NCEE 2010-4007 U.S. DEPARTMENT OF EDUCATION

# Impacts of a Violence Prevention Program for Middle Schools

## Findings From the First Year of Implementation

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

# Impacts of a Violence Prevention Program for Middle Schools

## Findings From the First Year of Implementation

## April 2010

**Authors**
Suyapa Silvia
Jonathan Blitstein
Jason Williams
RTI International

Chris Ringwalt
Pacific Institute for Research and Evaluation (PIRE)

Linda Dusenbury
William Hansen
Tanglewood Research, Inc.

**John Rice**
Project Officer
Institute of Education Sciences

NCEE 2010-4007
U.S. DEPARTMENT OF EDUCATION

ies 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**
John Q. Easton
*Acting Commissioner*

**April 2010**

This report was prepared for the National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, under contract no. ED01CO0052/0015 with RTI International, the Pacific Institute for Research and Evaluation (PIRE), and Tanglewood Research, Inc.



**To order copies of this report,**

- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 22207, Alexandria, VA 22304.
- 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 703-605-6794 or order online at www.edpubs.gov.

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.

# Contents

# List of Tables

# List of Figures

# Acknowledgments

This study represents a collaborative effort among participating schools; staff from RTI International (RTI); our colleagues at Pacific Institute for Research and Evaluation (PIRE); researchers at Tanglewood Research, Inc.; and the program developers for Responding in Peaceful and Positive Ways (RiPP), Wendy Bauers Northup and Aleta Meyer, and the program developer for Best Behavior at the University of Oregon, Jeff Sprague. We are especially grateful to the school principals, faculty, and students of the participating schools for their dedication to implementing the programs and for graciously participating in all data collection.

At Tanglewood Research, Bill Hansen and Linda Dusenbury and a group of dedicated site liaisons supported, coached, and monitored teachers and faculty during the implementation of the programs; they also coordinated training and technical assistance for large numbers of staff. The program developers were true partners in this effort and were extremely flexible and cooperative with the needs of the study as we worked to implement an intervention formed from the two programs.

At RTI, the evaluation was supported by a team of devoted school recruiters, site coordinators, and site visitors, including Becky Durocher, Linda Bailey-Stone, Betty Burton, Terri Dempsey, Connie Hobbs, Farley Bernholz, Kimrey Millar, Lisa McCaskill, Elizabeth Parish, and Linda Pucci. The team oversaw locally based field data collectors who worked tirelessly to coordinate evaluation activities and achieve high response rates for the student survey. Data collection logistics and training were ably supported by Margaret Searle and Lori Hill, under the leadership of Lisa McCaskill and Linda Bailey-Stone. Jon Blitstein led the analyses; Jason Williams and Gayle Bieler provided additional support. Early drafts of this report were reviewed by Gayle Bieler, Jeremy Bray, Gordon Brown, and Dennis Wallace.

The study's Technical Working Group provided helpful advice and guidance on the evaluation's design and analysis. Three panel members also provided expert advice on the selection of the programs and conducted extremely helpful reviews of the revised curriculum. We appreciate the contributions of Robert Boruch, Thomas Cook, David Cordray, Daniel Flannery, Dennis Gorman, Nancy Guerra, Jeffrey Smith, Pamela Orpinas, and William Shadish.

# Disclosure of Potential Conflicts of Interest

The research team for this evaluation consists of a prime contractor, RTI International (RTI), 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.

# Executive Summary

This is the first of two reports that summarize the findings from an impact evaluation of a violence prevention intervention for middle schools. This report discusses findings after 1 year of implementation. A forthcoming report will discuss the findings after 2 years and 3 years of implementation. In 2004, the U.S. Department of Education (ED) contracted with RTI International (RTI) 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 assesses the intervention's effects on student violence and victimization and whether these effects vary by levels of student risk. Tanglewood Research, which assisted in the process by which the programs under study were selected, 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.

## Study Background

The Office of Safe and Drug-Free Schools (OSDFS), through the Safe and Drug-Free Schools and Communities Act (SDFSCA, Title IV-A), supports both state and national programs intended to prevent and reduce the levels of drug use and violence in and around schools. State grants from SDFSCA provide funding to approximately 97 percent of all school districts for this purpose. 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). 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. Data from a recent National Center for Education Statistics report indicate that the rate of victimization for students aged 12 through 14 was 35 incidents per 1,000 students, which was higher than the rate for students aged 15 through 18, which was 23 incidents per 1,000 students (Dinkes, Kemp, and Baum 2009). Similarly, data from the School Survey on Crime and Safety (SSOCS) for the 2005–06 school year reveal that students were more likely to experience a violent event in middle schools (52 per 1,000) than in elementary (25 per 1,000) or secondary (26 per 1,000) schools (Nolle, Guerino, and Dinkes 2007). National Crime Victimization Survey (NCVS) data also indicate that bullying is a significant problem. In 2005, some 36 percent of students in grades 6 and 7, and 29 percent of students in grades 8 and 9, reported having been bullied in the previous 6 months (Dinkes, Cataldi, and Lin-Kelly 2007).

Among the violence prevention strategies for middle schools which have been evaluated through rigorous methods are at least eight curriculum-based programs, including RiPP (Meyer and Northup 2002a, 2002b, 2006), one of the two prevention programs selected for this study. Four evaluations of RiPP have been conducted by the program's developers in schools serving students in

grades 6 through 8. The evaluations found statistically significant decreases in students' approval of violent behavior, levels of peer provocation, physical aggression, victimization, and discipline violations, although results have not been consistently maintained at follow-up time points (Farrell, Valois, and Meyer 2002; Farrell, Meyer, and White 2001; Farrell, Meyer, et al. 2003; Farrell, Valois, et al. 2003). While evaluations of RiPP and similar curriculum programs have yielded significant results, their effect sizes tended to be 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).

A few other school-based programs have sought to prevent violence using whole-school strategies that seek to influence the school environment, for example, through clarifying rules for behavior, increasing supervision of the school grounds, or setting up positive behavior reward systems. Some of these programs have been evaluated using middle school–age students, but not with the same methodological rigor to which the curriculum-based programs have been subjected. Included among these is Positive Behavior Support (PBS [Sugai and Horner 1994; Sprague, Sugai, and Walker 1998]), whose components and strategies form the basis for Best Behavior (Sprague and Golly 2005), the second of the prevention programs selected for this evaluation. Two evaluations of PBS conducted with elementary and middle school students found reductions in students' aggressive social behavior and discipline violations and also found that students had increased knowledge of social skills (Metzler et al. 2001; Sprague et al. 2001). A few other programs combine both curriculum and whole-school approaches, but their evaluations, while yielding promising results, have not been very rigorous. More research is clearly indicated, particularly to determine the effects of programs that use a combination of classroom-based curricula and whole-school approaches to prevent school violence, which experts in the field of violence prevention suggest may help boost impacts. This study will yield information regarding the effects of a combined curriculum and whole-school intervention for middle school students.

The study's research questions for impacts and implementation in year one 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? |
| • With the goal of decreasing disruptive, aggressive, violent, and other delinquent behaviors, what other interventions or prevention programs do the treatment and control schools implement, other than the violence prevention program under study? |
| **Intervention impacts** |
| • Does the degree of violence differ 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 and aggression?[1] |

[1] The two programs that make up the intervention are universal programs and, as such, are not designed specifically for students who are already exhibiting serious violent behavior in school. However, the study addressed this research question because the impacts of universal programs on high-risk students are still of interest to practitioners.

## Implementation Findings

The key descriptive findings regarding the first year of implementation of the curriculum portion of the program include the following:

- **The first-year curriculum was delivered in its entirety to a majority of assigned classrooms.** Seventy percent of schools delivered all 16 lessons to all classrooms where the curriculum had been assigned, while another 15 percent of schools delivered all lessons in at least three-fourths of these classrooms.

- **Teachers in the majority of treatment schools followed the curriculum lesson scripts and adhered to the prescribed teaching strategies.** Teachers in 65 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. In addition, teachers in 55 percent of the intervention schools delivered lessons with few or no deviations from the prescribed teaching strategies (e.g., using role plays or small group discussions). Observations made by the liaisons were comparable for alignment with the teaching techniques (65 percent for liaisons, as compared with 55 percent for the evaluation team) but less so for deviations from the lesson plan (40 percent for liaisons, as compared with 65 percent for the evaluation team).

- **In a majority of schools, students were engaged with the curriculum.** Based on evaluation team observations, students were found to be engaged during the lesson activities, exercises, and discussions in 85 percent of the intervention schools. Observations made by the liaisons were comparable and indicated that students were engaged with the curriculum in 60 percent of the intervention schools.

The main implementation findings for the whole-school portion of the program (Best Behavior) include the following:

- **Half of the treatment schools had principals who were supportive of the whole-school portion of the program.** Principals at 50 percent of the treatment schools were rated as supportive on three out of four indicators, according to liaisons who helped implement the program.

- **The school management teams, charged with developing and disseminating the school rules and reward systems, met less frequently than stipulated by the program.** Although school management teams are to meet monthly during the school year, the average team met on five occasions during the first year of implementation.

- **By the end of the first year, the majority of treatment schools had instituted behavioral rules and rewards, which are key steps during the first year of Best Behavior.** By the end of the first year, 75 percent of treatment schools had school rules posted in the school, 75 percent had instituted a token reward system for adhering to school rules, and 50 percent had the school rules taught in classrooms.

- **The majority of teachers agreed that the rules were well defined and that they were clear with regard to the behaviors being targeted.** Eighty-four percent of teachers at intervention schools agreed or strongly agreed that school rules were clearly defined, 79 percent agreed or strongly agreed that rules emphasized rewarding desired behaviors, and

69 percent agreed or strongly agreed that rules emphasized consequences for undesired behaviors.

- **The primary treatment contrast between the treatment and control schools was the violence prevention program implemented by the study.** While no control school was implementing Best Behavior, an equal number of control and intervention schools used a discipline tracking system, one of the strategies encouraged by Best Behavior. In addition, an equal number of treatment and control schools implemented peer mediation programs. While no control schools implemented RiPP in the first year, slightly more treatment schools than control schools implemented stand-alone violence prevention curricula, other than RiPP. In addition, slightly more control schools than treatment schools were using security measures such as cameras and law enforcement officers.

## Impact Findings

The main findings regarding intervention impacts after year one are the following:

- **There were no statistically significant differences between intervention and control schools on self-reported student violence or victimization.** On average, 6th-graders in the intervention schools reported engaging in 2.91 violent acts at school in the past 30 days, compared with 6th-graders in control schools, who reported engaging in 2.88 violent acts at school in the past 30 days. On average, 6th-graders in both intervention and control schools reported being victimized 4.97 times in the past 30 days.
- **There were no statistically significant impacts on self-reported violence and victimization between high-risk youth at treatment schools and high-risk youth at control schools.** On average, 6th-graders in the intervention schools who were categorized as being at a high risk for violence reported at follow-up that they had engaged in 5.16 (change from baseline of 0.13) violent acts at school in the past 30 days. This is compared with high-risk 6th-graders in the control schools who reported at follow-up that they had engaged in 4.78 (change from baseline of 0.11) violent acts at school in the past 30 days. In addition, high-risk 6th-graders in the intervention schools reported being victimized an average of 6.28 (change from baseline of 0.47) times at school in the past 30 days. This is compared with high-risk 6th-graders in the control schools who reported being victimized an average of 6.23 (change from baseline of 0.30) times at school in the past 30 days.
- **There were no statistically significant impacts on either secondary or intermediate outcomes.** In addition, after 1 year 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 aggression.
- **There were no statistically significant impacts for either boys or girls on violence or victimization.** An exploratory subgroup analysis indicated that there were no statistically significant differences between boys in the treatment schools and boys in control schools for either self-reported violence or victimization. Likewise, there were no statistically significant

differences between girls in the treatment schools and girls in control schools on either self-reported violence or victimization. Also, the difference in impacts between boys and girls was not statistically significant for either self-reported violence or victimization.

As the implementation results document, the programs being evaluated as part of the study were not fully implemented with complete fidelity in the first year. This has the potential to limit the ability to find statistically significant differences between treatment and control schools.

## 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.

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 expected 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.

The lessons in the first-year RiPP program introduce 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 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. Also, Best Behavior was adapted to reinforce and complement the RiPP curriculum.

## Evaluation Design

Designed as a group-randomized control experiment, the study was conducted 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). In year one, the intervention was delivered schoolwide by school staff trained annually by the program developers. School staff also received technical assistance and were monitored throughout implementation by trained site liaisons under the guidance of Tanglewood Research.

### 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.[2] 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 is 871 and ranges between 462 and 1,404 students. Minority students compose 65 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 56 percent and ranges between 16 percent and 97 percent. There were no statistical differences between the intervention and control groups on these characteristics.

## Student Sample

The impact analysis for the main findings uses a cross-sectional sample in each school. The first-year analysis used data collected from a census of the entering 6th-grade class in fall 2006 (baseline) and from those who remained in the study schools as 6th-graders or were new in spring 2007 (first follow-up). The research team obtained written parental consent from 6th-grade students at baseline and from new students at the first follow-up. These data were used to answer questions about the effects of the intervention across the general student population after 1 year of program implementation. Over 70 percent of enrolled students received parental consent and completed a survey at baseline, and 77 percent completed a survey at the first follow-up. A total of 7,351 students provided survey data for this analysis. More than one-half of the students in the sample were either Hispanic or Black. About one-half of the students in the sample were male, and more than one-half 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. Unlike the remainder of the students in the full sample, this sample was to be followed longitudinally; 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 during spring 2007. If a student went to a different district, the student was not followed. Over 90 percent of the students identified as high risk at baseline completed a survey at the first follow-up ($N$ = 1,938).

## 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 40 middle 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 in intervention schools. Ninety-six percent ($N$ = 917) of the sampled teachers completed a survey.

[2] In the two school districts with three recruited schools, the schools first were rank ordered on the percentage of free or reduced-price lunches, from lowest to highest. The first two schools within each district formed a pair. The third school from each district formed the last pair, across districts. Schools were then randomly assigned within each pair to treatment or control conditions.

## Data Collection and Outcome Measures

Data for the study's outcome measures to estimate intervention impacts were collected through student surveys in both treatment and control schools. Student surveys provided data to address the main impact research questions regarding school violence. Baseline data collection for students occurred in fall 2006, prior to the introduction of the intervention; follow-up student data collection occurred in spring 2007. Teacher data were collected through a survey administered to a random sample of teachers at each school in spring 2007 to assess other program impacts besides main outcomes. 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.

The primary outcomes are student violence and student victimization, both measured through student surveys. For each of these two outcomes, two additional sub-indices were created to better understand any differences between intervention and control schools with regard to specific types of violence and aggression.

A second set of indices was created to examine possible secondary effects from the intervention (e.g., spillover effects), 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., aggression, victimization) as a function of condition (treatment 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 treatment schools relative to changes among high-risk youth in control schools. These models use data collected on the same sample of students at each measurement occasion. The repeated measures models for the high-risk sample 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 first follow-up for treatment 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.

The results are presented in terms of event rates (ERs) and event rate ratios (ERRs). ERs indicate 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) to 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 occurrences 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.

After 1 year of program implementation, there were no significant differences between the students in intervention and control schools on violence and victimization, both overall and for specific types of violence and victimization. These data are reported in table ES-2. In addition, there were no statistically significant program impacts among the subpopulation of high-risk youth, as measured by student violence and victimization (table ES-3).

**Table ES-2. Main program impacts on self-reported violence and victimization—Year one**

| Self-reported student outcome | Model-adjusted follow-up event rates (SE) | | Estimated impact[1] (95% CI) | Wald Chi-Square *p*-value |
|---|---|---|---|---|
| | Intervention group | Control group | | |
| Violence[2] (All items) | 2.91 (1.03) | 2.88 (1.03) | 1.01 (0.93, 1.10) | 0.79 |
| Without a weapon | 2.79 (1.03) | 2.76 (1.03) | 1.01 (0.93, 1.10) | 0.74 |
| With a weapon | 0.10 (1.09) | 0.11 (1.09) | 0.95 (0.78, 1.15) | 0.58 |
| Victimization[3] (All items) | 4.97 (1.02) | 4.97 (1.02) | 1.00 (0.95, 1.06) | 0.99 |
| Overt | 2.88 (1.02) | 2.86 (1.02) | 1.00 (0.94, 1.07) | 0.90 |
| Relational | 2.07 (1.02) | 2.11 (1.02) | 0.98 (0.93, 1.03) | 0.45 |
| Sample size (Schools) | 20 | 20 | | |
| Sample size[3] (7,351 students clustered within schools) | 3,619 | 3,732 | | |

[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] Based on count data.

[3] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 1 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 2007 (first follow-up).

**Table ES-3. Main program impacts on self-reported violence and victimization—Year one: High-risk subgroup (Via repeated measures)**

| Student self-reported outcome[1] | Model-adjusted baseline event rates[2] (SE) | | Model-adjusted follow-up event rates[2] (SE) | | Estimated impact (95% CI)[3] | Wald Chi-Square *p*-value |
|---|---|---|---|---|---|---|
| | Intervention group | Control group | Intervention group | Control group | | |
| Violence (All items) | 5.03 (1.03) | 4.67 (1.03) | 5.16 (1.03) | 4.78 (1.03) | 1.00 (0.92, 1.10) | 0.95 |
| Without a weapon | 4.76 (1.03) | 4.40 (1.02) | 4.89 (1.03) | 4.51 (1.03) | 1.00 (0.92, 1.09) | 0.95 |
| With a weapon | 0.26 (1.10) | 0.26 (1.10) | 0.26 (1.11) | 0.26 (1.10) | 0.99 (0.97, 1.33) | 0.99 |
| Victimization (All items) | 5.81 (1.03) | 5.93 (1.03) | 6.28 (1.03) | 6.23 (1.03) | 1.03 (0.97, 1.10) | 0.34 |
| Overt | 3.52 (1.02) | 3.60 (1.02) | 3.86 (1.02) | 3.81 (1.02) | 1.03 (0.96, 1.11) | 0.35 |
| Relational | 2.28 (1.02) | 2.32 (1.02) | 2.42 (1.02) | 2.41 (1.02) | 1.02 (0.95, 1.09) | 0.60 |
| Sample size: Schools | 20 | 20 | 20 | 20 | | |
| Sample size:[4] Students within 40 schools | 1,005 | 1,148 | 897 | 1,016 | | |

[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 groups. Ratios of impact estimates of 1.00 indicate no interaction between time and program group (i.e., no program impact).

