89 (0.52)	2.85 (0.54)
Attitudes toward violence[4]	2.59 (0.68)	2.63 (0.69)	2.86 (0.66)	2.88 (0.66)
Self-reported coping strategies (Positive)[3]	2.54 (0.64)	2.56 (0.63)	2.83 (0.60)	2.85 (0.57)
Self-reported coping strategies (Negative)[1,4]	1.10 (1.09)	1.11 (1.09)	1.08 (1.04)	0.98 (1.02)
Sample size (Schools)	18	18	18	18
Sample size (Students nested within schools)	1,402	1,511	1,425	1,552

[1] Based on count data.

[2] Scales based on continuous measures of the identified construct, unless otherwise indicated.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

NOTE: Student sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 3 percent to 4 percent.

SOURCE: Student survey, spring 2009 (third follow-up).

Table G-18. Unadjusted means and standard deviations for program impacts on teacher outcomes—Year three

Measure[1]	Event rates or scale means (standard deviations)	
	Intervention group	Control group
Teacher self-reported victimization[2,3]	0.55 (0.50)	0.57 (0.49)
Teacher safety concerns[3]	1.35 (0.54)	1.36 (0.58)
School consistency of enforcing behavioral rules[4]	3.00 (0.57)	2.94 (0.58)
Interactions with victims[4]	3.41 (0.50)	3.33 (0.56)
Interactions with aggressors[4]	3.53 (0.51)	3.54 (0.47)
Classroom management techniques[4]	3.65 (0.42)	3.62 (0.39)
Sample size (Schools)	18	18
Sample size (917 teachers nested within schools)[5]	429	428

[1] Reported as scale scores based on continuous measures of the identified construct, unless otherwise indicated.

[2] Teacher victimization is based on a dichotomous indicator. Results presented indicate the proportion of teachers who reported victimization by a student.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

[5] Teacher sample sizes vary due to item nonresponse. Missing data ranged from 0 percent to 1 percent.

SOURCE: Teacher survey, spring 2009.

Appendix H:
Impact Findings From Year Two

This appendix presents the impact findings from the interim data collection that occurred following the second year of a 3-year program to reduce student violence and victimization in middle schools. Results obtained after the second year of the study were similar to those presented in the main report after 3 years. Specifically, there were no statistically significant differences between intervention and control schools on self-reported violence or victimization or on secondary or intermediate outcomes such as student safety concerns, teacher victimization, and student attitudes toward violence. Further, no significant impacts on violence or victimization were found among the high-risk student subgroup.

H.1   Impact Findings After 2 Years of Participation

In the following tables, we present the results of analyses conducted to assess the impact of the intervention. We first examine the main outcomes that the intervention sought to influence: student violence and student victimization (table H-1). In addition, we examine the impacts on several other sets of outcomes. Table H-2 reports secondary outcomes on students. These outcomes include constructs that are hypothesized to be more immediately affected by the intervention, compared with the main outcomes of violence and victimization. These include students' attitudes toward violence, students' coping strategies for dealing with violence, students' clarity of understanding of the school rules, teacher expectations for student behavior, and school staff's response to student violence. Table H-2 also includes outcomes that, while not directly targeted by the program, might be expected to change as a result of the program (i.e., spillover effects), such as student safety concerns and student prosocial behaviors. Table H-3 then presents finding from the second teacher survey. These outcomes include teacher victimizations (by students), teacher safety concerns, rule enforcement and use of classroom management techniques, and interactions with violent or victimized students.

Tables H-4 and H-5 report the results of analyses conducted to assess the capacity of the Responding in Peaceful and Positive Ways (RiPP) and Best Behavior programs to promote individual change in a group of students identified as being at higher risk for perpetration of violent acts. Data are presented separately for two groups of high-risk youth. Table H-4 contains results of analyses based on a repeated measures model of students identified at baseline as high-risk, based on their responses to a number of attitude-and-belief questions regarding the appropriateness of using violence while not reporting any violent behaviors in the past 30 days (e.g., nonperpetrator group). Table H-5 contains results of analyses based on a repeated measures model of students identified at baseline as high-risk, based on their acknowledgment of engaging in violent behaviors in the past 30 days (e.g., perpetrator group).

