

## ABSTRACT

Title of Document: THE EFFECTS OF SCHOOL-BASED  
SOCIAL SKILLS PROGRAMMING ON  
ACADEMIC INSTRUCTION TIME AND  
STUDENT ACHIEVEMENT

Joseph F. Nese, Doctor of Philosophy, 2009

Directed By: Professor Gary D. Gottfredson,  
Counseling and Personnel Services

The purposes of this study are to examine: (a) the effects of a well-implemented, school-based, universal social skills intervention on time-spent in formal social skills instruction and academic instruction time in the classroom; and (b) the effects of time-spent in formal social skills instruction on student achievement. Twelve elementary schools were matched and one school of each pair was randomly assigned to the treatment. The sample included 1,724 students in 113 third, fourth, and fifth grade classrooms with low rates of social-behavioral problems. Multilevel data analyses (HLM) methods were used to investigate the school-based treatment effect of social skills programming on academic instruction time as well as the classroom-level effect of social skills instruction on student achievement. Results indicated that treatment students received significantly more formal instruction in social skills, and that the frequency of formal social skills instruction had a very small, negative effect on students' report card grades and standardized test scores when lessons were 30

minutes or less in duration. The effects were not consistent and were so small as to have little theoretical or practical significance. Research and policy implications are discussed.

THE EFFECTS OF SCHOOL-BASED SOCIAL SKILLS PROGRAMMING  
ON ACADEMIC INSTRUCTION TIME AND STUDENT ACHIEVEMENT

By

Joseph F. Nese

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2009

Advisory Committee:  
Professor Gary Gottfredson, Chair  
Associate Professor William Strein  
Professor Cheryl Holcomb-McCoy  
Assistant Professor Brian Johnson  
Associate Professor Robert Marcus

© Copyright by  
Joseph F. Nese  
2009

## Table of Contents

Table of Contents .....	ii
List of Tables .....	iii
Literature Review.....	1
Universal Social Skills Programming.....	2
Effects of Social Skills on Achievement .....	5
Second Step.....	13
Second Step Research.....	15
Importance of Instruction Time .....	24
Importance of Implementation.....	24
Research Questions.....	26
Methods.....	27
Sample.....	27
Random Assignment.....	27
Measures .....	28
<i>Student Variables</i> .....	28
<i>Classroom Variables</i> .....	30
<i>School Variables</i> .....	33
Analyses.....	33
<i>First Research Question</i> .....	33
<i>Second Research Question</i> .....	36
Results .....	39
First Research Question .....	39
Second Research Question.....	40
Discussion .....	44
Limitations .....	44
Implications.....	46
Author Note .....	49
Appendix.....	50
References.....	69

## List of Tables

Table 1: Treatment Effects for Second Step on Report Card Grades and Standardized Test Scores.....	53
Table 2: Descriptive Statistics.....	54
Table 3: Student Sample Baseline Differences Between Treatment and Control Schools.....	55
Table 4: Correlations of Student-Level Variable.....	56
Table 5: Differential Response Rates for Teacher Questionnaire Between Treatment and Control Groups.....	57
Table 6: Teacher Questionnaire Items about the Time-spent of in Formal Social Skills Instruction in the Classroom.....	58
Table 7: Report Card Items for the Composite Baseline Measure of Academic Achievement.....	59
Table 8: Report Card Items for the Outcome Report Card Grades Measure.....	60
Table 9: The Intraclass Correlations for the Time-Spent in Formal Social Skills Instruction Outcome Variables.....	61
Table 10: Difference between Treatment and Control Schools on Teacher Reports of the Delivery of Formal Social Skills Instruction in the Classroom.....	62
Table 11: Proportion of the Variance Explained Between Classrooms in the Achievement Outcome Measures by the Separate Models.....	63
Table 12: Within-Classroom Fixed Effects Model for Report Card Grades Outcome.....	64
Table 13: Effects of Time-spent in Social Skills Instruction on Report Card Grades.....	65
Table 14: Within-Classroom Fixed Effects Model for Standardized Test Scores Outcome.....	66
Table 15: Effects of Time-spent in Social Skills Instruction on Standardized Test Scores.....	67
Table 16: Correlations of Classroom-Level Variables and Aggregated Achievement Outcome Measures.....	68

## Literature Review

Social and emotional learning programs can improve social-emotional, behavioral, and health outcomes for youth (CASEL, 2008; Durlak & Wells, 1997; Wilson & Lipsey, 2007), but less research has examined the effects of social skills programs on achievement outcomes for youth (Rones & Hoagwood, 2000). Research about universal social skills programming generally demonstrates, and theory supports, that an increase in students' prosocial skills and a decrease in behavioral problems correlate with an increase in student achievement (Payton, Weissberg, Durlak, Dymnicki, Taylor, Schellinger, & Pachan, 2008). In other words, an increase in academic achievement is expected when the population is in need of the intervention or the intervention successfully increases social skills.