[4] Student sample sizes used in the analysis vary due to item nonresponse at baseline, follow-up, 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 survey, fall 2006 (baseline) and spring 2007 (first follow-up).

# Chapter 1. Overview of the Study

Under contract from the National Center for Education Evaluation and Regional Assistance of the U.S. Department of Education, RTI International (RTI) and its subcontractor, Pacific Institute for Research and Evaluation (PIRE), are conducting an impact evaluation of a middle school–based violence prevention program. This study examines the effects of a hybrid intervention model that combines curriculum-based and whole-school approaches. Using a group-randomized trial design, the evaluation will assess program effects on student aggression and victimization. The program was administered over the course of 3 successive school years by RTI and its subcontractor, 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 make up the intervention—Prevention Opportunities and University of Oregon—provided the program materials and conducted staff training; both are subcontractors on this study.

This chapter begins with a discussion of the federal programs that address school violence and the nature and extent of the problem of violence in the nation's middle schools. It then summarizes the research literature on the effectiveness of violence prevention programs. Finally, the chapter describes the interventions that were implemented and evaluated in the present study and then presents the key research questions that were addressed. Later chapters report findings from the first school year; findings from the second and third school years will be reported in a second report.

## 1.1 Grant Programs Supported by the Office of Safe and Drug-Free Schools

In 1986, Congress enacted the Drug-Free Schools and Communities Act (DFSCA) as Subtitle B of Title IV of the Anti-Drug Abuse Act of 1986, which has since endorsed the largest federal effort to prevent alcohol and other drug use among the nation's school-aged youth. As school safety became a more pressing concern—and after the president and the nation's governors adopted the Goals 2000: Educate America Act, P.L. 103-227, 108 Stat. 125–209 (1994), which included a goal for safe, drug-free, and disciplined schools—Congress reauthorized the DFSCA as Title I of the Safe and Drug-Free Schools and Communities Act (SDFSCA) of 1994. This reauthorization for the first time included violence prevention as a supported activity under the act. As of the end of the 1990s, state grants from SDFSCA were providing funding to approximately 97 percent of all school districts to implement programs that target youth alcohol and other drug use and also school safety. Of those, 90 percent implemented curricula that targeted youth violence (Hantman and Crosse 2000).

The next federal legislation of pertinence to this study is the No Child Left Behind Act (NCLB) of 2001, which reauthorized the Elementary and Secondary Education Act of 1965. Section 4121 of NCLB authorized the secretary of education to develop, implement, and evaluate scientifically valid, evidence-based programs to prevent violence and the illegal use of drugs among students, in response to state and local needs. NCLB legislation also included the "Principles of Effectiveness," which was originally developed by the Office of Safe and Drug-Free Schools (OSDFS) in 1998 (U.S. Department of Education 1998). The third of these principles directs

schools receiving OSDFS 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 [NIDA] 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 exists 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. School violence may be lethal, but it far more commonly involves fighting, bullying, physical attacks, and verbal threats made by students or their parents against teachers or other students. Recent data from the National Crime Victimization Survey (NCVS) show that in 2007, some 4 percent of students aged 12 through 18 reported that they had been victimized at school during the previous 6 months (Dinkes, Kemp, and Baum 2009).

Students are most likely to perpetrate or be subjected to violent behaviors in middle school grades. The rate of victimization in 2007 for students aged 12 through 14 was 35 incidents per 1,000 students, which was higher than the rate for students aged 15 through 18, which was 23 incidents per 1,000 students (Dinkes, Kemp, and Baum 2009). Similarly, data from the School Survey on Crime and Safety for the 2005–06 school year reveal that students were more likely to experience a violent event in middle schools (52 per 1,000) than in elementary (25 per 1,000) or secondary (26 per 1,000) schools (Nolle, Guerino, and Dinkes 2007). NCVS data also indicate that bullying is a significant problem. In 2005, some 36 percent of students in grades 6 and 7, and 29 percent of students in grades 8 and 9, reported having been bullied in the previous 6 months (Dinkes, Cataldi, and Lin-Kelly 2007).

## 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.

There are now at least eight violence prevention curricula that target students in middle school settings, including 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. RiPP has been subjected to four discrete evaluations by the program's developers. The

first of these comprised a quasi-experimental study administered in a single school, in which students in one 6th-grade pod received the program, while those in another did not. Compared with students in the comparison group, those in the treatment group reported less approval of violent behavior at immediate follow-up, as well as less physical aggression, lower levels of peer provocation, and increased peer support for nonviolent behavior (Farrell, Valois, and Meyer 2002). For the second study, 27 classes of 6th-graders in three urban middle schools were randomly assigned to an intervention or control group. Students exposed to RiPP scored higher at immediate follow-up on a test of knowledge and also manifested fewer in-school suspensions and disciplinary violations for violent offenses. These outcomes, however, were not maintained either 6 months or 1 year later (Farrell, Meyer, and White 2001). In a third evaluation, 21 classes of 7th-graders who had received RiPP the previous year were randomized either to an intervention group that received an additional year of RiPP or to a control group without exposure to another year of RiPP. As 8th-graders, students in the intervention group manifested fewer disciplinary code violations for violent offenses, but no effects were found on self-reported physical aggression. In a fourth evaluation that utilized a quasi-experimental design, eight schools were assigned to either an intervention or a treatment condition. Students in the intervention schools received RiPP in both the 6th- and 7th-grades and over time demonstrated lower levels of aggression and victimization, although these results were not noted consistently across four follow-up time points between the end of 6th grade and the beginning of 8th grade (Farrell, Valois, et al. 2003).

While evaluations of RiPP and similar prevention programs have yielded significant results, their effect sizes have tended to be modest. A meta-analysis of school-based violence prevention evaluations from a mix of experimental and quasi-experimental designs reported that the average effect size was $d = 0.10$ for classroom-based social skills programs (Wilson and Lipsey 2005). These findings suggest that curriculum-based programs for middle schools, in and of themselves, are unlikely to make lasting changes in students' violent behaviors.

A few school-based programs have sought to prevent violence by means of a whole-school approach. Some of these have been evaluated with middle school–age students. One example is the Olweus Bullying Prevention Program (Olweus 1983), which includes the formation of a schoolwide prevention committee that monitors areas of the school where bullying is prevalent and the establishment and enforcement of class norms and rules concerning bullying behavior. This program, which was subjected to a nonrandomized trial in a middle school setting, yielded statistically significant decreases in self-reports of taking part in bullying others, but no main effects on student victimization were observed (Bauer, Lozano, and Rivara 2007).

Another 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. One evaluation of PBS, which utilized data from three middle schools (one intervention school and two nonmatched comparison schools), demonstrated reductions in students' aggressive social behavior and referrals for discipline problems, as well as increases in perceptions of school safety (Metzler et al. 2001). A second evaluation of PBS, which also relied on a quasi-experimental design, matched nine intervention schools (six elementary and three middle schools) to six comparison schools (three elementary and three middle schools).

Intervention schools manifested reductions in office discipline referrals and improved knowledge of social skills, but no changes in perceptions of school safety were observed (Sprague et al. 2001).

Several programs have developed a hybrid approach that combines some of the characteristics of both whole-school and curriculum-based programs. One example of this approach is the Students for Peace project (Kelder et. al. 1996), which includes both a violence prevention curriculum and the formation of a School Health Promotion Council. This program was evaluated by means of a randomized controlled trial comprising eight middle schools, but yielded no intervention effects (Orpinas et al. 2000). Another example is Positive Action (Allred 1984), a K–12 program that includes a curriculum and also seeks to improve the overall climate of the schools. This program has been evaluated in middle schools, but only by means of a quasi-experimental design that relied on school-level archival data on student performance and disciplinary referrals/actions. This study of Positive Action found that middle schools with estimated higher proportions of Positive Action graduates scored significantly better than schools with proportionally fewer estimated Positive Action graduates (based on implementation of Positive Action at feeder elementary schools in prior years). Significant effects were reported for a variety of outcomes, including school achievement, absenteeism, and incidents per 100 students of problem behaviors (substance use, violence, disrespectful or disorderly conduct, and property crime) (Flay and Allred 2003). While the results of this study are promising, the methodological weaknesses inherent in quasi-experimental studies suggest the need to subject this and other promising hybrid programs to more controlled evaluations.

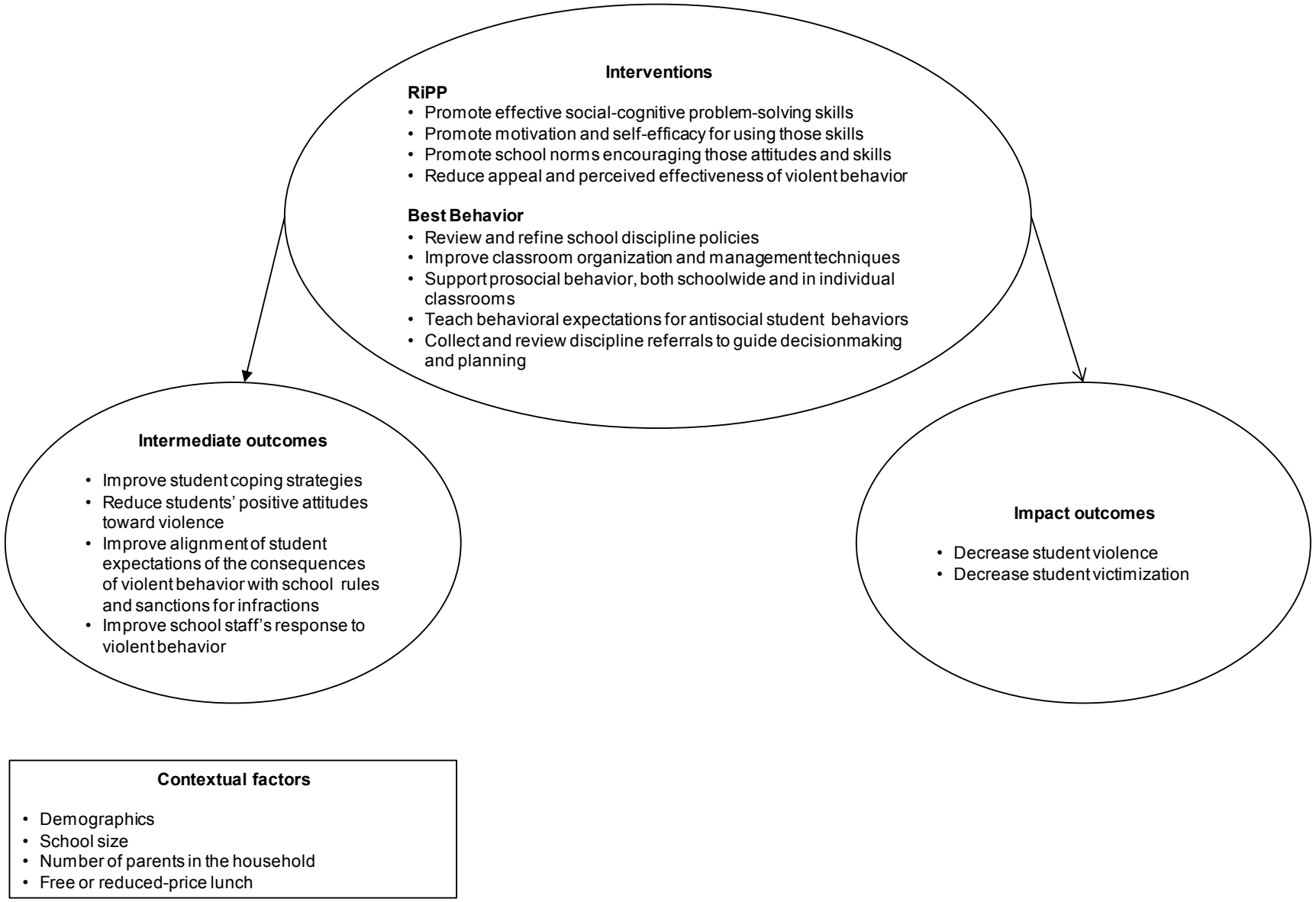
## 1.4 The Intervention

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[3]) are not designed as interventions to be used with students who are already exhibiting serious violent behavior in school. Nonetheless, the study examined outcomes among a subset of students at elevated risk for violence and aggression 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.

[3] 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**



### 1.4.1 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 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.

For this study, the RiPP curriculum comprises 16 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, as opposed to the outside professionals for whom the curriculum was originally developed. 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.

### 1.4.2 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 the school management team prior to the implementation and in follow-up years and were available throughout the study for assistance with resolving problems as they arose. 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.

## 1.5 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 |
|---|
| To be assessed at the end of each of the 3 implementation years:<br>• With the goal of decreasing disruptive, aggressive, violent, and other delinquent behaviors, what other interventions or prevention programs do the treatment and control schools implement, other than the violence prevention program under study?<br>• Is delivery of the violence prevention program consistent with its design and intended implementation?<br>To be assessed only after the last implementation year:<br>• What are the costs of the program for each participating school and student? |
| **Student outcomes and impacts** |
| To be assessed at the end of each of the 3 implementation years:<br>• Are there differences in the degree of violence in schools that implement the violence prevention program, relative to schools that do not implement it?<br>• What is the impact of the violence prevention program on students who are at elevated risk for violence and aggression?<br>To be assessed only after the final implementation year:<br>• What are the outcomes of the violence prevention program on students with varying years of program exposure or dosage? |

## 1.6 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 1 year 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 in the first year, training and technical assistance, challenges, and treatment contrast among all participating schools. Finally, chapter 4 presents findings from the impact analyses for both intermediate and impact outcomes.

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

The impact evaluation of a school-based violence prevention program examines the implementation and effects 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 and methods for selecting schools and respondents; it also provides an overview of data collection and the approaches to the data analysis.

## 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 and continued to be monitored by trained site liaisons under the guidance of the implementation subcontractor.

Student outcome data were collected 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., aggression, victimization) as a function of condition (treatment 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 and aggression, 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.

## 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 Responding in Peaceful and Positive Ways (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), 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.

Reviewers strongly recommended that the program selected should include both classroom components and whole-school approaches. 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 treatment 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) 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 refused to participate, 94 did not respond with a decision, and 34 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 during the 3 years of the study.

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 refused to participate, and 12 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 treatment 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 (RTI).

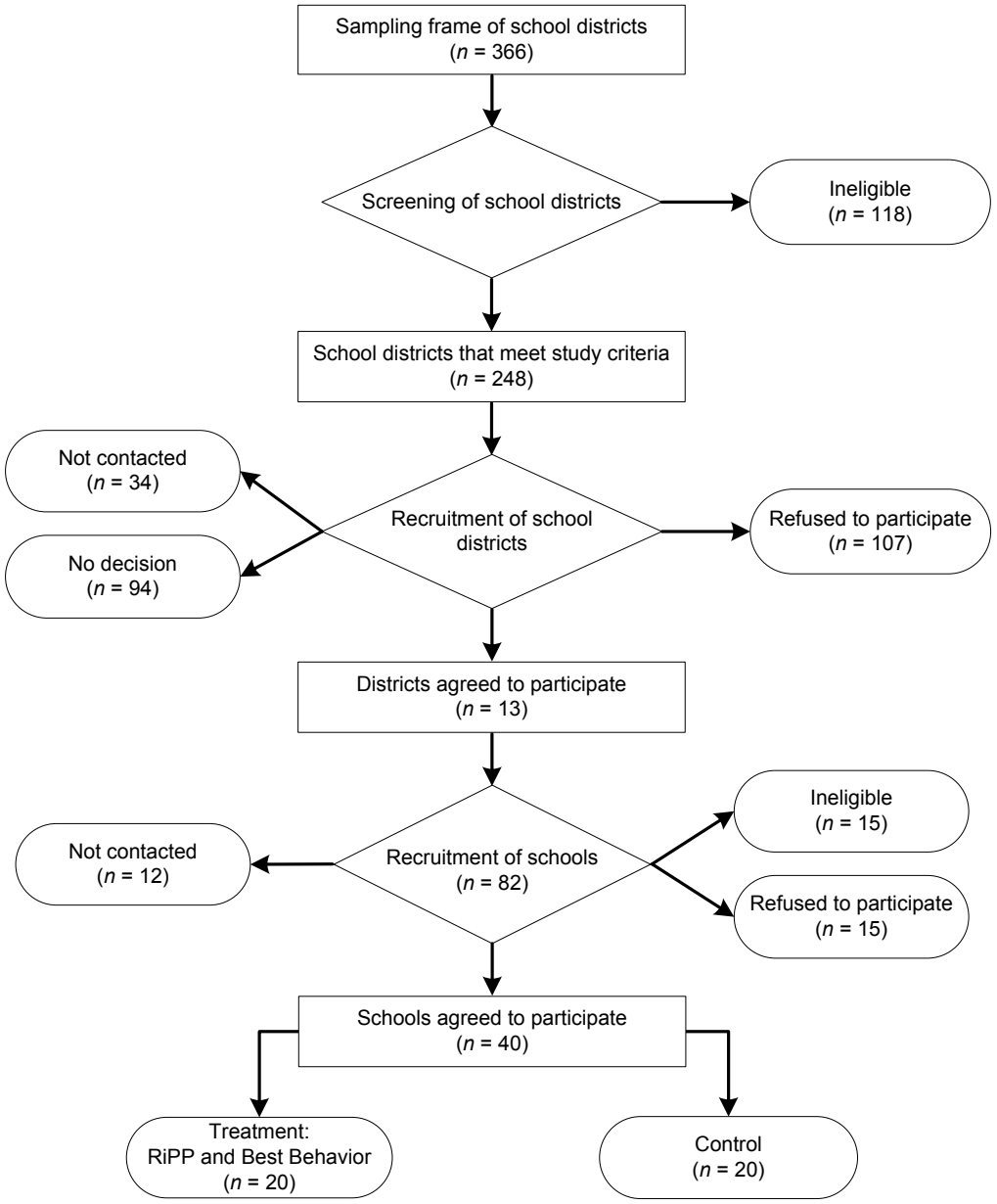
### 2.3.2 Random Assignment of Schools to Treatment

A critical feature of the impact evaluation is the random assignment of schools either to the treatment 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 treatment 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 treatment condition and the other to the control condition. Two districts included an odd number (three) of recruited schools. In these two districts, the schools were first rank ordered on the percentage of free or reduced-price lunches, and a pair was formed from the first two schools within each district. This left a third school, one in each district, which formed the last pair. As before, one school in each pair was randomly assigned to the treatment condition and the other to the control condition.