Table H-6 presents results of gender subgroup analyses based on the interim data collection. The youth aggression literature suggests that girls differ from boys in the types of aggression displayed, with relational aggression being more prevalent among girls than among boys (Orpinas and Horne 2006; Crick and Grotpeter 1995). Furthermore, research indicates that prevention programs may impact girls and boys differently (Farrell and Meyer 1997; Simon et al. 2002). These analyses were undertaken to explore whether or not the effects of the RiPP and Best Behavior programs affect boys and girls differently.

Table H-1.   Main program impacts on self-reported violence and victimization after 2 years of program delivery

Self-reported student outcome	Model-adjusted follow-up event rates (SE)		Estimated impact (95% CI)[1]	Wald *p*-value[2]	BH adjusted *p*-value[3]
	Intervention group	Control group			
Violence (All items)[4]	2.99 (0.10)	2.88 (0.10)	1.04 (0.96, 1.13)	0.3239	†
Without a weapon	2.87 (0.10)	2.77 (0.09)	1.04 (0.96, 1.12)	0.3367	0.013
With a weapon	0.11 (0.01)	0.10 (0.01)	1.10 (0.86, 1.40)	0.4154	0.025
Victimization (All items)[4]	4.72 (0.13)	4.73 (0.13)	1.00 (0.94, 1.06)	0.9247	†
Overt	2.77 (0.09)	2.76 (0.09)	1.00 (0.93, 1.09)	0.9051	0.025
Relational	1.94 (0.05)	1.98 (0.05)	0.98 (0.91, 1.05)	0.5606	0.013
Sample size (schools)	19	19			
Sample size (6,354 students clustered within schools)[5]	3,131	3,223			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald *p*-value (unadjusted) ≤.05, two-tailed test.

‡ Statistically significant at Wald *p*-value ≤ B-H critical *p*-value, two-tailed test.

[1] Program impact estimated as a model-adjusted event rate ratio (ERR) for intervention versus controls at follow-up, with 95 percent confidence limits. Impact estimates of 1.00 indicate no difference between intervention and control groups.

[2] The Wald *p*-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical *p*-value.

[3] Critical *p*-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional *p*-value when determining statistical significance. Where the Wald *p*-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[4] Based on count data.

[5] Student sample sizes used in the analysis vary due to item nonresponse at follow-up and/or covariate nonresponse. Missing data ranged from 2 percent to 4 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, treatment condition (intervention vs. control), race/ethnicity, sex, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student survey, fall 2006 (baseline) and spring 2008 (second follow-up).

Table H-2. **Secondary program impacts on student-level outcomes after 2 years of program delivery**

Self-reported student outcome	Model-adjusted follow-up event rates or scale means (SE)		Estimated impact (95% CI)	Wald p-value
	Intervention group	Control group		
Safety concerns[1,4]	1.92 (0.06)	1.91 (0.06)	0.01 (−0.07, 0.09)	0.7576
Prosocial behaviors extended[1,2]	2.91 (0.05)	2.93 (0.05)	−0.03 (−0.08, 0.02)	0.2738
Prosocial behaviors received[1,2]	2.85 (0.04)	2.85 (0.04)	−0.01 (−0.06, 0.05)	0.8170
Perceived behavioral expectations[1,2]	2.91 (0.03)	2.87 (0.03)	0.04 (−0.03, 0.12)	0.2601
Attitudes toward violence[1,2]	2.75 (0.02)	2.79 (0.02)	−0.04 (−0.10, 0.02)	0.1796
Coping strategies (Negative)[3,4]	0.99 (0.03)	0.95 (0.03)	1.05 (0.97, 1.12)	0.1958
Coping strategies (Positive)[1,2]	2.72 (0.02)	2.71 (0.02)	0.01 (−0.04, 0.07)	0.6230
Sample size (schools)	19	19		
Sample size (6,354 students clustered within schools)[5]	3,131	3,223		