Gottfredson, et al. (in preparation) delivered a well-implemented, school-based universal social skills program, *Second Step*, to a population with low rates of social-behavioral problems and found no positive effects across measures. It is hypothesized here that formal skills instruction may detract from academic instruction time in the classroom. The first goal of this study is to examine the program effects on the amount of time-spent implementing formal social skills instruction and academic instruction time in the classroom. The second goal of this study is to examine the effects of time-spent in formal social skills instruction on student report card grades and standardized test scores. Classroom instruction time leads to greater student achievement (Bodovski & Farkas, 2007; Connell, Spencer, & Aber, 1994; Connell & Wellborn, 1991; Fisher & Berliner, 1985; Marks, 2000; Skinner, Wellborn, & Connell, 1990), and prevention programs with the biggest effects are

those that are implemented with greater fidelity and quantity (D. C. Gottfredson, Fink, Skroban, & Gottfredson, 1997; Weissberg, Kumpfer, & Seligman, 2003; Wilson & Lipsey, 2007). It is hypothesized that an increase in time-spent in social competency instruction may lead to less opportunity to engage in classroom academic instruction and as a result, decreased achievement.

### *Universal Social Skills Programming*

Universal social skills programs are designed to promote social-emotional and behavioral outcomes for students. Social and emotional learning (SEL) is a conceptual framework designed to guide programs that promote children's social competencies and problem-solving skills and support academic achievement. SEL programs are designed to enhance students' skills to recognize and manage their emotions, appreciate the perspectives of others, develop positive problem-solving competencies, and handle interpersonal situations effectively (Collaborative for Academic, Social, and Emotional Learning, 2003). SEL outcomes are sought through effective classroom instruction, student engagement in positive activities in and out of the classroom, and through family and community involvement. Interventions that promote students' social skills and reduce problem behaviors are increasingly endorsed, and schools have been identified as potential environments for implementation (Mayer, 1995). For example, the National Association for School Psychologists (2006) endorses a three-level prevention service delivery model (i.e., universal, targeted, intensive), based on the theoretical prevention-intervention framework established in a report by the Institute of Medicine (Munoz, Mrazek, & Haggerty, 1996).

A review of the literature by Greenberg et al. (2003) indicated that well-designed and well-implemented prevention programs can promote positive academic, social, emotional, and health behaviors. Durlak and Wells (1997) conducted a meta-analysis reviewing 177 primary prevention programs designed to prevent behavioral and social problems. Their review suggested that most programs reduced problem behaviors and increased competencies, with effect sizes ranging from 0.24 to 0.39. In a similar review, 219 studies on the effectiveness of school-based violence prevention programs were synthesized using meta-analysis techniques (Wilson & Lipsey, 2005). Of these, 61 were universal programs, for which an effect size of 0.61 was found, suggesting that these programs can be effective in reducing aggressive and disruptive behaviors. Additionally, the Conduct Problems Prevention Research Group (CPPRG, 2000) implemented a multi-site, multi-component preventive intervention program for students at high risk for long-term antisocial behavior. The program included both a classroom-level program (universal) and an individual student at-risk program (indicated), and was randomized at the school level. After one year of implementation, CPPRG found “moderate positive effects” on indicated students’ social and emotional outcomes. At the universal level, treatment schools showed lower overall levels of aggression and higher ratings of classroom atmosphere quality. Thus, social competency programs can improve social-emotional and behavioral outcomes for youth (CASEL, 2008).

The primary outcomes of studies examining school-based social competency programs and prevention interventions, however, are predominantly behavioral in nature. Reviews and meta-analyses of these programs less often examine the program

effects on academic achievement – the most common foci for school-based prevention programs are social competencies and interaction skills, and coping/stress management (Consortium on the School-Based Promotion of Social Competence, 1994). A review of school-based mental health services research concluded that program evaluations of these services have typically ignored or underemphasized school-relevant outcomes such as student achievement (Rones & Hoagwood, 2000).

The current climate of education however, emphasizes academic achievement outcomes beyond all else, including socialization outcomes. Accordingly, prevention programs that are designed to increase social competency and decrease problem behaviors must also aim to improve students' academic achievement. Social and emotional learning programs and social competency interventions that do not aim to improve achievement test scores may have difficulty garnering support. The Institute of Education Sciences (IES, 2008), for example, names specific requirements for applicants seeking grants in 2009 for Social/Behavioral programs. In order to receive funding, applicants must address:

interventions (e.g., curriculum, classroom management, teacher professional development) that are implemented in schools and are intended to improve the social and behavioral context for academic learning in schools or other education delivery entities from kindergarten through high school; or measures of children's behaviors and teacher classroom management practices that are predictive of academic learning from kindergarten through high school.