**Figure 2. Flow diagram detailing recruitment and randomization**




SOURCE: Data are from recruitment records maintained by the research team.

### 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. A majority of the districts were located in large urban or suburban areas, with only three sites in rural districts.

**Table 2. Characteristics of participating districts**

| District | Region | District location | District enrollment | Number of participating schools |
|---|---|---|---|---|
| Total | | | | 40 |
| 1 | West | Urban | Greater than 100,000 | 2 |
| 2 | Midwest | Urban | 20,000 and fewer | 2 |
| 3 | Midwest | Urban | 20,000 and fewer | 2 |
| 4 | Midwest | Urban | 20,000 and fewer | 2 |
| 5 | Northeast | Urban | 20,000 and fewer | 2 |
| 6 | Northeast | Suburban | 20,000 and fewer | 2 |
| 7 | South | Suburban | 21,000 to 100,000 | 3 |
| 8 | South | Suburban | Greater than 100,000 | 6 |
| 9 | South | Rural | 20,000 and fewer | 4 |
| 10 | South | Rural | 20,000 and fewer | 2 |
| 11 | South | Rural | 20,000 and fewer | 2 |
| 12 | South | Urban | 21,000 to 100,000 | 5 |
| 13 | South | Urban | 21,000 to 100,000 | 6 |

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.

All participating schools were middle schools that included only grades 6 through 8. As shown in table 3, the average enrollment in these schools is 871 and ranges between 462 and 1,404 students. Minority students comprise 65 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 56 percent and ranges between 16 percent and 97 percent. Intervention and control schools were compared for baseline differences using a variation on the model described in section 2.7. Baseline demographics were predicted by intervention status in mixed models with the school as the unit of analysis. There are no statistical differences between the intervention and control groups on these characteristics. 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

Our selection plan includes all 6th-graders in the 2006–07 school year, a refreshed census of all 7th-graders in the 2007–08 school year, and all 8th-graders in the 2008–09 school year. 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.

**Table 3. School demographics**

| Characteristic | All schools | Intervention | Control | Difference | *p*-value |
|---|---|---|---|---|---|
| Enrollment | | | | | |
| Mean | 871 | 848 | 895 | −47 | 0.50 |
| Range | 462–1,404 | 462–1,404 | 634–1,209 | | |
| Race/ethnicity minority (%) | | | | | |
| Mean percentage | 64.54 | 66.83 | 62.24 | 4.59 | 0.63 |
| Range | 14.70–99.90 | 15.21–99.90 | 14.70–99.83 | | |
| Students eligible for free or reduced-price lunches (%) | | | | | |
| Mean percentage | 55.64 | 55.68 | 55.59 | 0.09 | 0.99 |
| 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 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.

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, the impact analysis for the first year uses data collected from a census of the entering 6th-grade class in fall 2006, from those who remained in the study schools ("stayers") as 6th-graders in spring 2007 and from students who entered the schools during the 2006–07 school year. These data will be used to answer questions about the effects of the intervention across the general student population after 1 year of program implementation.

**Table 4. Impact data collection points for students**

| Data collection year | Grade level | | |
|---|---|---|---|
| | 6th grade | 7th grade | 8th grade |
| First year | Census of 6th grade (fall); 6th-grade "stayers," new students, and ITT high-risk subsample[1] (spring) | | |
| Second year | | 7th-grade "stayers," new students, and ITT high-risk subsample (spring) | |
| Third year | | | 8th-grade "stayers," new students, and ITT high-risk subsample[1] (spring) |

[1] Intent-to-treat (ITT) subsample of students at high risk for aggressive and violent behaviors.
NOTE: Data were collected at both treatment and control schools.

The data collection schedule follows the grade-progression pattern of the intervention program. 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 treatment schools, as compared with students in a matched group of schools that did not include the program.

Impact analyses in the second and third years will be based on the 7th- and 8th-grade classes, respectively, and will include students who remain in the study schools and also new students enrolled in each of those years whose parents provide consent for the survey. Analyses included in this report are based on measures after 1 year of program implementation only.

### 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 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 surveyed again at the end of the first implementation year. 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 during spring 2007.

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. In examining risk, the study considered both overt behaviors and self-reported attitudes and intentions regarding violence. The sample of high-risk youth includes perpetrators who reported acts of violence or aggression in the past 30 days; students who exhibit risk, based on their attitudes and intentions regarding violence; and students who fall into both of the above groups.

To identify students who fall into the perpetrator group, the study employed item response theory (IRT [Nunnally and Bernstein 1994]) and examined the 14 perpetration variables included in the baseline survey (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;
- 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 perpetrator group.

To identify high-risk students who did not report perpetration in the past 30 days, the study examined a number of 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.

Factor analysis 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 a latent class analysis (LCA [McCutcheon 1987; Collins and Lanza 2009]), 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) demonstrating a clear potential for future violence. 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.[4] These youth represent a potential perpetrator group. Combining the results of LCA and the perpetrator analysis, the research team identified 2,391 high-risk youth. Because this number is approximately 10 percent greater than the number of high-risk youth that the study team anticipated identifying and would overpower the study with regard to the high-risk sample, the researchers randomly selected 60 high-risk youth from schools that provided 71 or more potential cohort members. This approach provides a high-risk cohort that corresponds to the target minimum detectable effects (MDEs) for the high-risk sample (see appendix B). Additionally, it allowed for cost control by limiting the number of students the study needed to follow into nonprogram schools within the same districts.

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. As shown in table 4, the same subset of high-risk students identified in fall 2006 was assessed in the second and third years, based on data collected in spring 2008 and spring 2009.

### 2.4.3 Selection of the 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 40 middle 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 in intervention schools.

[4] 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). Two composite variables were formed accordingly, to represent positive or negative reactions.

## 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 baseline and follow-up at the end of the first year of program implementation are shown in table 5. In fall 2006, 70 percent of the 10,717 students in the 6th grade received written parental consent for participation in the study. A survey was completed by 99 percent of students with consent, for an overall response rate of 70 percent. Thus, the baseline sample of 6th-grade students participating in the study represented 70 percent of the total number of students enrolled in the 6th grade, based on the response rate. At the follow-up in spring 2007, the percentage of students receiving consent increased to 79 percent as new students were invited to participate. The survey was completed by 97 percent of this group, so the overall response rate increased to 77 percent.

**Table 5. Student response rates, total sample**

| | 6th-grade enrollment | Percent consented of enrolled | Percent surveyed of consented | Percent surveyed of enrolled |
|---|---|---|---|---|
| Baseline | | | | |
| Total sample | 10,717 | 70 | 99 | 70 |
| School median | 262 | 71 | 99 | 71 |
| School range | 129–445 | 36–93 | 94–100 | 37–92 |
| First follow-up | | | | |
| Total sample | 9,872 | 79 | 97 | 77 |
| School median | 240 | 81 | 98 | 80 |
| School range | 135–399 | 58–96 | 90–100 | 52–95 |

SOURCE: Data collection records maintained by the research team.

Response rates across schools ranged from 37 percent to 92 percent at baseline, with a median of 71 percent. Three schools from the same district exhibited the lowest response rates (below 50 percent). At the first follow-up, the response rates ranged between 52 percent and 95 percent, with a median of 80 percent. Among the paired treatment and control schools, response rates differed by no more than 11 percentage points, with half the pairs differing by 5 or fewer percentage points.

As shown in table 6, a two-tailed *t*-test, derived using multilevel regression, indicated that the mean response rates between treatment (71 percent) and control groups (70 percent) were not statistically different at baseline or first follow-up (77 percent and 78 percent, respectively). Student mobility in and out of study schools was examined by measuring the rate of students who exited or entered the sample at the first follow-up. Across all schools, there was a 30 percent increase in new students for whom consent was attempted at the first follow-up and a 5 percent loss of sample students from baseline due to students transferring out of study schools. These rates were not statistically different for treatment and control groups.

**Table 6. Student response and mobility rates, by group**

| | Total | Intervention | Control | Difference | *t*-test | *p*-value |
|---|---|---|---|---|---|---|
| Response rate at baseline | 70.75 | 71.19 | 70.31 | 0.90 | 0.22 | 0.8278 |
| Response rate at first follow-up | 77.33 | 76.79 | 77.88 | −1.08 | −0.29 | 0.7765 |
| Rate of exiting[1] students | 4.63 | 4.83 | 4.44 | 0.39 | 0.27 | 0.7920 |
| Rate of entering[2] students | 30.12 | 30.75 | 29.50 | 1.25 | 0.16 | 0.8736 |

[1] Exiting students are defined as those who consented at baseline but left the school by the first follow-up.

[2] Entering students are defined as those who were identified as new to the school at the first 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 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 first follow-up, 90 percent of the 2,153 high-risk students completed a survey, while 3 percent did not complete a survey due to refusal or repeated absences, and 7 percent did not complete a survey because they moved out of the school district. Among treatment and control schools, the average school nonresponse rate for high-risk students at the first follow-up was 11 percent in both groups; rates were not statistically different across the two groups ($t = -0.55$, $p = 0.580$).

### 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-six percent of the 960 teachers sampled in spring 2007 responded to the survey, for a total of 917 teachers. The teacher sample comprised approximately equal numbers of teachers in treatment ($N = 465$) and control schools ($N = 452$). Average school-level response rates were 97 percent and 94 percent for treatment and control schools, respectively; this difference was not statistically significant ($t = 0.83$, $p = 0.411$). Among responding teachers, 11 percent had taught the RiPP curriculum during the year, while the majority (89 percent) were non-RiPP teachers.

## 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 treatment 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 attendance, 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**

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

### 2.6.1 Student Surveys

Student surveys were conducted in both treatment and control schools. Questions asked about students' social competency skills, aggressive or disruptive conduct, attitudes toward violence, victimization by violence or bullying, 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 all 6th-graders in the fall and spring semesters of the 2006–07 school year and included students in the high-risk subsample if they remained at the same school. Students in the high-risk subsample who had left the school they attended during baseline were still surveyed, if possible, during the follow-up, 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:

- aggression,
- victimization,
- safety concerns,
- behavioral expectations,
- prosocial behaviors,
- self-reported coping strategies, and
- attitudes toward violence.

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 2007; a new random sample of teachers was selected in each of the two subsequent years. 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 study schools. The study is 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 will compare 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) prior to the study have been excluded. The violence prevention coordinator 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 in winter 2007 so that respondents could describe activities that were ongoing. Subsequent interviews were conducted once each year during the 2007–08 and 2008–09 school years.

### 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 treatment schools. 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.

### 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. 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.

### 2.6.6 Curriculum Implementation Records

Assessment of RiPP implementation records provided information on implementation of the RiPP curriculum and was collected only in treatment schools. The study provided RiPP teachers with standard reporting forms to record implementation information, including program attendance, number of participants, and sessions covered or topics addressed. Reporting forms were collected after every fourth lesson during implementation in each treatment 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 treatment school classrooms in which RiPP had been implemented. Observations were designed to minimize disruption to classroom implementation of RiPP.

### 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 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.

### 2.6.9 Year-End Implementation Reports (Liaisons)

At year end, the implementation team rated each school on its progress with Best Behavior, based on the expected progression with the program's key elements, and the extent of implementation of RiPP across all grades. These data provided additional information to that collected by the evaluation team regarding the fidelity and implementation of the programs.

## 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).

RiPP fidelity criteria are presented in table 8. Program objectives for RiPP included the delivery of all 16 of the curriculum's lessons to all three grades targeted. Additional fidelity criteria included 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. An additional fidelity criterion for RiPP was 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. Another fidelity criterion was that teachers were expected to use the appropriate RiPP materials during each of the lessons. Finally, the study measured student engagement to gauge the extent of participant (student) responsiveness to the lessons.

**Table 8. RiPP fidelity criteria—Year one**

| RiPP fidelity criteria | Measure | Source |
|---|---|---|
| All first-year lessons completed | Number of RiPP lessons delivered by year's end | Curriculum implementation records |
| Teachers follow the scripted lesson plan | The teacher follows the written plans for the lesson; minimizes the extent to which activities are added, changed, or omitted; and presents the material in the sequence as written | Evaluation team classroom observations |
| Teachers use appropriate RiPP materials for the lesson | The teacher uses or displays the appropriate materials during the lesson | Evaluation team classroom observations |
| Teachers use the specific RiPP teaching techniques prescribed in each lesson | The teacher uses the teaching modalities outlined in the lesson plan | Evaluation team classroom observations |
| Students are engaged during the lesson | The students are engaged in lesson activities, exercises, or discussions | Evaluation team classroom observations |

The fidelity criteria for Best Behavior are shown in table 9. For Best Behavior, a primary objective 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. Another aspect of fidelity for Best Behavior was the active support and commitment of the school principal for the program. This support should be communicated at faculty meetings and other contexts—and through attendance at the Best Behavior training workshops and the school Best Behavior management team meetings, whenever possible. A third 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. These needs assessments were essential for formulating measurable goals that were made part of the action plans for the year. A fourth objective for Best Behavior was to implement a positive reward system to reinforce desirable behaviors. 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. A final objective for Best Behavior was to communicate the schoolwide expectations and rules for behavior that the team had worked to create and define, both for the school in general and for specific settings in the school (e.g., classroom, cafeteria, hallways). These 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.

**Table 9. Fidelity criteria for Best Behavior—Year one**

| Best Behavior fidelity criteria | Measure | Source |
|---|---|---|
| The management team meets as a group at least once a month | Frequency of management team meetings | Monthly Implementation Progress Reports |
| The school principal is committed to and supportive of Best Behavior | The principal is actively involved in the school management team through support, promoting the program, attending the training, and attending the meetings | Year-End Implementation Reports |
| A schoolwide needs assessment is conducted | A schoolwide survey of school staff is conducted, and results are reviewed to identify priorities for improvement of the school discipline system | Year-End Implementation Reports |
| A discipline/reward system has been developed | The positive reinforcement system is in the school | Year-End Implementation Reports |
| The school rules/rewards system can be easily understood | Clarity of the rules and rewards system | Teacher Survey |

## 2.8 Construction of Impact 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.

### 2.8.1 Measurement Model

The survey measures for aggression 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 aggression 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 aggression 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 allows the assessment of the absence/presence of each event. A simple sum score can then be 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 sub-indices were created to better understand any differences between intervention and control schools with regard to specific types of violence and aggression.

#### *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." 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 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 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, 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 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.

#### *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. 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.

#### *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.

#### *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.

#### *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.

### 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.

#### *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.

#### *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. 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. The first scale, interactions with victims of aggression, 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.

## 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., aggression, victimization) as a function of condition (treatment 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:

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

This indicates that, given the predicted value $\mu_{i:kp}$, the outcome measure of student $i$ ($i = 1 \ldots m$) located in the $k^{th}$ 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, $\sigma^2$. The expectations of these values are expressed as:

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

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 $k^{th}$ condition ($k = 0, 1$) and the $p^{th}$ pair ($p = 1 \ldots 20$), over a given period of time ($\omega_p$). To simplify the following, we note that $\omega$ is a nonvarying time period in this study ($\omega_p$ = 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:

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

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, $\text{Var}\left(Y_{i:kp} \mid \mu_{i:kp}\right)$ 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

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

### 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,

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

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 (3) it follows that

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

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

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

For outcomes that meet the assumptions of linearity, we employ general linear mixed models where the expectation for $Y_{i:kp}$ in (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 (9) is the appropriate form.

### 2.9.3 Matched Nested Cross-Sectional Models for Full Sample and Gender Subgroups Analyses

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 $p^{th}$ pair and the $k^{th}$ condition is the same as noting that the student is nested in the $j^{th}$ 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 treatment 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 $i^{\text{th}}$ student nested in the $p^{\text{th}}$ pair and the $k^{\text{th}}$ condition. The model includes an intercept parameter $(\beta_{0:kp})$, which can be interpreted as the mean response across all students located in the $j^{\text{th}}$ 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:[5]

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

**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 ($\bar{Y}_{(t-1)}$).

An indicator variable (COND) identifies each school as a member of either the treatment or control condition; its coefficient, $\gamma_{0:01}$, accounts for the sum of the marginal differences among the pairs of treatment and control schools. The model includes two random effects. The first random effect, ($u_{0:0p}$), 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.

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

### 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.

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

Fixed effects associated with gammas ($\gamma$) represent school-level effects, while those associated with betas ($\beta$) represent student-level effects.

$$\eta_{i:kp} = \gamma_{0:00} + \gamma_{0:01}\text{COND} + \gamma_{0:02}\text{SIZE} + \gamma_{0:03}\overline{Y}_{(t-1)}$$
$$+ \beta_{1:kp}\text{GENDER}_{0:kp} + \beta_{2:kp}\text{RACE}_{0:kp} + \beta_{3:kp}\text{HH}_{0:kp} + \boldsymbol{u}_{0:0p} + \boldsymbol{u}_{0:kp} + \boldsymbol{\varepsilon}_{i:kp} \quad (12)$$

As shown in (12), $\boldsymbol{\varepsilon}_{i:kp} + \boldsymbol{u}_{0:0p} + \boldsymbol{u}_{0:kp}$ 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):

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

### 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 aggression and victimization among high-risk youth in treatment 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 $i^{th}$ student measured on occasion *t*, nested in the $j^{th}$ school of the $k^{th}$ 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:[6]

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

**Student-level models (level two).** At the student level, each of the parameters ($\beta$) 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 $i^{th}$ student in the $j^{th}$ school of the $k^{th}$ 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.