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

[1] Scales based on continuous measures of the identified construct, unless otherwise indicated. Results presented include the estimated group-specific scale means and standard errors (SEs) at follow-up, estimated program impact for intervention versus control (difference in scale means, with 95 percent confidence limits), and the Wald p-value indicating statistical significance of the program impact. Impact estimates of 0.00 indicate no difference between intervention and control conditions.

[2] Higher scores indicate better outcomes.

[3] Based on count data. Results include the estimated group-specific event rates and SEs at follow-up, estimated program impact for intervention versus control (event rate ratio [ERR], with 95 percent confidence limits), and Wald p-value indicating statistical significance of the program impact. Impact estimates of 1.00 indicate no difference between intervention and control conditions.

[4] Lower scores indicate better outcomes.

[5] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 2 percent to 6 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function for count data) and linear mixed models (PROC MIXED, for continuous data) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, race/ethnicity, gender, number of parents in household, and school size. CI = confidence interval.

SOURCE: Student survey, fall 2006 (baseline) and spring 2008 (second follow-up).

Table H-3. Secondary program impacts on teacher outcomes after 2 years of program delivery

Teacher self-reported outcome[1]	Model-adjusted odds or scale means (SE)		Estimated impact (95% CI)	Wald Chi-square p-value
	Intervention group	Control group		
Teacher self-reported victimization[2,3]	0.65 (0.04)	0.61 (0.04)	0.54 (0.45, 0.63)	0.3658
Teacher safety concerns[3]	1.44 (0.04)	1.41 (0.04)	0.3 (−0.07, 0.13)	0.5564
School consistency of enforcing behavioral rules[4]	2.89 (0.06)	2.89 (0.06)	−0.01 (−0.14, 0.12)	0.9189
Interactions with victims[4]	3.39 (0.03)	3.34 (0.03)	0.05 (−0.03, 0.13)	0.2225
Interactions with aggressors[4]	3.52 (0.03)	3.52 (0.03)	−0.01 (−0.08, 0.06)	0.8565
Classroom management techniques[4]	3.68 (0.02)	3.65 (0.02)	0.03 (−0.03, 0.09)	0.2652
Sample size (Schools)	19	19		
Sample size (879 teachers nested within schools)[5]	435	444		

* Statistically significant at Wald *p*-value (unadjusted) ≤.05, two-tailed test.

[1] Reported as scale scores, unless otherwise indicated, based on continuous measures of the identified construct. Results presented include the estimated group-specific scale means and standard errors (SEs), estimated program impact (difference in scale means for intervention vs. control, with 95 percent confidence limits), and the Wald *p*-value indicating statistical significance of the program impact. Impact estimates of 0.00 indicate no difference between intervention and control condition.

[2] Teacher victimization is based on a dichotomous indicator. Results presented include the estimated group-specific odds and SEs, estimated program impact for intervention versus control (odds ratio, with 95 percent confidence limits), and the Wald *p*-value indicating statistical significance of the program impact. Impact estimates of 1.00 indicate no difference between intervention and control conditions.

[3] Lower scores indicate better outcomes.

[4] Higher scores indicate better outcomes.

[5] Teacher sample sizes used in the analysis vary due to item nonresponse. Missing data ranged from 0 percent to 1 percent.

NOTE: Program impacts are estimated using restricted maximum likelihood (MIXED procedure) or pseudo-likelihood (GLIMMIX procedure), controlling for the random assignment of schools to program condition from pairs matched within district on the level of free or reduced-price lunches received by students. CI = confidence interval.

SOURCE: Teacher survey, spring 2008.

Table H-4. Main program impacts on self-reported violence and victimization, after 2 years of program delivery: High-risk, nonperpetrator subgroup (Via repeated measures)

Student self-reported outcome[1]	Model-adjusted baseline event rates (SE)[2]		Model-adjusted follow-up event rates (SE)[2]		Estimated impact (95% CI)[3]	Wald p-value[4]	BH adjusted p-value[5]
	Intervention group[2]	Control group[2]	Intervention group[2]	Control group[2]			
Violence (All items)	1.21 (0.21)	1.31 (0.21)	4.27 (0.47)	3.46 (0.40)	0.75 (0.44, 1.30)	0.3008	†
Not weapons-related	1.21 (0.20)	1.30 (0.20)	3.99 (0.45)	3.26 (0.38)	0.76 (0.45, 1.30)	0.3146	0.0125
Weapons-related[6]	0.00 (0.00)	0.00 (0.00)	0.12 (0.05)	0.09 (0.05)	1.28 (0.32, 5.01)	0.7208	0.0250
Victimization (All items)	3.32 (0.36)	3.66 (0.37)	4.88 (0.45)	4.29 (0.40)	0.80 (0.56, 1.13)	0.1896	†
Overt	1.95 (0.25)	2.08 (0.25)	2.91 (0.32)	2.62 (0.29)	0.84 (0.56, 1.28)	0.4083	0.0250
Relational	1.36 (0.14)	1.55 (0.15)	1.97 (0.17)	1.67 (0.15)	0.74 (0.53, 1.04)	0.0801	0.0125
Sample size (Schools)	19	19	19	19			
Sample size (Students within schools)[7]	82	88	82	88			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald p-value ≤ B-H critical p-value, two-tailed test.

[1] Based on count data.

[2] Group by time-specific event rates.

[3] Program impact (with 95 percent confidence limits) estimated via difference-in-difference models comparing change across time in the intervention versus control group. Ratios of impact estimates of 1.00 indicate no interaction between time and program group (i.e., no program impact).

[4] The Wald p-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical p-value.

[5] Critical p-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional p-value when determining statistical significance. Where the Wald p-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[6] Parameter estimates cannot be obtained due to convergence issues. This may be due to lack of variability of the response variable and/or model complexity.

[7] Student sample sizes used in the analysis vary due to item nonresponse at baseline, follow-up, and/or covariate nonresponse. Missing data ranged from 3 percent to 5 percent, with 240 missing at follow-up.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included gender, race/ethnicity, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student surveys limited to a high risk-subgroup of 6th-graders surveyed in fall 2006 and followed up in spring 2008.

Table H-5. **Main program impacts on self-reported violence and victimization, after 2 years of program delivery: High-risk, perpetrator subgroup (Via repeated measures)**

Student self-reported outcome[1]	Model-adjusted baseline event rates (SE)[2]		Model-adjusted follow-up event rates (SE)[2]		Estimated impact (95% CI)[3]	Wald p-value[4]	BH adjusted p-value[5]
	Intervention group[2]	Control group[2]	Intervention group[2]	Control group[2]			
Violence (All items)	5.37 (0.18)	4.91 (0.16)	4.73 (0.16)	4.42 (0.15)	1.02 (0.90, 1.16)	0.7282	†
Not weapons-related	5.13 (0.17)	4.64 (0.15)	4.50 (0.15)	4.20 (0.14)	1.03 (0.91, 1.17)	0.6109	0.0250
Weapons-related	0.24 (0.03)	0.27 (0.03)	0.23 (0.03)	0.22 (0.02)	0.85 (0.55, 1.31)	0.4532	0.0125
Victimization (All items)	6.01 (0.17)	6.16 (0.16)	5.77 (0.16)	5.97 (0.15)	1.01 (0.92, 1.11)	0.8493	†
Overt	3.64 (0.12)	3.76 (0.11)	3.59 (0.11)	3.67 (0.11)	0.99 (0.89, 1.10)	0.8369	0.0250
Relational	2.37 (0.06)	2.39 (0.05)	2.18 (0.06)	2.30 (0.05)	1.05 (0.97, 1.14)	0.2376	0.0125
Sample Size (Schools)	19	19	19	19			
Sample Size (Students within schools)[6]	657	764	657	764			

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald p-value ≤ B-H critical p-value, two-tailed test.