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

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

The second student-level model in equation (16) describes $\beta_{1i:j:k}$, the change or growth over time, of the $i^{th}$ student in the $j^{th}$ school of the $k^{th}$ 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 $j^{th}$ school of the $k^{th}$ 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 treatment 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

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

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 $j^{\text{th}}$ school of the $k^{\text{th}}$ 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 treatment 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 ($\lambda$) represent school-level effects, while those associated with gammas ($\gamma$) represent student-level effects.

$$\begin{aligned}\eta_{ti:j:k} &= \lambda_{00:0:k} + \lambda_{00:1:k}\text{COND} + \lambda_{00:0:k}\text{TIME} + \lambda_{00:0:k}\text{TIME*COND} \\ &+ \lambda_{00:2:k}\text{SIZE} + \gamma_{00:j:k}\text{SEX} + \gamma_{00:j:k}\text{RACE} + \gamma_{00:j:k}\text{HH} \\ &+ \mathbf{u_{0i:j:k}} + u_{00:j:k} + u_{10:j:k}\mathbf{TIME} + \left[u_{1i:j:k}\mathbf{TIME} + e_{ti:j:k}\right]\end{aligned} \tag{19}$$

In (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, $\boldsymbol{u_{0i:j:k}} + \boldsymbol{u_{00:j:k}} + \boldsymbol{u_{10:j:k}}\mathbf{TIME} + \boldsymbol{e_{ti:j:k}}$ 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:

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

## 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.

HLMs account for the correlation of teachers within schools and for schools within matched pairs assigned to condition. Covariates, including demographic variables on teachers (e.g., gender, tenure, and education) and school-level characteristics (e.g., rural vs. urban, number of policy violations by students), may be added to the model; variables that account for significant reduction in systematic variation or those that alter the magnitude of the coefficient identifying program impact by more than between 10 percent and 15 percent are retained.

### 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.

## 2.10 Observed Precision

Statistical power calculations guide a number of key decisions in the design and execution of the research study.[7] 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.

[7] 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 D.

The models employed for these analyses partition the total variation into within-subjects (i.e., school) variation, represented by the term $\varepsilon_{i:kp}$,[8] and between-subjects variation, represented by the terms $u_{0:0\,p}$ 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:

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

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 $\lambda$, the mean of the outcome. The between-subjects components of variation are scaled as equation (21):

$$\sigma^2_{u_{0:kp}(\text{scaled})} \cong \frac{\sigma^2_{u_{0:kp}(\text{unscaled})}}{\lambda^2_{i:kp}} \text{ and } \sigma^2_{u_{0:0\,p}(\text{scaled})} \cong \frac{\sigma^2_{u_{0:0\,p}(\text{unscaled})}}{\lambda^2_{i:kp}}, \quad (21)$$

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 (22):

$$\sigma^2_{\varepsilon_{i:kp}(\text{scaled})} \cong \frac{\sigma^2_{\varepsilon_{i:kp}(\text{scaled})}}{\lambda^2_{i:kp}}, \quad (22)$$

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 treatment 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.

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

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 nl(1.00) = 0.00. This fact, a statistical necessity, means that direct translation of standardized MDEs to ERR-based MDEs is not straightforward.

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

MDE for violent behavior (all items), 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 1.13 to 1.00 (negative program effect). As an example, 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 treatment schools would have to be 2.56 events in the past 30 days *or less* to achieve statistically significant positive 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**

| | ICCs | | ERRs as MDE[1] | |
|---|---|---|---|---|
| Self-reported student outcome | Unadjusted | Adjusted | Positive program effect | Negative program effect |
| Violence (All items) | 0.01109 | 0.006269 | 0.89 | 1.13 |
| Without a weapon | 0.01178 | 0.006713 | 0.89 | 1.13 |
| With a weapon | 0.0042e-10 | 0.0067e-2 | 0.74 | 1.35 |
| Victimization (All items) | 0.01467 | 0.007296 | 0.92 | 1.08 |
| Overt | 0.01540 | 0.007092 | 0.91 | 1.10 |
| Relational | 0.01089 | 0.005275 | 0.93 | 1.08 |

[1] Values below 1.00 indicate positive program effects; that is, event rates (ERs) in treatment 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 treatment schools.

SOURCE: Student survey, spring 2007.

# Chapter 3. Implementation of the Violence Prevention Program

This chapter provides details about the implementation of the combined intervention in the 20 treatment schools. 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 concludes with a comparison of programs offered in intervention schools and control schools for violence prevention or similar goals.

The key descriptive findings regarding the first year of implementation of the curriculum portion of the program include the following:

- **The first-year curriculum was delivered in its entirety to a majority of assigned classrooms.** Seventy percent of schools delivered all 16 lessons to all classrooms where the curriculum had been assigned, while another 15 percent of schools delivered all lessons in at least three-fourths of these classrooms.

- **Teachers in the majority of treatment schools followed the curriculum lesson scripts and adhered to the prescribed teaching strategies.** Teachers in 65 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. In addition, teachers in 55 percent of the intervention schools delivered lessons with few or no deviations from the prescribed teaching strategies (e.g., using role plays or small group discussions). Observations made by the liaisons were comparable for alignment with the teaching techniques (65 percent for liaisons, as compared with 55 percent for the evaluation team) but less so for deviations from the lesson plan (40 percent for liaisons, as compared with 65 percent for the evaluation team).

- **In a majority of schools, students were engaged with the curriculum.** Based on evaluation team observation, students were found to be engaged during the lesson activities, exercises, and discussions in 85 percent of the intervention schools. Observations made by the liaisons were comparable and indicated that students were engaged with the curriculum in 60 percent of the intervention schools.

The key descriptive findings regarding the first year of implementation of the whole-school portion of the program include the following:

- **One-half of the treatment schools had principals who were supportive of the whole-school portion of the program.** Principals at 50 percent of the treatment schools were rated as supportive on three out of four indicators, according to liaisons who helped implement the program.

- **The school management teams, charged with developing and disseminating the school rules and reward systems, met less frequently than stipulated by the program.**

Although school management teams are to meet monthly during the school year, the average team met on five occasions during the first year of implementation.

- **By the end of the first year, the majority of treatment schools had instituted behavioral rules and rewards as part of the whole-school portion of the intervention. In addition, a majority of teachers agreed that the rules were well defined and that they were clear with regard to the behaviors being targeted.** By the end of the first year, 75 percent of treatment schools had school rules posted in the school, 75 percent had instituted a token reward system for adhering to school rules, and 50 percent had the school rules taught in classrooms. In addition, 84 percent of teachers at intervention schools agreed or strongly agreed that school rules were clearly defined, 79 percent agreed or strongly agreed that rules emphasized rewarding desired behaviors, and 69 percent agreed or strongly agreed that rules emphasized consequences for undesired behaviors.

## 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,[9] 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, Responding in Peaceful and Positive Ways (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

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

- 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.

The RiPP developers also revised their curriculum so that the number of lessons in the program was reduced. This was done to address concerns about the competing demands on classroom time. This 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 vs. 49 lessons).

The revised Best Behavior whole-school approach and the revised RiPP curriculum were piloted in a middle school in North Carolina during the 2005–06 school year. The implementation team worked very closely with the school management team and teachers during the pilot year 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. The key components of the RiPP curriculum are described below.

#### 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. In the second year of the project, 6th-graders would also receive the RiPP first-year curriculum, while the 7th- and 8th-graders would both receive the RiPP second-year curriculum. In the third year of the project, each grade would receive 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) would have 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; and
- posters of the RiPP problem-solving model to display in the classroom.

Table 11 provides a description of the curriculum lessons and teaching techniques for RiPP-6. Included with each lesson 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 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.

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). Students 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, participants 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.

**Table 11. Content and sequence of the first-year RIPP curriculum lesson plans**

| Content | Interactive techniques |
|---|---|
| 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 |
| 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 |

### Teaching Strategies

RiPP lessons are highly interactive. 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 learn what students already know. The teacher may also learn what students erroneously think of as true. Brainstorming is fast paced and is often 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 cooperative learning is to encourage students to increase learning by allowing students to gain from each others' efforts. Each person has different 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, which creates a positive pressure to perform and is highly motivating.

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. Typically, role play is not sufficient for building mastery once new skills have been explained and demonstrated. Practice in real-world settings is needed for true mastery to emerge. However, role play does help students gain a mastery of technique.

In RiPP, teachers use interactive techniques 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.

## 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. It is based on the Positive Behavior Support approach (PBS) (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/). 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.[10] 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.

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

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

necessary skills that will ensure desired behavior change; (4) effective incentives and motivational systems provided to encourage students to behave 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.

### 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 comprising a core group of staff: namely, the school principal or other senior-level building administrator with authority to make policy decisions; representative teachers from each grade; guidance counselors; school psychologists; and other stakeholders. Usually, the school management team is formed by the school principal in consultation with the implementation support team. 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:

1. **Form a school management team.** A representative building leadership team is formed to guide program evaluation and 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.
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.
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, training was designed so that participants 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 (80 percent) relative to didactic instruction (20 percent). 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 and fall prior to the beginning of the school year. 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 lasted for 2 days and was led by two trainers working together. The trainers 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.

Overall, 188 teachers were trained to administer the RiPP curriculum in the first year. Seventy-nine percent was the site-level average of teachers trained during the developer-led sessions, while 21 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).

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

| | Portion of the intervention | |
|---|---|---|
| Type of training | Curriculum[1] | Whole-school[2] |
| Group training led by developer | 79 | 89 |
| Alternative training[3] | 21 | 11 |

[1] For the curriculum, designated staff includes teachers assigned to teach RiPP.

[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.

SOURCE: Implementation team's attendance records for training.

For RiPP, the need for alternative training 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 and distributed to all liaisons. One teacher who implemented the curriculum schoolwide to all grade levels was trained by the implementation team over multiple telephone calls and visits to her school because she needed to begin teaching before the RiPP developers could arrange a visit to her school.

### 3.2.2 Best Behavior Training

The goal of the 2-day training 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 aggression. 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 lasted for 2 days and was led by one of two program developers. A total of 216 school

staff were trained in the skills required to form a functioning school management team. 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 (see 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 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, 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:

> Meeting facilitation. 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.

Liaisons hired for the project were a group of experienced former or current educators. All but three were retired teachers or principals—and thus fully appreciated the challenges that teachers

and administrators encountered as they implemented RiPP and Best Behavior in their schools. One liaison was a graduate student who had also worked on another research study at the school, one worked for an education consulting firm and was familiar with health and character education programs, and one was a college-level educator. Eleven of the 13 liaisons were recommended by the principal or district coordinator; 1 was recommended by another liaison; and 1 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 seven of them frequently served as mentors for new teachers in the district.

## 3.3 Program Implementation

This section reports the findings for program fidelity, assessed in treatment 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 section begins by describing the contextual background for the treatment 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 three specific issues that site liaisons had documented in the Monthly Implementation Progress Reports as having resulted in challenges for program delivery: administrator turnover, failure to meet adequate yearly progress (AYP) on academic end-of-grade testing, and budgetary constraints. The specific challenges discussed in this section were brought up in response to the question, "What, if any, are the issues at the school that are impacting implementation of Best Behavior and RiPP?"

#### Administrator Turnover

During the first year of the project, 7 of the 20 principals at treatment schools were so new that they were not the principal who had agreed to participate in the project, as documented on Monthly Implementation Progress Reports. In addition, two principals were absent for extended time periods due to personal reasons. The Best Behavior developers (Sprague and Golly 2005) underscore that the 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.

#### Failure to Meet Annual Yearly Progress

Seven of the 20 treatment schools were on academic probation and did not meet standards for AYP during the first year of the project, as documented on the liaisons' Year-End Implementation Reports. According to the Monthly Implementation Progress 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.

#### Budgetary Challenges

Based on the Monthly Implementation Progress Reports completed by site liaisons, all 20 treatment 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, a major component of the Best Behavior program. For example, five schools had to obtain funding for Best Behavior rewards from community sources. One school 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 pay them overtime or otherwise compensate them. In three schools, there were academic booster programs (e.g., Athena, Avid, and Renaissance) that competed with Best Behavior for privileges.

### 3.3.2 RiPP Delivery and Fidelity

Delivery of RiPP began in early fall 2006, once RiPP training and the student baseline surveys were completed.[11] All schools in the treatment condition began delivering RiPP within 3 months (range of 1 to 3 months) of the beginning of the school year. Table 13 displays the variety of classes selected by the schools for delivering the RiPP curriculum. During the first year of implementation, RiPP was taught most often as part of social studies (50 percent) or health (25 percent). Other subjects chosen were physical education (20 percent) and science and language arts (each 15 percent). Six out of 20 schools combined several of these 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.

**Table 13. Placement of RiPP curriculum in classes**

| Class | Percent of intervention schools |
|---|---|
| Health | 25 |
| Physical education | 20 |
| Social studies | 50 |
| Science | 15 |
| Language arts | 15 |
| Allied arts (e.g., art; technology education) | 5 |

NOTE: Percentages total more than 100 percent because schools could implement in more than one type of class.
SOURCE: RiPP implementation records.

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 treatment schools received the program by the end of the first semester, and the remainder received the program by the end of the school year.

#### Amount of RiPP Curriculum Delivered

One criterion for RiPP faithful implementation is delivery of the entire first-year curriculum. Table 14 shows the extent to which intervention schools delivered the full program (all 16 RiPP lessons) in the first year of implementation. Delivery of the full program in a school varied to the extent that individual classes did not complete all lessons; however, 19 out of 20 treatment schools

[11] There was one exception where the curriculum was begun prior to the baseline survey. In this case, the curriculum was delivered by a single teacher in one treatment school, and it was necessary to begin teaching the program to the first group of students at the beginning of the school year so that the teacher would have time to complete delivery to all students by the end of the year.

delivered all RiPP lessons to a majority of the assigned classrooms. Specifically, 14 of the study's 20 intervention schools (or 70 percent) delivered all 16 lessons to all classrooms; 3 of the schools delivered the full program to at least three-quarters of the classrooms; 2 of the schools completed the full program in at least one-half of the classrooms; and 1 school completed the full program in less than one-half of the classrooms.

**Table 14. RiPP curriculum dosage—Year one**

| Percent of classrooms completing all 16 lessons | Number of intervention schools |
|---|---|
| 0–49 | 1 |
| 50–74 | 2 |
| 75–99 | 3 |
| 100 | 14 |

SOURCE: RiPP implementation records.

## Fidelity of RiPP Delivery in the Classroom

Several measures of fidelity concern the extent to which teachers' delivery of RiPP was aligned with the following program requirements: adherence to scripted lessons, using appropriate materials for the lessons, using the correct teaching techniques, and having students engaged during the session. Fidelity alignment for adherence to prescribed teaching methods was coded as follows: well aligned was indicated by three classes displaying the technique; moderately aligned was indicated by two classes displaying the technique; and poorly aligned was indicated by one class displaying the technique. The measure for use of correct materials was scored from one (not at all well) to seven (very well). Average scores greater than five were coded as well aligned; scores between three and five were coded as moderately aligned; and scores below three were coded as poorly aligned. The measures for adherence to scripted lesson plans and for student engagement were scored from one (not at all) to four (always/very much). 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. The delivery of the curriculum at a majority of intervention schools was well aligned on all four of the measures. Specifically, 65 percent of schools were rated as well aligned with respect to teachers following lesson instructions; 90 percent were well aligned with regard to teachers using the correct RiPP materials for the lesson; 55 percent of schools were well aligned with regard to teachers using the correct teaching techniques for each RiPP lesson; and 85 percent were well aligned with respect to student engagement during lesson activities.

**Table 15. RiPP curriculum fidelity assessments based on evaluation team observations—Year one**

| Fidelity criteria | Alignment with program requirements[1] | Percent of schools |
|---|---|---|
| Teachers follow the scripted lesson plan | Well aligned | 65 |
| | Moderately aligned | 35 |
| | Poorly aligned | 0 |
| Teachers use appropriate RiPP materials | Well aligned | 90 |
| | Moderately aligned | 10 |
| | Poorly aligned | 0 |
| Teachers use the prescribed RiPP teaching methods | Well aligned | 55 |
| | Moderately aligned | 40 |
| | Poorly aligned | 5 |
| Students are engaged during the lesson | Well aligned | 85 |
| | Moderately aligned | 15 |
| | Poorly aligned | 0 |

[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.

SOURCE: Calculations are from observations of randomly selected RiPP classes conducted by the research team.

The RiPP fidelity data collected by the evaluation team's classroom observation provided only a snapshot of the alignment of program delivery and RiPP fidelity because they were collected at each intervention school over only 1 week by observing three teachers at each school. The observational data collected by the implementation liaisons over the entire year while periodically observing all RiPP teachers were also examined to determine the extent to which they agreed with observational data from the evaluation team. These data from the liaison reports indicate that 40 percent of the intervention schools were well aligned with the written curriculum and that 65 percent of the schools were rated as well aligned with respect to adherence to the specific teaching techniques called for in the curriculum. Finally, the liaisons observed that in 60 percent of the intervention schools, students were well engaged in the session.

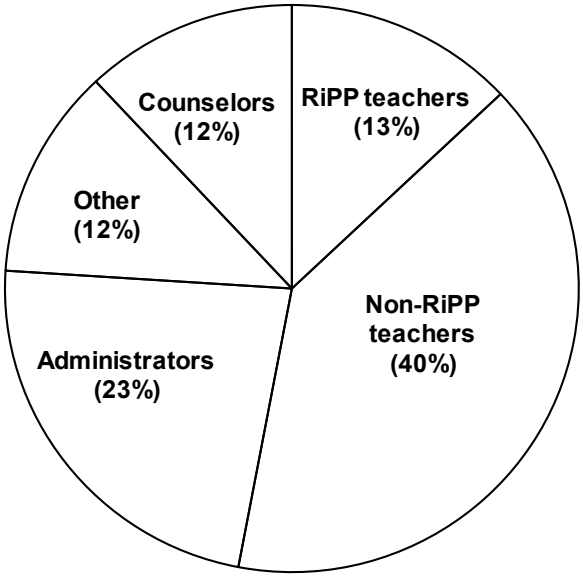
To better understand the teacher modifications to the RiPP curriculum during classroom delivery, data were also collected from RiPP teachers through interviews conducted at midyear by the evaluation team. Of the 24 percent of teachers who reported modifying the core curriculum content, 52 percent indicated that they had left out portions of lessons, and another 24 percent reported condensing the material for a lesson. These modifications took place because, according to some teachers, there was insufficient time to complete RiPP lessons. Indeed, less than one-half (45 percent) of teachers reported that the time was sufficient to cover all the material in any given lesson.