[1] Based on count data.

[2] Group by time-specific event rates.

[3] Program impact (with 95 percent confidence limits) estimated via difference-in-difference models comparing change across time in the intervention versus control group. Ratios of impact estimates of 1.00 indicate no interaction between time and program group (i.e., no program impact).

[4] The Wald p-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical p-value.

[5] Critical p-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional p-value when determining statistical significance. Where the Wald p-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant.

[6] Student sample sizes used in the analysis vary due to item nonresponse at baseline, follow-up, and/or covariate nonresponse. Missing data ranged from 4 percent to 5 percent, with 240 missing at follow-up.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included gender, race/ethnicity, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student surveys limited to a high-risk subgroup of 6th-graders surveyed in fall 2006 and followed up in spring 2008.

Table H-6. Main program impacts on self-reported violence and victimization, after 2 years of program delivery: Gender subgroups

Student self-reported outcome[1]	Model-adjusted follow-up event rates (SE)				Impact on boys (95% CI)[2]	Impact on girls (95% CI)[2]	Ratio of impacts (95% CI)[3]	Wald Chi-Square p-value[4]	BH adjusted p-value[5]
	Boys		Girls						
	Intervention group	Control group	Intervention group	Control group					
Violence (All items)	3.00 (0.13)	2.75 (0.11)	2.73 (0.12)	2.65 (0.11)	1.09 (0.97, 1.22)	1.03 (0.92, 1.15)	1.06 (0.93, 1.21)	0.3831	†
Not weapons-related	2.81 (0.12)	2.60 (0.11)	2.63 (0.11)	2.57 (0.10)	1.08 (0.96, 1.21)	1.02 (0.92, 1.15)	1.05 (0.93, 1.20)	0.3988	0.0250
Weapons-related	0.18 (0.02)	0.14 (0.02)	0.10 (0.01)	0.08 (0.01)	1.27 (0.95, 1.68)	1.28 (0.89, 1.84)	0.99 (0.64, 1.54)	0.9623	0.0125
Victimization (All items)	4.40 (0.14)	4.23 (0.13)	3.90 (0.12)	4.14 (0.12)	1.04 (0.96, 1.13)	0.94 (0.87, 1.02)	1.10 (1.00, 1.22)	0.0469	†
Overt	2.71 (0.10)	2.55 (0.09)	2.10 (0.08)	2.30 (0.08)	1.06 (0.96, 1.17)	0.91 (0.82, 1.01)	1.16 (1.04, 1.30)	0.0115*‡	0.0250
Relational	1.71 (0.05)	1.68 (0.04)	1.81 (0.05)	1.83 (0.05)	1.02 (0.93, 1.10)	0.99 (0.91, 1.07)	1.03 (0.94, 1.13)	0.5336	0.0125
Sample size (Schools)	19	19	19	19					
Sample size (Students nested within schools)[6]	1,530	1,593	1,589	1,618					

† Not applicable. The Benjamini-Hochberg (B-H) correction is applied to a family of independent tests. Accordingly, it would be inappropriate to include the main impact indices in the family of adjusted tests.

* Statistically significant at Wald p-value (unadjusted) ≤ .05, two-tailed test.

‡ Statistically significant at Wald p-value ≤ B-H critical p-value, two-tailed test.

[1] Based on count data.

[2] Model-adjusted event rate ratios (ERRs) for intervention versus controls at follow-up, separately for boys versus girls. Impact estimates of 1.00 indicate no difference between intervention and control conditions.