### 3.3.3 Best Behavior Program Fidelity

#### School Management Team Composition

School management teams ranged in size between 7 and 16 members, with an average of 10 members. Figure 3 shows the site-level averages with regard to the composition of the Best Behavior school management teams in the first year of implementation. On average, non-RiPP teachers constituted 40 percent of teams; RiPP teachers constituted 13 percent (to facilitate alignment between RiPP and Best Behavior); administrators accounted for 23 percent; counselors made up 12 percent; and others constituted 12 percent.

**Figure 3. Composition of the Best Behavior teams**




NOTE: *N* = 209. Percentages are based on site-level averages.
SOURCE: Implementation team records.

#### Frequency of School Management Team Meetings

The Best Behavior program was considered to start implementation the month that the school management team began having regularly scheduled meetings. In the first year of Best Behavior implementation, during the 2006–07 school year, the average lapse between the beginning of the school year and Best Behavior implementation was 3 months (range of 1 to 7 months). Three schools were delayed by 7 months, and one school was delayed 6 months.

One aspect of fidelity is that the team at each school meets at least monthly. 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 5 times (range of 2 to 13 times). Fifty percent of the schools met five or more times during the 9 months. One school held 13 meetings as a result of meeting weekly during periods of increased program activity.

#### Principal Support of Best Behavior Program

Principal support is another aspect of program fidelity for Best Behavior. 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. 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 16 summarizes the level of principal support and commitment, as measured by perceived support (by the liaisons) (80 percent), perceived use of leadership to promote the program (65 percent), and whether or not the principal attended the Best Behavior training workshop (80 percent) and was a member of the school management team (55 percent). Of the 20 principals, one-half (10) were supportive in three or four ways, while 6 principals were only supportive in two ways, and 4 principals gave minimal support.

**Table 16. Principal support of Best Behavior—Year one**

| Type of principal support | Number of schools | Percent of schools |
|---|---|---|
| Principal was supportive (as perceived by the liaisons) | 16 | 80 |
| Principal used leadership to promote program (as perceived by the liaisons) | 13 | 65 |
| Principal attended Best Behavior workshop | 16 | 80 |
| Member of school management team and attended meetings on occasion | 11 | 55 |

SOURCE: Year-End Implementation Report for the 2006–07 school year. Training records maintained by the implementation team.

### Needs Assessment Completed

Completion of a schoolwide needs assessment was an additional fidelity criteria for year one. The status of this milestone is shown in table 17. According to liaisons' Year-End Implementation Reports, by the end of year one, a needs assessment was completed by 95 percent of the schools.

### Creation of Schoolwide Rules and Rewards System

The creation of a schoolwide rules and rewards system, another fidelity criterion, was measured through liaisons' observation of whether schools completed several tasks, which can be found in table 17. An early task for the school management team was to identify and adopt a set of rules of behavior by which all of the school could be guided. At the end of the first year, 80 percent of schools had completed this task, while 15 percent of schools were still working on it, and one school (5 percent) had not yet begun this task. Seventy-five percent of treatment schools had posted schoolwide rules and expectations for behavior in their respective schools. A related subsequent task was to develop lesson plans to teach students and faculty about each of the selected rules and expectations. This development task was completed by 60 percent of schools by the end of year one, while 50 percent of schools had used teachers to teach the rules and expectations systematically to students in the classroom. Regarding the reward system for recognizing appropriate behavior, 80 percent of schools completed development of the system, while slightly fewer (75 percent) implemented it by the end of the year.

**Table 17. Needs assessment; creation of schoolwide rules; and rewards system—Year one**

| Measure | Status (Percent of schools) | | |
|---|---|---|---|
| | Not initiated | In progress | Completed |
| Needs assessment | | | |
| Needs assessment completed | 0 | 5 | 95 |
| Schoolwide rules | | | |
| Rules and schoolwide behavioral expectations are defined | 5 | 15 | 80 |
| Rules posted in the school | 10 | 15 | 75 |
| Lesson plans are developed for teaching about behavioral expectations in all school settings (e.g., cafeteria, gym, restrooms) | 15 | 25 | 60 |
| Rules/expectations taught systematically | 20 | 30 | 50 |
| Rewards system | | | |
| A schoolwide system is defined for recognizing and rewarding appropriate, expected behaviors | 5 | 15 | 80 |
| Reinforcement system implemented | 5 | 20 | 75 |

SOURCE: Site liaison reports.

## 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 18 displays results from the teacher survey administered between February and May of the first implementation year. Fifty-four percent of the teachers indicated that when school rules were broken, it was clear to school staff what consequences would follow. Sixty-nine percent of teachers reported that school rules emphasized the consequences for negative behaviors, 79 percent said that their school rules reinforced desirable behaviors, and 84 percent said that rules for student behavior were clearly defined in their schools.

**Table 18. Teacher reports of clarity of school rules in intervention schools**

| | Percent of teachers agreeing |
|---|---|
| School rules for student behavior are clearly defined | 84 |
| The school rules emphasize reinforcing desired behaviors | 79 |
| The school rules emphasize consequences for undesired behaviors | 69 |
| When a school rule is broken, it is clear to school staff what consequences should follow | 54 |

NOTE: *N* = 465. Response options ranged from one ("strongly agree") to four ("strongly disagree"); "strongly agree" and "agree" responses were combined.

SOURCE: Teacher survey, spring 2007.

## 3.4 Implementation Challenges

This section describes the challenges encountered when implementing each of the two programs and how the problems faced by many of the schools contributed to these challenges.

RiPP was somewhat easier to launch than Best Behavior, perhaps because it is more structured and localized within the school. According to the Year-End Implementation Reports, the major challenge that RiPP implementation faced was limited time, which was often a result of competing priorities. This was true in all 20 schools. The first strategy used by the implementation project team to minimize this problem was addressed during the team's initial negotiations with the principal. Principals were asked to be careful to consider testing and other priorities as they identified placement of RiPP within a particular subject. However, teachers who were then asked to teach the curriculum did not always agree with their principal's decision, and this sometimes translated into teacher resistance. According to the Year-End Implementation Reports, while virtually all teachers (with two exceptions) saw the value of RiPP, one in four teachers expressed concern about the time that RiPP took away from his or her respective curricula. This was particularly true when schools decided to implement RiPP in core subjects, especially when those subject areas were tested. This was also the case when RiPP was implemented in health and physical education courses. For example, health is often very limited in schools, and health teachers have multiple mandates. In one school, the health teacher only saw students for nine 90-minute sessions during the school year, and this teacher also was responsible for covering health topics such as drug use and sexuality during that time. Physical education teachers said that students were unhappy to be in classroom activities when they were used to participating in physical activities during this class.

A general challenge to Best Behavior involved competing philosophies. Specifically, liaisons noted that some teachers and administrators in at least five schools (25 percent) had a philosophical conflict with reinforcing students for good behaviors—which is the underlying philosophy and theory (driven by research) of Best Behavior. These faculty members disagreed with the idea that they should reward students for good behavior. Some of these faculty raised concerns that rewarding desired behavior would weaken intrinsic motivation. At other times, some faculty expressed a simple objection that they should not have to reward students for good behavior.

While this philosophical opposition was relatively rare, it became most problematic when a member of the Best Behavior team expressed this view. When this happened, the Best Behavior trainers worked hard during the training workshop to address the faculty member's concerns. For example, they would present extensive research showing the benefits of rewarding expected behaviors and research showing that rewards did not diminish intrinsic motivation. The liaisons continued to address any concerns following training. Unfortunately, it is difficult to change philosophical views, and the team had only minimal success persuading faculty members who were reluctant because of philosophical differences, despite extensive efforts. We found it was more productive to try to remove skeptical faculty from the Best Behavior team. (All faculty who were oppositional were happy not to have team responsibilities.) In one case, a principal had a philosophical difference that made it impossible for Best Behavior to move forward in his school. While it was helpful to keep faculty with strong philosophical differences off of the team, it was also important to note that the presence in the school of individuals with opposing philosophical positions created problems for the program. For example, it was very difficult to get these faculty members to actively participate in giving rewards. Further, to the extent these faculty members had power and influence among their colleagues, they created tension that was not productive.

## 3.5 Treatment Contrast

This section provides information concerning the variety of interventions in the study's intervention and comparison schools that is potentially related to the study's outcomes of interest. As table 19 indicates, more treatment schools than control schools implemented stand-alone violence prevention curricula, other than RiPP, during the first year. Specifically, three intervention schools and one control school were administering a gang resistance program, while two intervention schools and one control school were administering a character education program. Individual presentations that were not part of a curriculum (most often, speakers, a video, or a lesson) provided information on bullying, harassment, and dating violence. These types of sessions were used by one school pair, three intervention schools, and two control schools. In addition, while no control school was implementing Best Behavior, an equal number of control and intervention schools used a discipline tracking system, one of the strategies encouraged by the Best Behavior.

Finally, with regard to violence prevention strategies other than curricula and whole-school approaches, slightly more control schools used security measures, such as cameras and law enforcement officers, compared with treatment schools. An equal number of treatment and control schools implemented peer mediation programs, although the majority of these occurred outside of the matched pairs of schools.

**Table 19. Number of schools implementing strategies related to violence prevention**

| Violence prevention strategy | Number of intervention-control pairs | Number of intervention schools only | Number of control schools only |
|---|---|---|---|
| Classroom-based education | | | |
| Responding in Peaceful and Positive Ways (RiPP) (Intervention) | 0 | 20 | 0 |
| Gang Resistance Education and Training (GREAT): Taught by law enforcement officers, focusing on skills to avoid gang pressure and involvement | | | |
| Grades 6 and 7; total of 13 sessions | 0 | 3 | 1 |
| Character education: Wide variety of programs focused on core ethical values, such as respect, justice, civic virtue, and responsibility | | | |
| Infused in subjects or taught in lessons (range of 1 to 12 lessons) | 1 | 2 | 1 |
| Individual sessions (lesson, speaker, or video) on bullying, harassment, dating violence; not part of a curriculum | | | |
| Number of sessions ranges from 1 to 10 per year | 1 | 3 | 2 |
| Whole-school reward/disciplinary approach | | | |
| Best Behavior (Intervention) | 0 | 20 | 0 |
| Discipline tracking system | 1 | 3 | 3 |
| Non-classroom-based prevention | | | |
| Peer mediation, conflict resolution, student court | 3 | 4 | 4 |
| Security | | | |
| Security cameras, metal detectors | 11 | 1 | 3 |
| Law enforcement officers, security guards, school resource officers | 14 | 0 | 3 |
| Controlled access, visitor check-in | 20 | 0 | 0 |

SOURCE: Violence prevention coordinator interview, midyear 2006–07.

# Chapter 4.
# Impacts of the Violence Prevention Program

This chapter presents the impact findings following the first year of a 3-year program to reduce student violence and victimization in middle schools. 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 (i.e., spillover effects): student and teacher safety concerns, student prosocial behaviors, and teacher victimization. Another group of outcomes examined are those hypothesized to be more immediately affected by the intervention, compared with the main outcomes of aggression and victimization. These include students' attitudes toward violence, students' self-reported coping strategies for dealing with aggression, students' clarity of understanding of the school rules, teacher expectations for student behavior, and school staff's response to student aggression.

This chapter begins with a description of the student characteristics at baseline across treatment 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 subsample and results of exploratory analyses for the gender subgroups. The chapter ends with a brief conclusion. Unadjusted means and standard deviations for all impact variables are included in appendix F.

## 4.1 Student Characteristics at Baseline

This section presents the demographic characteristics of the student sample and the baseline outcome measures across treatment and control groups. Table 20 shows the student demographic characteristics at baseline for the intervention and control groups.[12] These data were obtained through the baseline student survey, which was administered in fall 2006. More than one-half of the students in the sample were either Hispanic or Black. About one-half of the students in the sample were male, and more than one-half 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.

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

**Table 20. Baseline demographic characteristics of the student sample**

| Characteristic | Intervention group | Control group | Difference | *t*-test[1] | *p*-value |
|---|---|---|---|---|---|
| Sample size, grade 6 | 3,705 | 3,896 | | | |
| Race/ethnicity (%) | | | | | |
| Hispanic | 26.02 | 24.25 | 1.77 | 0.22 | 0.83 |
| Black, non-Hispanic | 27.83 | 25.25 | 2.58 | 0.29 | 0.77 |
| White, non-Hispanic | 30.16 | 36.36 | −6.20 | −0.70 | 0.49 |
| Other or mixed[2] | 16.09 | 14.19 | 1.90 | 0.86 | 0.39 |
| Gender (%) | | | | | |
| Male | 49.20 | 48.86 | 0.36 | 0.17 | 0.87 |
| Single-adult household (%) | 58.67 | 59.54 | −0.87 | −0.29 | 0.77 |

[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 0.05.

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[13] were compared with 6th-grade demographic data for the entire grade obtained from the Common Core of Data (CCD) for the same year. Comparisons for race/ethnicity should be viewed with caution, as the two sources of information used different ways to collect the data.[14] Table 21 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 White, Black, Hispanic, and male students were observed in the population, compared with the sample of study participants. Similar differences were observed among intervention schools, with the exception of the percentage of Black students, which was not significantly different between sample and population. Among control schools, there were differences only with respect to the percentage of Hispanic students, which was found to be significantly higher in the population. In all instances, the mean percentages differ by no more than 1 or 2 points between sample and population, except for the Hispanic variable, which differs by 9.5 points for the intervention group, 6.7 points for the control group, and 8.1 points overall. These results reveal that the students from whom outcome data were collected were not representative of the middle school student population on these specific characteristics.

[13] The baseline sample of 6th-grade students participating in the study represented 70 percent of the total number of students enrolled in the 6th grade, based on the response rate.

[14] 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, then about race, while the school data collected through CCD do not measure ethnicity separately from race.

**Table 21. Difference in baseline demographic characteristics between the student sample and population**

| Characteristic | 6th-grade population | 6th-grade study participants | Difference | *t*-test | *p*-value |
|---|---|---|---|---|---|
| Total sample | | | | | |
| Race/ethnicity (Mean %) | | | | | |
| Hispanic | 32.63 | 24.56 | 8.07 | 18.48 | 0.00* |
| Black, non-Hispanic | 27.55 | 26.75 | 0.80 | 2.02 | 0.05* |
| White, non-Hispanic | 34.83 | 33.34 | 1.49 | 3.41 | 0.00* |
| Gender (Mean %) | | | | | |
| Male | 50.70 | 49.00 | 1.70 | 2.98 | 0.01* |
| Intervention group | | | | | |
| Race/ethnicity (Mean %) | | | | | |
| Hispanic | 34.39 | 24.92 | 9.47 | 14.63 | 0.00* |
| Black, non-Hispanic | 28.84 | 28.26 | 0.58 | 1.03 | 0.32 |
| White, non-Hispanic | 32.09 | 30.35 | 1.74 | 2.87 | 0.01* |
| Gender (Mean %) | | | | | |
| Male | 51.29 | 49.08 | 2.21 | 2.69 | 0.02* |
| Control group | | | | | |
| Race/ethnicity (Mean %) | | | | | |
| Hispanic | 30.91 | 24.15 | 6.76 | 11.46 | 0.00* |
| Black, non-Hispanic | 26.26 | 25.24 | 1.02 | 1.81 | 0.09 |
| White, non-Hispanic | 37.58 | 36.32 | 1.26 | 2.00 | 0.06 |
| Gender (Mean %) | | | | | |
| Male | 50.14 | 48.91 | 1.23 | 1.54 | 0.14 |

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 0.05. 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 treatment 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 22, students' responses were not statistically different between intervention and control schools, with one exception.[15] 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.

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

These analyses suggest that randomization of schools to condition resulted in generally similar groups.

**Table 22. Baseline measures for the student sample: 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) | 1.97 (0.13) | 1.90 (0.13) | 0.08 | 0.41 | 0.68 |
| Violence: Weapons-related | 0.08 (0.01) | 0.08 (0.01) | 0.00 | 0.23 | 0.82 |
| Violence: Not weapons-related | 1.89 (0.12) | 1.82 (0.12) | 0.08 | 0.44 | 0.67 |
| Victimization (All items) | 3.88 (0.14) | 4.09 (0.14) | −0.21 | −1.09 | 0.28 |
| Victimization: Overt | 2.19 (0.09) | 2.31 (0.09) | −0.12 | −1.00 | 0.32 |
| Victimization: Relational | 1.69 (0.06) | 1.78 (0.06) | −0.09 | −1.16 | 0.25 |
| Other outcomes[3] | | | | | |
| Self-reported coping strategies: Positive | 2.87 (0.02) | 2.90 (0.01) | −0.03 | −1.51 | 0.14 |
| Self-reported coping strategies: Negative[2] | 0.59 (0.04) | 0.59 (0.03) | 0.00 | 0.08 | 0.94 |
| Attitudes toward violence | 2.99 (0.03) | 3.01 (0.03) | −0.03 | −0.57 | 0.57 |
| Prosocial behaviors: Extended to others | 2.93 (0.03) | 2.95 (0.03) | −0.03 | −0.71 | 0.48 |
| Prosocial behaviors: Received from others | 2.77 (0.02) | 2.78 (0.02) | −0.01 | −0.32 | 0.75 |
| Behavioral expectations | 3.16 (0.02) | 3.16 (0.02) | 0.01 | 0.24 | 0.81 |
| Safety concerns | 1.88 (0.03) | 1.96 (0.03) | −0.08 | −2.11 | 0.04* |
| Sample size (Schools) | 20 | 20 | | | |
| Sample size[4] (Students) | 3,705 | 3,896 | | | |

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to 0.05.

[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).