[3] ERRs of program impacts for boys versus girls, with 95 percent confidence limits. Ratios of impact estimates of 1.00 indicate no interaction between gender and program group.

[4] The Wald p-value is unadjusted for multiple comparisons and should be interpreted in conjunction with the B-H critical p-value.

[5] Critical p-values in this column account for multiple comparisons based on the B-H method for controlling false discovery and should be used in place of the traditional p-value when determining statistical significance. Where the Wald p-value is equal to or less than the B-H critical value, the test of the program impact can be interpreted as statistically significant. [6] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 2 percent to 4 percent.

NOTE: Generalized linear mixed models (SAS PROC GLIMMIX, Poisson distribution with log link function) were used to evaluate the program impact while accounting for the clustering of students within schools. Covariates in the model included the baseline school mean of the response variable, gender, race/ethnicity, number of parents in household, and school size. CI = confidence interval. SE = standard error.

SOURCE: Student survey, fall 2006 (baseline) and spring 2008 (second follow-up).

H.2 Observed Precision—Year Two

Model-based estimates of minimum detectable effects (MDEs) are presented in table H-7. MDEs indicate the smallest differences that would allow us to reject the null hypothesis with confidence (i.e., 80 percent statistical power). MDEs are presented as event rate ratios (ERRs) that compare the mean number of events in intervention schools to the mean number of events in control schools. When the two event rates (ERs) are equal, ERR is 1.00. As the ER among intervention schools moves away from the ER in the control schools, ERR moves away from 1.00. Accordingly, the two ERRs listed for each outcome in table H-7 provide an interval space around the null value of 1.00 and identify the minimum program impacts that would be viewed statistically significant. For the current study, ERRs below 1.00 indicate positive program effects, while those above 1.00 indicate negative program effects.

The MDE for violent behavior (all items) in table H-7, for example, indicates that to reject the null hypothesis and view the ER among intervention schools as different from the ER among control schools would require a ratio of at least 0.89 to 1.00 (positive program effect) or 1.12 to 1.00 (negative program effect). To further illustrate, consider the case of an intervention effect where the mean ER among the students in the control schools is 2.89 events in the past 30 days. Here, the mean ER among students in the intervention schools would have to be 2.44 events in the past 30 days *or less* to achieve statistically significant positive program effects.

Table H-7. Observed intraclass correlation coefficients (ICCs) and minimum detectable effects (MDEs) reported as event rate ratios (ERRs) for self-reported violence and victimization—Year two

	ICCs		ERRs as MDE[1]	
Self-reported student outcome	Unadjusted	Adjusted	Positive program effect	Negative program effect
Violence (All items)	0.00527	0.004041	0.89	1.12
Not weapons-related	0.00571	0.004300	0.89	1.12
Weapons-related	0.00442	0.000362	0.71	1.40
Victimization (All items)	0.00797	0.005704	0.92	1.09
Overt	0.00835	0.009076	0.89	1.12
Relational	0.01642	0.011009	0.90	1.11

[1] Values below 1.00 indicate positive program effects; that is, event rates (ERs) in intervention schools are below those in control schools. Values greater than 1.00 indicate negative program effects; that is, ERs in control schools are below those in intervention schools.
SOURCE: Student survey, spring 2008.

Estimations used to derive MDEs presented in table H-7 involved a number of assumptions having to do with values derived from the empirical model. These assumptions include the following:

- The scale parameter observed in data reflects the true population parameter.

- The model-based variance components are true for population.

- The parameter estimates are asymptotically normal.

These assumptions underscore one of the main differences between linear and generalized linear modeling. In the former, covariance parameter values are assumed constant and independent of location; this assumption is untenable for the latter. It is important also to bear in mind that ERRs are calculated based on coefficients that have been estimated in the natural logarithmic (ln) scale where $\ln(1.00) = 0.00$. This fact, a statistical necessity, means that direct translation of standardized MDEs to ERR-based MDEs is not straightforward.