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 23 presents the outcome measures at baseline for students identified as high risk. 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 23. Baseline measures for the student sample: High-risk subgroup**

| Measure | Baseline event rates or scale means (SE) | | Difference | *t*-statistic | *p*-value[1] |
|---|---|---|---|---|---|
| | Intervention group | Control group | | | |
| Main outcomes | | | | | |
| Violence (All items) | 5.03 (0.13) | 4.63 (0.12) | 0.40 | 2.21 | 0.03* |
| Violence: Weapons-related | 0.28 (0.03) | 0.26 (0.02) | 0.02 | 0.46 | 0.65 |
| Violence: Not weapons-related | 4.76 (0.12) | 4.37 (0.10) | 0.38 | 2.47 | 0.02* |
| Victimization (All items) | 5.78 (0.12) | 6.00 (0.12) | −0.22 | −1.26 | 0.22 |
| Victimization: Overt | 3.52 (0.08) | 3.65 (0.08) | −0.13 | −1.25 | 0.22 |
| Victimization: Relational | 2.29 (0.06) | 2.36 (0.05) | −0.07 | −0.93 | 0.36 |
| Other outcomes[3] | | | | | |
| Self-reported coping strategies: Positive | 2.54 (0.03) | 2.63 (0.02) | −0.09 | −2.69 | 0.01* |
| Self-reported coping strategies: Negative[2] | 1.26 (0.04) | 1.20 (0.04) | 0.05 | 1.03 | 0.31 |
| Attitudes toward violence | 2.54 (0.03) | 2.60 (0.03) | −0.06 | −1.60 | 0.12 |
| Prosocial behaviors: Extended to others | 2.70 (0.03) | 2.74 (0.03) | -0.04 | -0.92 | 0.36 |
| Prosocial behaviors: Received from others | 2.69 (0.03) | 2.67 (0.03) | 0.02 | 0.48 | 0.64 |
| Behavioral expectations | 3.01 (0.02) | 3.02 (0.02) | −0.01 | −0.31 | 0.76 |
| Safety concerns | 1.94 (0.03) | 2.06 (0.03) | −0.12 | −2.45 | 0.02* |
| Sample size (Schools) | 20 | 20 | | | |
| Sample size[4] (Students) | 1,005 | 1,148 | | | |

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to 0.05.

[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).

## 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. 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.

## 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 first year of program implementation.

Analyses examining these outcomes are presented in table 24. 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 column 4 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 the first year of implementation. In each case, using the Wald statistic, results indicated that the rates of behavior reported by students in treatment schools were not statistically different from the rates of behavior reported by students in schools not receiving the intervention.

**Table 24. Main program impacts on self-reported violence and victimization—Year one**

| Self-reported student outcome | Model-adjusted follow-up event rates (SE) | | Estimated impact[1] (95% CI) | Wald Chi-Square *p*-value |
|---|---|---|---|---|
| | Intervention group | Control group | | |
| Violence[2] (All items) | 2.91 (0.09) | 2.88 (0.09) | 1.01 (0.93, 1.10) | 0.79 |
| Without a weapon | 2.79 (0.09) | 2.76 (0.09) | 1.01 (0.93, 1.10) | 0.74 |
| With a weapon | 0.10 (0.01) | 0.11 (0.01) | 0.95 (0.78, 1.15) | 0.58 |
| Victimization[2] (All items) | 4.97 (0.09) | 4.97 (0.09) | 1.00 (0.95, 1.06) | 0.99 |
| Overt | 2.88 (0.07) | 2.86 (0.06) | 1.00 (0.94, 1.07) | 0.90 |
| Relational | 2.07 (0.03) | 2.11 (0.03) | 0.98 (0.93, 1.03) | 0.45 |
| Sample size (Schools) | 20 | 20 | | |
| Sample size[3] (7,351 students clustered within schools) | 3,619 | 3,732 | | |

[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] Based on count data.

[3] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 1 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 2007 (first 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 after 1 year of program implementation, 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 aggression, and
- negative (inappropriate) self-reported coping strategies when faced with aggression.

Analyses examining these student outcomes are presented in table 25. 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).

**Table 25. Secondary program impacts on student-level outcomes—Year one**

| Self-reported student outcome[1] | Model-adjusted follow-up event rates or scale means (SE) | | Estimated impact (95% CI) | Wald Chi-Square *p*-value |
|---|---|---|---|---|
| | Intervention group | Control group | | |
| Safety concerns[2] | 1.93 (0.04) | 1.91 (0.03) | 0.02 (−0.05, 0.08) | 0.67 |
| Prosocial behaviors extended[3] | 2.88 (0.03) | 2.89 (0.03) | −0.01 (−0.07, 0.04) | 0.54 |
| Prosocial behaviors received[3] | 2.78 (0.03) | 2.78 (0.02) | 0.00 (−0.04, 0.05) | 0.86 |
| Perceived behavioral expectations[3] | 3.02 (0.02) | 2.99 (0.02) | 0.03 (−0.02, 0.09) | 0.16 |
| Attitudes toward violence[3] | 2.84 (0.02) | 2.85 (0.02) | −0.01 (−0.06, 0.04) | 0.71 |
| Self-reported coping strategies (Negative)[2,4] | 0.85 (0.02) | 0.81 (0.02) | 1.04 (0.97, 1.12) | 0.22 |
| Self-reported coping strategies (Positive)[3] | 2.81 (0.02) | 2.78 (0.01) | 0.03 (−0.02, 0.08) | 0.17 |
| Sample size (Schools) | 20 | 20 | | |
| Sample size[5] (7,351 students clustered within schools) | 3,619 | 3,732 | | |

[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 3 percent to 5 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 2007 (first follow-up).

Results for intermediate student outcomes indicated that there were no significant impacts after 1 year 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 treatment schools were not significantly different from the rates reported by students in control schools after 1 year of program implementation.

To assess the influence of the intervention on outcomes for teachers after 1 year of program implementation, 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 26. The models predict the average response at follow-up, adjusting for school size. Across all teacher measures, there were no significant impacts observed after the first year of implementation. 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 26. Secondary program impacts on teacher outcomes—Year one**

| 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] | 0.78 (0.19) | 0.60 (0.19) | 1.19 (0.83, 1.72) | 0.32 |
| Teacher safety concerns[3] | 1.44 (0.05) | 1.45 (0.05) | −0.01 (−0.15, 0.12) | 0.85 |
| School consistency of enforcing behavioral rules[4] | 2.13 (0.06) | 2.18 (0.06) | −0.05 (−0.22, 0.12) | 0.53 |
| Interactions with victims[4] | 3.34 (0.03) | 3.32 (0.03) | 0.02 (−0.05, 0.10) | 0.53 |
| Interactions with aggressors[4] | 3.54 (0.03) | 3.57 (0.03) | −0.03 (−0.08, 0.03) | 0.37 |
| Classroom management techniques[4] | 3.64 (0.03) | 3.65 (0.03) | −0.01 (−0.09, 0.07) | 0.79 |
| Sample size (Schools) | 20 | 20 | | |
| Sample size[5] (917 teachers nested within schools) | 465 | 452 | | |

[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 2007.

## 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 is being 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 and aggression?"

Table 27 contains results of analyses based on a repeated measures model of high-risk student violence and victimization. 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. The table includes the model-adjusted group-specific average rates of violent behaviors and victimization at baseline and first follow-up. Estimated program impacts reflect the net difference of the within-group change from pretest to first follow-up for treatment versus controls. Overall, 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) during the first year of implementation. These analyses indicate that the observed changes from baseline to first 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.

## 4.6 Exploratory Analysis—Impacts on the Gender Subgroup

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 effects of the RiPP and Best Behavior programs affect boys and girls differently.

The outcome measures at baseline for gender subgroups appear in appendix E. There were no differences at baseline between boys in treatment and control groups, or between girls in treatment and control groups. Table 28 reports on 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 for girls are reported as the within-gender ERRs comparing students receiving the RiPP and Best Behavior programs with students in control schools. None of these ERRs was statistically significant, indicating that there were no impacts on either boys or girls. The differences in impacts between boys and girls were obtained from the interaction of gender with treatment condition and estimated as the ratio of the two within-gender comparisons, with boys as the indicated group and girls as the reference group. Overall, there were no significant gender-by-treatment interaction effects after 1 year of the 3-year intervention, indicating no statistically significant differences between the impacts of the program on boys and on girls.

**Table 27. Main program impacts on self-reported violence and victimization—Year one: High-risk subgroup (Via repeated measures)**

| Student self-reported outcome[1] | Model-adjusted baseline event rates[2] (SE) | | Model-adjusted follow-up event rates[2] (SE) | | Estimated impact (95% CI)[3] | Wald Chi-Square *p*-value |
|---|---|---|---|---|---|---|
| | Intervention group | Control group | Intervention group | Control group | | |
| Violence (All items) | 5.03 (0.13) | 4.67 (0.12) | 5.16 (0.14) | 4.78 (0.13) | 1.00 (0.92, 1.10) | 0.95 |
| Without a weapon | 4.76 (0.12) | 4.40 (0.11) | 4.89 (0.13) | 4.51 (0.11) | 1.00 (0.92, 1.09) | 0.95 |
| With a weapon | 0.26 (0.03) | 0.26 (0.02) | 0.26 (0.03) | 0.26 (0.02) | 0.99 (0.97, 1.33) | 0.99 |
| Victimization (All items) | 5.81 (0.12) | 5.93 (0.12) | 6.28 (0.13) | 6.23 (0.12) | 1.03 (0.97, 1.10) | 0.34 |
| Overt | 3.52 (0.08) | 3.60 (0.08) | 3.86 (0.09) | 3.81 (0.09) | 1.03 (0.96, 1.11) | 0.35 |
| Relational | 2.28 (0.05) | 2.32 (0.04) | 2.42 (0.05) | 2.41 (0.05) | 1.02 (0.95, 1.09) | 0.60 |
| Sample size (Schools) | 20 | 20 | 20 | 20 | | |
| Sample size[4] (Students within 40 schools) | 1,005 | 1,148 | 897 | 1,016 | | |

[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 groups. Ratios of impact estimates of 1.00 indicate no interaction between time and program group (i.e., no program impact).

[4] Student sample sizes used in the analysis vary due to item nonresponse at baseline, follow-up, 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 survey, fall 2006 (baseline) and spring 2007 (first follow-up).

**Table 28. Main program impacts on self-reported violence and victimization—Year one: Gender subgroups**

| | Model-adjusted follow-up event rates (SE) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Boys | | Girls | | | | | |
| Student self-reported outcome[1] | Intervention group | Control group | Intervention group | Control group | Impact on boys[2] (95% CI) | Impact on girls[2] (95% CI) | Ratio of impacts[3] (95% CI) | Wald Chi-Square *p*-value |
| Violent and aggressive behaviors (All items) | 3.28 (0.15) | 3.34 (0.15) | 3.14 (0.14) | 3.02 (0.14) | 0.98 (0.89, 1.08) | 1.04 (0.95, 1.14) | 0.94 (0.85, 1.05) | 0.29 |
| Without a weapon | 3.08 (0.14) | 3.12 (0.14) | 3.01 (0.14) | 2.90 (0.13) | 0.99 (0.90, 1.08) | 1.04 (0.94, 1.14) | 0.95 (0.86, 1.06) | 0.35 |
| With a weapon | 0.18 (0.02) | 0.21 (0.02) | 0.11 (0.01) | 0.10 (0.01) | 0.86 (0.69, 1.07) | 1.13 (0.84, 1.51) | 0.76 (0.52, 1.13) | 0.16 |
| Victimization (All items) | 5.41 (0.16) | 5.51 (0.16) | 4.97 (0.15) | 4.87 (0.14) | 0.98 (0.92, 1.05) | 1.02 (0.95, 1.09) | 0.96 (0.90, 1.04) | 0.29 |
| Overt | 3.38 (0.11) | 3.45 (0.11) | 2.79 (0.10) | 2.71 (0.09) | 0.98 (0.91, 1.06) | 1.03 (0.95, 1.12) | 0.95 (0.88, 1.04) | 0.23 |
| Relational | 2.02 (0.06) | 2.06 (0.06) | 2.15 (0.06) | 2.15 (0.06) | 0.98 (0.92, 1.04) | 1.00 (0.94, 1.06) | 0.97 (0.91, 1.05) | 0.53 |
| Sample size (Schools) | 20 | 20 | 20 | 20 | | | | |
| Sample size[4] (Students nested within schools) | 1,772 | 1,818 | 1,829 | 1,902 | | | | |

[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] Student sample sizes used in the analysis vary due to item nonresponse at follow-up, covariate nonresponse, or both. Missing data ranged from 3 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 2007 (first follow-up).

## 4.7 Conclusions

The purpose of this study is 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. This report presents the background of the study and findings after 1 year 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 1 year 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 in the first year. This has the potential to limit the ability to find statistically significant differences between treatment and control schools.

The key impact findings are summarized as follows:

- 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).
- After 1 year of exposure to the RiPP and Best Behavior intervention, student measures for other outcomes—including safety concerns, prosocial behaviors, perception of behavioral expectations, attitudes toward violence, and self-reported strategies for coping with aggression—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.
- Subgroup analyses indicated that the main impacts for student violence and victimization were similar for boys and girls.
- After 1 year of program implementation, there were no statistically significant impacts of the RiPP and Best Behavior programs among the subpopulation of high-risk youth, as measured by student violence and victimization.

This report is based on data from the first of the intervention's 3 years; as such, the findings reported here do not represent conclusive information about the impact of the two programs examined in this study. As designed, the intervention was to be implemented over an additional 2 years, during which time additional RiPP lessons were delivered to students and school management teams worked toward completing additional milestones for Best Behavior. During the second and third years, RiPP-7 and RiPP-8 were introduced to 7th- and 8th-graders, respectively, and at least one group of students (those who began the study as 6th-graders) received the complete

sequence of RiPP by the end of the third year. Regarding Best Behavior, the study's goal in the second and third years was for school staff and management teams to address the challenges experienced in year one and continue progress toward the goals they set at the beginning of the study. A second report, combining results for years two and three, is forthcoming and will provide impact estimates of the intervention after 2 years and 3 years.

# References


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

Bauer, N., Lozano, P., and Rivara, F. (2007). The Effectiveness of the Olweus Bullying Prevention Program in Public Middle Schools: A Controlled Trial. *Journal of Adolescent Health, 40*: 266–274.

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.

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

Collins, L.M., & 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.

Dinkes, R., Cataldi, E.F., and Lin-Kelly, W. (2007). *Indicators of School Crime and Safety: 2007* (NCES 2008-021/NCJ 219553). National Center for Education Statistics and Bureau of Justice Statistics, Institute of Education Sciences, U.S. Departments of Education and Justice. Washington, DC.

Dinkes, R., Kemp, J., and Baum, K. (2009). *Indicators of School Crime and Safety: 2008* (NCES 2009–022/NCJ 226343). 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.

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.

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://www.ed.gov/rschstat/eval/other/progressinprevention.pdf.

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.

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.

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.

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.

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.

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.

Nolle, K.L., Guerino, P., and Dinkes, R. (2007). *Crime, Violence, Discipline, and Safety in U.S. Public Schools: Findings From the School Survey on Crime and Safety: 2005–06* (NCES 2007-361). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

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

Olweus, D. (1983). *Olweus Bullying Prevention Program*. Center City, MN: Hazelden Publishing. Retrieved December 7, 2009, from http://www.olweus.org.

Orpinas, P. (2008). *Measurement Manual: Description of Aggression, Victimization, and Social Skills Scales.* Retrieved November 21, 2008, from http://www.publichealth.uga.edu/hpb/research/hpb_youth_violence_orpinas.html.

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.

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

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.

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.

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.

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

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

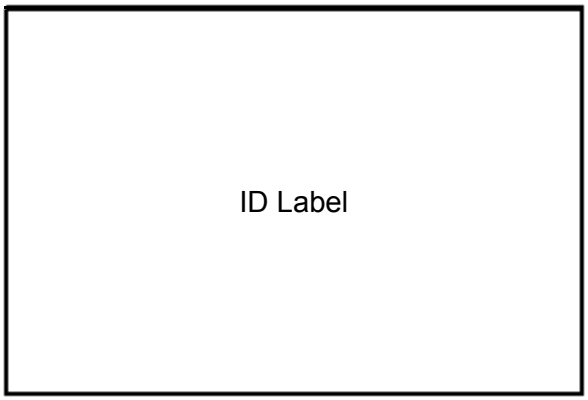
U.S. Department of Health and Human Services. (2001). *Youth Violence: A Report of the Surgeon General.* Rockville, MD: Author, Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Substance Abuse and Mental Health Services Administration, Center for Mental Health 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.

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

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

5. Which of these people do you live with most 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

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

## 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 | ○ | ○ | ○ | ○ |
| 17. Say "thanks" or "you're welcome" to you | ○ | ○ | ○ | ○ |
| 18. Say or do something that made you feel good | ○ | ○ | ○ | ○ |
| 19. Invite you to participate in a game, group conversation, or a class activity | ○ | ○ | ○ | ○ |
| 20. Say a compliment (praise, kind word) to you | ○ | ○ | ○ | ○ |
| 21. Offer to help you | ○ | ○ | ○ | ○ |
| 22. Share something with you | ○ | ○ | ○ | ○ |
| 23. Act friendly with you | ○ | ○ | ○ | ○ |
| 24. Show interest in your ideas or activities | ○ | ○ | ○ | ○ |

| 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 | ○ | ○ | ○ | ○ |
| 26. Actually hurt you by hitting, pushing, slapping, or shoving you | ○ | ○ | ○ | ○ |
| 27. Threaten you (but not actually injure you) with a weapon such as a gun, knife, or club | ○ | ○ | ○ | ○ |
| 28. Actually injure you with a weapon such as a gun, knife, or club | ○ | ○ | ○ | ○ |
| 29. Yell at you when they were angry | ○ | ○ | ○ | ○ |
| 30. Throw something at you to hurt you | ○ | ○ | ○ | ○ |
| 31. Pick a fight with you | ○ | ○ | ○ | ○ |
| 32. Take, damage, or destroy on purpose something that belonged to you | ○ | ○ | ○ | ○ |
| 33. Try to force you to do something that you didn't want to do | ○ | ○ | ○ | ○ |
| 34. Leave you out from a group or activity on purpose | ○ | ○ | ○ | ○ |
| 35. Tell lies, spread rumors, or say mean things about you to other students | ○ | ○ | ○ | ○ |
| 36. Call you an insulting name or word | ○ | ○ | ○ | ○ |
| 37. Make fun of you in front of other people just to be mean | ○ | ○ | ○ | ○ |

**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 | ○ | ○ | ○ | ○ |
| 39. Say "thanks" or "you're welcome" to a kid from your school | ○ | ○ | ○ | ○ |
| 40. Say or do something that made a kid from your school feel good | ○ | ○ | ○ | ○ |
| 41. Invite a kid from your school to participate in a game, group conversation, or activity | ○ | ○ | ○ | ○ |
| 42. Compliment (praise, say a kind word) a kid from your school | ○ | ○ | ○ | ○ |
| 43. Offer to help a kid from your school | ○ | ○ | ○ | ○ |
| 44. Share something with a kid from your school | ○ | ○ | ○ | ○ |
| 45. Act friendly with a kid from your school | ○ | ○ | ○ | ○ |
| 46. Show interest in the ideas or activities of a kid from your school | ○ | ○ | ○ | ○ |

| 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 | ○ | ○ | ○ | ○ | ○ |
| 48. Stand up for someone who was being "bullied" | ○ | ○ | ○ | ○ | ○ |

| 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 | ○ | ○ | ○ | ○ |

# 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 | ○ | ○ | ○ | ○ |
| 2. Any hallways or stairs in the school | ○ | ○ | ○ | ○ |
| 3. Any part of the school cafeteria | ○ | ○ | ○ | ○ |
| 4. Any school restroom | ○ | ○ | ○ | ○ |
| 5. In any classroom | ○ | ○ | ○ | ○ |
| 6. School parking lot, athletic fields, or other places outside school buildings | ○ | ○ | ○ | ○ |

**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 | ○ | ○ | ○ | ○ |
| 8. A student actually hit, slap, shove, or push another student to hurt him or her | ○ | ○ | ○ | ○ |
| 9. A student threaten (but not actually injure) another student with a weapon such as a knife, gun, or club | ○ | ○ | ○ | ○ |
| 10. A student actually injure another student with a weapon such as a knife, gun, or club | ○ | ○ | ○ | ○ |
| 11. A student get angry and yell at another student | ○ | ○ | ○ | ○ |
| 12. A student throw something at another student to hurt him or her | ○ | ○ | ○ | ○ |
| 13. A student pick a fight with another student | ○ | ○ | ○ | ○ |
| 14. A student take, damage, or destroy on purpose something that belonged to another student | ○ | ○ | ○ | ○ |
| 15. A student try to force another student to do something he or she didn't want to do | ○ | ○ | ○ | ○ |
| 16. A student leave out another student on purpose from a group or activity | ○ | ○ | ○ | ○ |

| | Never | Almost never | Sometimes | Often |
|---|---|---|---|---|
| 17. A student tell lies, spread rumors, or say mean things about someone to other students | ○ | ○ | ○ | ○ |
| 18. A student call another student an insulting name or word | ○ | ○ | ○ | ○ |
| 19. A student make fun of another student in front of him or her just to be mean | ○ | ○ | ○ | ○ |
| 20. A student sexually harass another student | ○ | ○ | ○ | ○ |
| 21. A student disrupt my class due to misbehavior | ○ | ○ | ○ | ○ |

**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 | ○ | ○ | ○ | ○ |
| 23. Threatened with physical harm by a student from your school | ○ | ○ | ○ | ○ |
| 24. Physically attacked or injured by a student from your school | ○ | ○ | ○ | ○ |

**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 | ○ | ○ | ○ | ○ |
| 26. Say “thanks” or “you’re welcome” to another student | ○ | ○ | ○ | ○ |
| 27. Say or do something that made another student feel good | ○ | ○ | ○ | ○ |
| 28. Invite another student to participate in a game, group conversation, or a class activity | ○ | ○ | ○ | ○ |
| 29. Say a compliment (praise, kind word) to another student | ○ | ○ | ○ | ○ |
| 30. Offer to help another student | ○ | ○ | ○ | ○ |

| | | | | |
|---|---|---|---|---|
| 31. Share something with another student | ○ | ○ | ○ | ○ |
| 32. Act friendly with another student | ○ | ○ | ○ | ○ |

| | Never □ | Almost never □ | Sometimes □ | Often □ |
|---|---|---|---|---|
| 33. Show positive interest in another student's ideas or activities | ○ | ○ | ○ | ○ |
| 34. Stop someone from getting in a fight | ○ | ○ | ○ | ○ |
| 35. Stand up for someone who was being "bullied" | ○ | ○ | ○ | ○ |
| 36. Voluntarily apologize to another student | ○ | ○ | ○ | ○ |

**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 | ○ | ○ | ○ | ○ | ○ |
| 38. Create an "open door" policy for students who are the target of aggression | ○ | ○ | ○ | ○ | ○ |
| 39. Include victimized or isolated children in group projects | ○ | ○ | ○ | ○ | ○ |
| 40. Use classroom routines that reduce the opportunity for acting out behaviors | ○ | ○ | ○ | ○ | ○ |
| 41. Reward small improvements toward desired behavior | ○ | ○ | ○ | ○ | ○ |
| 42. Use a behavior plan that provides students choices and consequences for their choices | ○ | ○ | ○ | ○ | ○ |
| 43. Provide opportunities for students to confidentially report aggressive acts | ○ | ○ | ○ | ○ | ○ |
| 44. Assist students victimized by aggressive peers in identifying skills and behaviors they can use in these situations | ○ | ○ | ○ | ○ | ○ |
| 45. Address aggressive situations in the classroom immediately | ○ | ○ | ○ | ○ | ○ |

| 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 | ○ | ○ | ○ | ○ | ○ |
| 47. Confront students who make inappropriate comments | ○ | ○ | ○ | ○ | ○ |
| 48. Consult with school administrators for support | ○ | ○ | ○ | ○ | ○ |
| 49. Use self-calming techniques during the school day | ○ | ○ | ○ | ○ | ○ |
| 50. Model dignity and respect at school | ○ | ○ | ○ | ○ | ○ |
| 51. Provide positive reinforcement for prosocial behavior | ○ | ○ | ○ | ○ | ○ |

**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 | ○ | ○ | ○ | ○ |
| 53. The school makes sure that students know the rules for student behavior | ○ | ○ | ○ | ○ |
| 54. The school rules emphasize reinforcing desired behavior | ○ | ○ | ○ | ○ |
| 55. The school rules emphasize consequences for undesired behavior | ○ | ○ | ○ | ○ |
| 56. When a school rule is broken, it is clear to the school staff what consequences should follow | ○ | ○ | ○ | ○ |
| 57. Teachers at my school consistently enforce the rules | ○ | ○ | ○ | ○ |
| 58. Teachers at my school punish students the same way for breaking the same rule no matter who the students are | ○ | ○ | ○ | ○ |
| 59. Teachers compliment or reward students when they follow the rules | ○ | ○ | ○ | ○ |
| 60. Teachers at my school reward students the same way for following the same rule no matter who the students are | ○ | ○ | ○ | ○ |

| | 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. This was based on 8 items from the full set of 14 perpetration items. 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 items**

| Number of severe 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).

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:

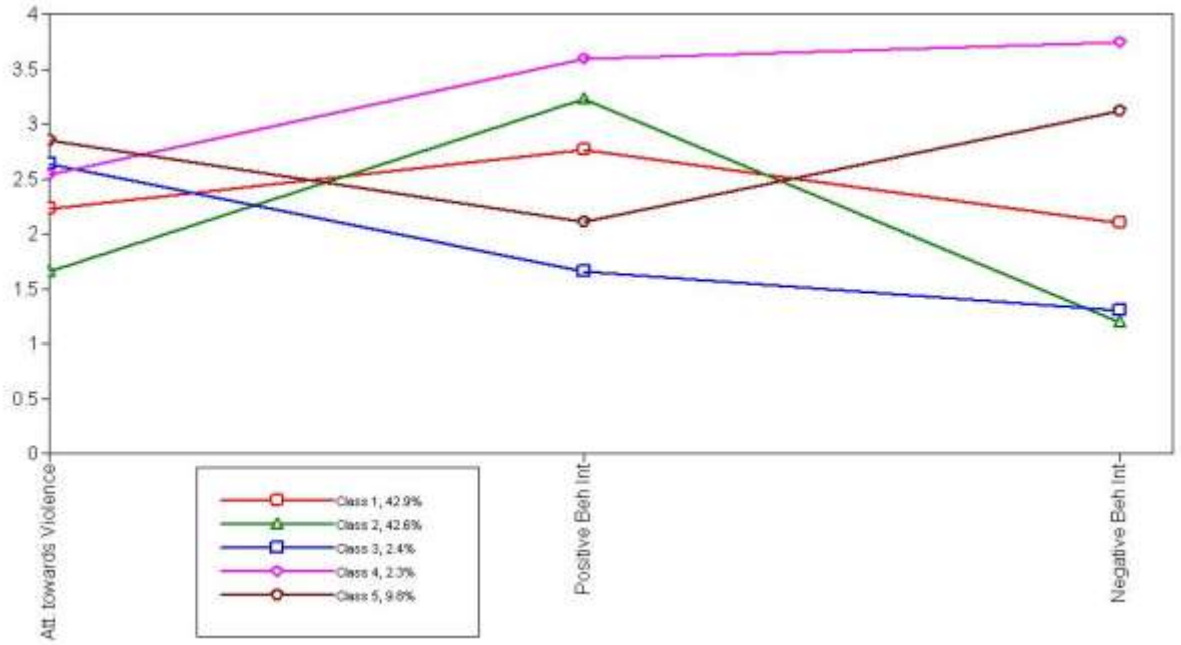
- attitudes endorsing violence (items Q65–Q72);
- reactions to anger—positive (items Q73–Q77, Q81–Q84); and
- reactions to anger—negative (items Q78–Q80).

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 students in the high-risk group.

Combining criteria (high-risk class or high-risk perpetration) yields the breakdown in table B-3, with a final risk assignment frequency of 5,212 (68.55 percent) classified as not high risk and 2,391 (31.45 percent) as high-risk students.

**Table B-3. Risk classification using joint criteria**

| Risk type | *N* | Percent |
|---|---|---|
| None | 5,212 | 68.55 |
| Risk, perpetrator | 1,669 | 21.95 |
| Risk, class | 256 | 3.37 |
| Risk, both | 466 | 6.13 |

SOURCE: Student survey, fall 2006 (baseline).

Table B-4 displays the agreement in risk classification by each of the two criteria used to assign risk. Only about 20 percent of those classified as high risk were positive for risk, according to both criteria (466 out of 2,391).

**Table B-4. Cross-tabulation of risk classification by class or perpetration**

| | Perpetration | | |
|---|---|---|---|
| Class | High risk | Not high risk | Total |
| High risk | 466 (6.13%) | 256 (3.37%) | 722 (9.50%) |
| Not high risk | 1,669 (21.95%) | 5,212 (38.55%) | 6,881 (90.50%) |
| Total | 2,135 (28.08%) | 5,468 (71.92%) | |

SOURCE: Student survey, fall 2006 (baseline).

Students that met either criterion were classified as high risk. However, not all high-risk students were followed for the longitudinal high-risk sample. A maximum of 60 high-risk students per school were followed for this sample. This sample size limit led to the random selection of 60 students in nine of the larger schools, in which more than 60 students were classified as high risk at baseline.

Table B-5 provides self-reported demographic data for the high-risk subgroup at baseline. Over 30 percent of students were Black, 27 percent were Hispanic, and 24 percent were White. More than one-half of the students (58 percent) were male, and just over 54 percent lived in single-adult households.

**Table B-5. Baseline demographic characteristics of the high-risk subgroup**

| Characteristic | |
|---|---|
| Sample size, high-risk subgroup | 2,391 |
| Race/ethnicity (%) | |
| Hispanic | 27.49 |
| Black, non-Hispanic | 30.77 |
| White, non-Hispanic | 24.04 |
| Other or mixed[1] | 16.70 |
| Gender (%) | |
| Male | 57.56 |
| Single-adult household (%) | 54.41 |

[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 |
|---|---|---|---|---|
| Aggression: Overall | Past 30 days: Threats or actual aggression, at school | Student survey | Adapted from the Problem Behavior Frequency Scales (Farrell et al. 2000) | 14 |
| Aggression: Weapons-related | Past 30 days: Threats or actual aggression, 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 |
| Aggression: Not weapons-related | Past 30 days: Threats or actual aggression, 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 |
| 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 |

See notes at end of table.

**Table C-1. Sources of outcome measures for the student and teacher surveys—Continued**

| Outcome measure | Description | Instrument type | Source | # items in scale |
|---|---|---|---|---|
| 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 2008) | 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 2008) | 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 |
| 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 |

See notes at end of table.

**Table C-1. Sources of outcome measures for the student and teacher surveys—Continued**

| Outcome measure | Description | Instrument type | Source | # items in scale |
|---|---|---|---|---|
| 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 aggression. 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 and aggression (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 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 aggression and 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.

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.

**Table C-2. Student survey: Outcomes addressing key research questions**

| Outcome measure | Items | Variable type | # items in scale | Cronbach α[1] |
|---|---|---|---|---|
| Student violence: Overall | Past 30 days: Threats or actual aggression, at school | Count | 14 | 0.84 |
| Student aggression: Weapons-related | In the PAST 30 DAYS, how often did YOU: | Count | 3 | 0.72 |
| | -Bring a weapon to school | | | |
| | -Threaten with a weapon (knife, club, or gun) | | | |
| | -Actually injure with a weapon (knife, club, or gun) | | | |
| Student aggression: Not weapons-related | In the PAST 30 DAYS, how often did YOU: | Count | 11 | 0.90 |
| | -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 | | | |
| Student victimization: Overall | Past 30 days: Threats or actual victimization, at school | Count | 13 | 0.89 |
| Student victimization: Overt | In the PAST 30 DAYS, how often did OTHER STUDENTS AT YOUR SCHOOL: | Count | 9 | 0.84 |
| | -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 | | | |
| Student victimization: Relational | In the past 30 days, how often did OTHER STUDENTS AT YOUR SCHOOL: | Count | 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 | # items in scale | Cronbach α[1] |
|---|---|---|---|---|
| Student safety concerns | The next questions ask about how safe you feel at school:<br>-How often do you worry that someone from your school will attack or hurt you<br>-How often do you worry that someone from your school will bully you | Continuous | 2 | — |
| Prosocial behaviors: Extended to others | In the PAST 30 DAYS, how often did YOU:<br>-Say/do something nice to a student at school<br>-Say "thanks" or "you're welcome" to a student at school<br>-Say/do something that made a student feel good<br>-Invite a student to participate in a game, conversation, activity<br>-Compliment another student<br>-Offer to help a student<br>-Share something<br>-Act friendly with another student<br>-Show interest in the ideas/activities of another student | Continuous | 9 | 0.90 |
| Prosocial behaviors: Received from others | In the PAST 30 DAYS, how often did OTHER STUDENTS AT YOUR SCHOOL:<br>-Say or do something nice to you<br>-Say "thanks" or "you're welcome" to you<br>-Say or do something that made you feel good<br>-Invite you to participate in a game, group conversation, or class activity<br>-Say a compliment (praise, kind word) to you<br>-Offer to help you<br>-Act friendly with you<br>-Show interest in your ideas or activities | Continuous | 9 | 0.86 |
| Perceived behavioral expectations | The following statements could describe a school. Think about your school over the PAST 30 DAYS:<br>-Everyone knows the rules<br>-The school rules are strictly enforced<br>-If a rule is broken, students know the consequences<br>-Punishments are the same regardless of who you are<br>-Students are complimented or rewarded for following rules | Continuous | 10 | 0.74 |

See notes at end of table.

**Table C-3. Student survey: Other outcome measures—Continued**

| Outcome measure | Items | Variable type | # items in scale | Cronbach α[1] |
|---|---|---|---|---|
| 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: | Continuous | 9 | 0.84 |
| | -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 | | | |
| 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: | Count | 3 | 0.77 |
| | -Yell at the person | | | |
| | -Break something | | | |
| | -Hit or threaten to hurt the person | | | |
| Student attitudes toward violence[2] | The next few questions ask about your feelings and how you get along in general: | Continuous | 6 | 0.78 |
| | -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 | | | |

[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 | # items in scale | Cronbach α[1] |
|---|---|---|---|---|
| Teacher victimization | How often in the PAST 6 MONTHS have YOU been:<br>-Verbally abused by a student at school<br>-Threatened with physical harm by a student<br>-Physically attacked or injured by a student | Count | 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:<br>-The entrance to the school<br>-Any hallway or stairs<br>-Any part of the school cafeteria<br>-Any school restroom<br>-Any classroom<br>-The school parking lot, fields, or other places outside the school | Continuous | 6 | 0.89 |
| Teacher interactions with victims | In the PAST 30 DAYS, how often did you:<br>-Model strategies for solving conflicts<br>-Create an open-door policy for students who are targets of aggression<br>-Include victimized or isolated children in group activities<br>-Assist victimized students in identifying skills and behaviors they can use in these situations | Continuous | 4 | 0.75 |
| Teacher interactions with aggressors | In the PAST 30 DAYS, how often did you:<br>-Address aggressive situations in the classroom immediately<br>-Maintain calmness when faced with an aggressive or disruptive student<br>-Confront students who make inappropriate comments<br>-Consult with school administrators for support | Continuous | 4 | 0.71 |
| Classroom management techniques | In the PAST 30 DAYS, how often did you:<br>-Use classroom routines that reduce the opportunity for acting out behaviors<br>-Reward small improvements toward desired behavior<br>-Use a behavior plan that provides students with choices and consequences for their choices<br>-Model dignity and respect at school<br>-Provide positive reinforcement for prosocial behavior | Continuous | 5 | 0.68 |

See notes at end of table.

**Table C-4. Teacher survey: Other outcome measures—Continued**

| Outcome measure | Items | Variable type | # items in scale | Cronbach α[1] |
|---|---|---|---|---|
| School consistency of enforcing behavioral rules | Statements that could describe your school:<br>-The school rules for student behavior are clearly defined<br>-The school makes sure that students know the rules for student behavior<br>-The school rules emphasize reinforcing desired behavior<br>-The school rules emphasize consequences for undesired behavior<br>-When a school rule is broken, it is clear to the school staff what consequences should follow<br>-Teachers at my school consistently enforce the rules<br>-Teachers at my school punish students the same way for breaking the same rule, no matter who the students are<br>-Teachers compliment or reward students when they follow the rules<br>-Teachers at my school reward students the same way for following the same rule, no matter who the students are<br>-Teachers or other adults at the school try to prevent or stop bullying<br>-Teachers treat students with respect<br>-Teachers know the procedure for reporting bullying and violence<br>-Administrators and/or teachers periodically review school rules to determine if they need modification<br>-Administrators at my school are supportive of teachers in creating a safe school<br>-Teachers receive adequate training in classroom management/discipline strategies<br>-Administrators at my school consistently enforce the rules | Continuous | 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.

Using these binary and polytomous 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 aggression 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.

**Table C-5. Correspondence of measurement models for primary outcomes**

| | | Polytomous CFA | | | Binary CFA | | |
|---|---|---|---|---|---|---|---|
| | Correlation of CFA factor scores | 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).

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-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: 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 aggression, drug use, peer provocation, nonphysical aggression 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 treatment or control conditions (i.e., 20 schools in each condition) and surveying 243 students at each school provided sufficient power for testing the null hypothesis regarding treatment effectiveness on schools (table D-1).[16]

[16] Estimations of statistical precision are based on assumptions appropriate for a normally distributed variable and are considered approximate for nonnormally distributed variables.

**Table D-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 $\alpha$ = 0.05 and a two-tailed test. Calculations are based on intraclass correlation coefficient (ICC) = 0.045.
SOURCE: RTI calculations.

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.

Table D-2 provides a power table for the nested cohort analysis that will be used to assess the subsample of high-risk youth.[17] Using assumptions noted in section 2.6 regarding attrition rates across the 3 years of the program, the study team anticipates 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 will 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 D-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 $\alpha$ = 0.05 and a two-tailed test. Calculations are based on intraclass correlation coefficient (ICC) = 0.019.
SOURCE: RTI calculations.

[17] Estimations of statistical precision are based on assumptions appropriate for a normally distributed variable and are considered approximate for nonnormally distributed variables.

In table D-2, statistical power could be improved even further by including covariates (such as baseline data) in the models. In summary, tables D-1 and D-2 demonstrate that a total of 40 schools are required to detect the target minimal detectable effects.

# Appendix E: Baseline Measures for Gender Subgroups

We conducted analyses to assess whether or not there were any differences between treatment and control groups on the student measures at baseline, for each of the two gender subgroups. Tables E-1 and E-2 present the outcome measures at baseline for boys and girls, respectively. In both cases, there were no significant differences between intervention and control groups on any of the baseline measures.

**Table E-1. Baseline measures for the student sample: 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) | 2.12 (0.14) | 2.12 (0.14) | −0.01 | −0.04 | 0.96 |
| Violence: Weapons-related | 0.10 (0.01) | 0.11 (0.01) | −0.01 | −0.42 | 0.67 |
| Violence: Not weapons-related | 2.01 (0.13) | 2.01 (0.13) | 0.00 | 0.02 | 0.98 |
| Victimization (All items) | 4.12 (0.14) | 4.49 (0.16) | −0.37 | −1.73 | 0.09 |
| Victimization: Overt | 2.47 (0.10) | 2.69 (0.10) | −0.22 | −1.56 | 0.13 |
| Victimization: Relational | 1.65 (0.06) | 1.80 (0.06) | −0.15 | −1.74 | 0.09 |
| Other outcomes[3] | | | | | |
| Self-reported coping strategies: Positive | 2.77 (0.02) | 2.82 (0.02) | −0.05 | −1.57 | 0.12 |
| Self-reported coping strategies: Negative[2] | 0.66 (0.04) | 0.71 (0.04) | −0.05 | −0.94 | 0.35 |
| Attitudes toward violence | 2.89 (0.03) | 2.89 (0.03) | −0.00 | −0.07 | 0.94 |
| Prosocial behaviors: Extended to others | 2.83 (0.03) | 2.84 (0.03) | −0.01 | −0.19 | 0.85 |
| Prosocial behaviors: Received from others | 2.68 (0.03) | 2.68 (0.03) | 0.00 | 0.02 | 0.98 |
| Behavioral expectations | 3.12 (0.02) | 3.12 (0.02) | −0.00 | −0.00 | 1.00 |
| Safety concerns | 1.81 (0.03) | 1.87 (0.03) | −0.06 | −1.68 | 0.10 |
| Sample size (Schools) | 20 | 20 | | | |
| Sample size[4] (Students) | 1,821 | 1,905 | | | |

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to 0.05.

[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 E-2. Baseline measures for the student sample: 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) | 1.82 (0.15) | 1.65 (0.14) | 0.18 | 0.88 | 0.38 |
| Violence: Weapons-related | 0.06 (0.01) | 0.05 (0.01) | 0.01 | 0.96 | 0.34 |
| Violence: Not weapons-related | 1.77 (0.14) | 1.60 (0.13) | 0.17 | 0.87 | 0.39 |
| Victimization (All items) | 3.62 (0.17) | 3.68 (0.17) | −0.06 | −0.27 | 0.79 |
| Victimization: Overt | 1.90 (0.10) | 1.92 (0.10) | −0.02 | −0.14 | 0.89 |
| Victimization: Relational | 1.72 (0.07) | 1.77 (0.07) | −0.05 | −0.53 | 0.60 |
| Other outcomes[3] | | | | | |
| Self-reported coping strategies: Positive | 2.97 (0.02) | 2.99 (0.02) | −0.02 | −0.61 | 0.55 |
| Self-reported coping strategies: Negative[2] | 0.53 (0.04) | 0.47 (0.04) | 0.06 | 1.05 | 0.30 |
| Attitudes toward violence | 3.08 (0.04) | 3.13 (0.04) | −0.05 | −0.84 | 0.41 |
| Prosocial behaviors: Extended to others | 3.02 (0.03) | 3.06 (0.03) | −0.05 | −1.09 | 0.28 |
| Prosocial behaviors: Received from others | 2.85 (0.03) | 2.87 (0.03) | −0.03 | −0.63 | 0.53 |
| Behavioral expectations | 3.20 (0.02) | 3.19 (0.02) | 0.01 | 0.30 | 0.76 |
| Safety concerns | 1.94 (0.04) | 2.04 (0.04) | −0.10 | −1.92 | 0.06 |
| Sample size (Schools) | 20 | 20 | | | |
| Sample size[4] (Students) | 1,869 | 1,979 | | | |

[1] Statistical significance is indicated by * if the *p*-value is less than or equal to 0.05.

[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 F: Unadjusted Means and Standard Deviations for Impact Variables

**Table F-1. Unadjusted means and standard deviations for student outcomes—Baseline**

| | Event rates or scale means (standard deviations) | |
|---|---|---|
| Measure | Intervention group | Control group |
| Main outcomes[1] | | |
| Violence (All items) | 2.03 (2.88) | 1.95 (2.72) |
| Violence: Weapons-related | 0.09 (0.40) | 0.08 (0.40) |
| Violence: Not weapons-related | 1.94 (2.68) | 1.86 (2.52) |
| Victimization (All items) | 3.88 (3.41) | 4.11 (3.42) |
| Victimization: Overt | 2.20 (2.24) | 2.31 (2.26) |
| Victimization: Relational | 1.68 (1.48) | 1.80 (1.49) |
| Other outcomes[2] | | |
| Safety concerns[3] | 1.89 (0.92) | 1.96 (0.94) |
| Prosocial behaviors extended[4] | 2.92 (0.69) | 2.95 (0.68) |
| Prosocial behaviors received[4] | 2.76 (0.66) | 2.78 (0.65) |
| Perceived behavioral expectations[4] | 3.16 (0.44) | 3.16 (0.42) |
| Attitudes toward violence[4] | 2.99 (0.68) | 3.01 (0.66) |
| Self-reported coping strategies (Positive) [1,3] | 2.87 (0.65) | 2.90 (0.63) |
| Self-reported coping strategies (Negative) [4] | 1.84 (0.75) | 1.82 (0.75) |
| Sample size (Schools) | 20 | 20 |
| Sample size (Students) | 3,705 | 3,896 |

[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 F-2. Unadjusted means and standard deviations for high-risk student outcomes—Baseline**

| | Event rates or scale means (standard deviations) | |
|---|---|---|
| Measure | Intervention group | Control group |
| Main outcomes[1] | | |
| Violence (All items) | 5.08 (3.26) | 4.40 (2.81) |
| Violence: Weapons-related | 0.29 (0.68) | 0.27 (0.68) |
| Violence: Not weapons-related | 4.79 (2.95) | 4.40 (2.81) |
| Victimization (All items) | 5.81 (3.38) | 6.00 (3.31) |
| Victimization: Overt | 3.53 (2.32) | 3.65 (2.30) |
| Victimization: Relational | 2.29 (1.41) | 2.36 (1.38) |
| Other outcomes[2] | | |
| Safety concerns[3] | 1.94 (0.92) | 2.06 (0.96) |
| Prosocial behaviors extended[4] | 2.70 (0.69) | 2.74 (0.67) |
| Prosocial behaviors received[4] | 2.68 (0.65) | 2.67 (0.65) |
| Perceived behavioral expectations[4] | 3.00 (0.47) | 3.02 (0.43) |
| Attitudes toward violence[4] | 2.53 (0.65) | 2.59 (0.65) |
| Self-reported coping strategies (Positive)[1,3] | 2.54 (0.68) | 2.63 (0.67) |
| Self-reported coping strategies (Negative)[4] | 2.36 (0.79) | 2.30 (0.78) |
| | 20 | 20 |
| Sample size: Students within 40 schools | 1,005 | 1,148 |

[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 F-3. Unadjusted means and standard deviations for student outcomes, by gender subgroup—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) | 2.18 ( 2.99) | 2.16 (2.89) | 1.88 (2.77) | 1.73 (2.52) |
| Violence: Weapons-related | 0.12 (0.47) | 0.12 (0.47) | 0.06 (0.33) | 0.05 (0.30) |
| Violence: Not weapons-related | 2.07 (2.74) | 2.16 (2.89) | 1.81 (2.62) | 1.68 (2.40) |
| Victimization (All items) | 4.14 (3.56) | 4.48 (3.60) | 3.62 (3.23) | 3.74 (3.20) |
| Victimization: Overt | 2.48 (2.37) | 2.68 (2.40) | 1.92 (2.06) | 1.95 (2.05) |
| Victimization: Relational | 1.65 (1.48) | 1.80 (1.50) | 1.71 (1.48) | 1.79 (1.48) |
| Other outcomes[2] | | | | |
| Safety concerns[3] | 1.82 (0.90) | 1.87 (0.92) | 1.95 (0.94) | 2.04 (0.95) |
| Prosocial behaviors extended[4] | 2.81 (0.71) | 2.84 (0.70) | 3.02 (0.66) | 3.06 (0.64) |
| Prosocial behaviors received[4] | 2.67 (0.67) | 2.68 (0.65) | 2.85 (0.65) | 2.87 (0.63) |
| Perceived behavioral expectations[4] | 3.12 (0.44) | 3.13 (0.43) | 3.20 (0.43) | 3.19 (0.41) |
| Attitudes toward violence[4] | 2.89 (0.68) | 2.90 (0.68) | 3.08 (0.67) | 3.13 (0.63) |
| Self-reported coping strategies (Positive)[1,3] | 2.76 (0.67) | 2.81 (0.69) | 2.98 (0.61) | 2.99 (0.58) |
| Self-reported coping strategies (Negative)[4] | 1.93 (0.76) | 1.94 (0.77) | 1.75 (0.72) | 1.71 (0.70) |
| Sample size (Schools) | 20 | 20 | 20 | 20 |
| Sample size (Students nested within schools) | 1,821 | 1,905 | 1,869 | 1,979 |

[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 F-4. Unadjusted means and standard deviations for student outcomes—Year one follow-up**

| | Event rates or scale means (standard deviations) | |
|---|---|---|
| Measure | Intervention group | Control group |
| Main outcomes[1] | | |
| Violence (All items) | 3.02 (3.38) | 2.91 (3.23) |
| Violence: Weapons-related | 0.13 (0.49) | 0.12 (0.47) |
| Violence: Not weapons-related | 2.89 (3.15) | 2.78 (3.01) |
| Victimization (All items) | 4.90 (3.64) | 5.07 (3.62) |
| Victimization: Overt | 2.86 (2.46) | 2.94 (2.45) |
| Victimization: Relational | 2.04 (1.50) | 2.13 (1.50) |
| Other outcomes[2] | | |
| Safety concerns[3] | 1.92 (0.94) | 1.98 (0.96) |
| Prosocial behaviors extended[4] | 2.90 (0.71) | 2.94 (0.70) |
| Prosocial behaviors received[4] | 2.79 (0.70) | 2.79 (0.69) |
| Perceived behavioral expectations[4] | 3.02 (0.52) | 2.98 (0.53) |
| Attitudes toward violence[4] | 2.83 (0.72) | 2.87 (0.72) |
| Self-reported coping strategies (Positive)[1,3] | 2.80 (0.71) | 2.80 (0.66) |
| Self-reported coping strategies (Negative)[4] | 2.06 (0.81) | 2.01 (0.79) |
| Sample size (Schools) | 20 | 20 |
| Sample size (Students) | 3,619 | 3,732 |

[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, first follow-up (spring 2007).

**Table F-5. Unadjusted means and standard deviations for high-risk student outcomes—Year one follow-up**

| | Event rates or scale means (standard deviations) | |
|---|---|---|
| Measure | Intervention group | Control group |
| Main outcomes[1] | | |
| Violence (All items) | 5.17 (3.82) | 4.75 (3.57) |
| Violence: Weapons-related | 0.27 (0.69) | 0.26 (0.68) |
| Violence: Not weapons-related | 4.90 (3.50) | 4.49 (3.27) |
| Victimization (All items) | 6.21 (3.64) | 6.26 (3.61) |
| Victimization: Overt | 3.82 (2.52) | 3.84 (2.51) |
| Victimization: Relational | 2.39 (1.43) | 2.43 (1.42) |
| Other outcomes[2] | | |
| Safety concerns[3] | 1.95 (0.97) | 2.02 (0.98) |
| Prosocial behaviors extended[4] | 2.71 (0.71) | 2.75 (0.70) |
| Prosocial behaviors received[4] | 2.74 (0.69) | 2.69 (0.69) |
| Perceived behavioral expectations[4] | 2.85 (0.53) | 2.82 (0.54) |
| Attitudes toward violence[4] | 2.45 (0.70) | 2.54 (0.70) |
| Self-reported coping strategies (Positive)[1,3] | 2.53 (0.73) | 2.59 (0.68) |
| Self-reported coping strategies (Negative)[4] | 2.43 (0.81) | 2.36 (0.80) |
| | 20 | 20 |
| Sample size: Students within 40 schools | 897 | 1,016 |

[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 2007 (first follow-up).

**Table F-6. Unadjusted means and standard deviations for student outcomes, by gender subgroup—Year one follow-up**

| | Event rates or scale means (standard deviations) | | | |
|---|---|---|---|---|
| | Boys | | Girls | |
| Measure | Intervention group | Control group | Intervention group | Control group |
| Main outcomes[1] | | | | |
| Violence (All items) | 3.09 (3.51) | 3.07 (3.40) | 2.94 (3.25) | 2.75 (3.05) |
| Violence: Weapons-related | 0.17 (0.57) | 0.17 (0.56) | 0.09 (0.40) | 0.08 (0.37) |
| Violence: Not weapons-related | 2.92 (3.21) | 2.90 (3.11) | 2.85 (3.09) | 2.68 (2.92) |
| Victimization (All items) | 5.11 (3.80) | 5.39 (3.78) | 4.68 (3.46) | 4.77 (3.43) |
| Victimization: Overt | 3.14 (2.57) | 3.30 (2.58) | 2.59 (2.31) | 2.59 (2.27) |
| Victimization: Relational | 1.97 (1.51) | 2.08 (1.51) | 2.10 (1.49) | 2.18 (1.49) |
| Other outcomes[2] | | | | |
| Safety concerns[3] | 1.87 (0.93) | 1.96 (0.98) | 1.97 (0.95) | 2.00 (0.93) |
| Prosocial behaviors extended[4] | 2.80 (0.72) | 2.81 (0.73) | 3.01 (0.68) | 3.07 (0.65) |
| Prosocial behaviors received[4] | 2.68 (0.68) | 2.66 (0.69) | 2.90 (0.69) | 2.92 (0.67) |
| Perceived behavioral expectations[4] | 2.98 (0.54) | 2.92 (0.53) | 3.06 (0.51) | 3.04 (0.51) |
| Attitudes toward violence[4] | 2.72 (0.72) | 2.74 (0.72) | 2.93 (0.71) | 2.99 (0.71) |
| Self-reported coping strategies (Positive)[1,3] | 2.67 (0.74) | 2.68 (0.68) | 2.93 (0.65) | 2.91 (0.62) |
| Self-reported coping strategies (Negative)[4] | 2.11 (0.81) | 2.11 (0.79) | 1.99 (0.80) | 1.91 (0.78) |
| Sample size (Schools) | 20 | 20 | 20 | 20 |
| Sample size (Students nested within schools) | 1,772 | 1,818 | 1,829 | 1,902 |

[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 2007 (first follow-up).

**Table F-7. Unadjusted means and standard deviations for teacher outcomes—Year one follow-up**

| | Event rates or scale means (standard deviations) | |
|---|---|---|
| Measure[1] | Intervention group | Control group |
| Teacher self-reported victimization[2,3] | 0.67 (0.47) | 0.63 (0.48) |
| Teacher safety concerns[3] | 1.43 (0.58) | 1.45 (0.62) |
| School consistency of enforcing behavioral rules[4] | 2.12 (0.58) | 2.18 (0.57) |
| Interactions with victims[4] | 3.34 (0.55) | 3.32 (0.54) |
| Interactions with aggressors[4] | 3.54 (0.44) | 3.57 (0.44) |
| Classroom management techniques[4] | 3.64 (0.41) | 3.65 (0.44) |
| Sample size (Schools) | 20 | 20 |
| Sample size[5] (917 teachers nested within schools) | 465 | 452 |

[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 2007.