It seems that IES intends only to fund social/behavioral intervention research designed to directly or indirectly support academic outcomes. Thus, social competency and behavioral interventions that improve achievement are not only important for students, but also for educators, researchers, and policy makers.

### *Effects of Social Skills on Achievement*

Research generally supports the positive relation between social-emotional competency and academic achievement but the evaluations of school-based social competency interventions are less clear. Measures of intelligence have modest but consistent negative correlations with delinquency, but the link between the two may be largely mediated by performance in and attitudes toward school (G. D. Gottfredson, 1981). It is well documented that school performance and delinquency are related, as research dating from 1936 to the present has found negative, moderately-sized associations between school performance and delinquency (Kellam, Brown, Rubin, & Ensminger, 1983; G. D. Gottfredson, 1981; Healy & Bronner, 1936). Studies have shown that student academic performance is inversely related to antisocial behavior (Hawkins, Farrington, & Catalano, 1998; Nelson, Martella, & Marchand-Martella, 2002; Maguin & Loeber, 1996) and that higher aggression is associated with academic difficulties (Cantrell & Prinz, 1985; Cole, Lochman, Terry, & Hyman, 1992). In a study by Orpinas et al. (2000), middle-school students who did less well academically were significantly more likely to engage in violence-related behaviors, including aggression, substance use, and weapon-carrying behavior.

Students' social competence and low rates of administrative disciplinary actions both contribute to a positive school climate, which is inversely related to

lower levels of school crime and violence. Improved student behavior is associated with a reduction in disruptions and thus an environment more focused on academics (Nelson, Martella, & Marchand-Martella, 2002; Verdugo & Schneider, 1999). There is also a clear, strong empirical association between student social competency and academic performance (DiPerna & Elliot, 1999; Hawkins, Von Cleve, & Catalano, 1991; Kupersmidt, Coie, & Doge, 1990; Malecki & Elliot, 2002; Welsh, Parke, Widaman, & O’Niel, 2001; Wentzel, 1991). It is reasonable to speculate that gains in academic achievement are expected when the population is in need of prosocial skills or when the intervention works to increase social skills and decrease problem behaviors in schools.

Theory extends, and research generally demonstrates, that an increase in prosocial skills can lead to an increase in student achievement. A 2008 research brief by CASEL cites a meta-analysis in review (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2008) that summarizes the effects of universal school-based SEL programs. This meta-analysis reports a decrease in conduct problems and an increase in social/emotional skills and prosocial behaviors, and a 0.11 effect size increase in achievement test scores and improved report card grades (Payton, Weissberg, Durlak, Dymnicki, Taylor, Schellinger, & Pachan, 2008). The authors state that these positive effects on student achievement are comparable to or exceed the benefits on similar outcomes found in other meta-analyses of psychosocial or educational interventions for school-aged youth (DuBois, Holloway, Valentine, & Cooper, 2002; Durlak & Wells, 1997; Haney & Durlak, 1998; Hill, Bloom, Black, & Lipsey, 2007). Each of these meta-analyses reported increases in social competency and academic

achievement, supporting the idea that a change in the one can be related to a change in the other. But, these meta-analyses do not specifically target school-based, universal social skills programs, and include, for example, individual mentoring programs and hospital-based studies. The sample of school-based, universal social skills programs reviewed in these analyses is actually quite small. Thus, the face value of these meta-analyses can be misleading, and the literature examining the effects of universal social skills programming on student achievement are not as straightforward or robust as the results of these analyses suggest.

The Institute of Education Sciences at the U.S. Department of Education established the What Works Clearinghouse (WWC) in 2002 to provide educators, policymakers, researchers, and the public with a source of scientific evidence about what works in education. The WWC evaluates school-based social skills interventions intended to improve student outcomes related to positive character development, prosocial behavior, and academic performance.

A search of the WWC database reveals only three character education interventions that had effectiveness ratings for academic achievement outcomes (What Works Clearinghouse, 2008). The first program had two supporting studies that met evidence with reservations, but indicated no positive evidence of program effects on academic achievement.

The second program had evidence of potentially positive effects on academic achievement, with no overriding contrary evidence, but the extent of that evidence was determined to be small. Two studies reviewed the effects of this program, one met the WWC evidence standards (Dietsch, Bayha, & Zheng, 2005), and the other

met the standards with reservations (DeVargas, 1998). The 1998 study by DeVargas did not even include an academic achievement outcome, which is not an unusual practice for such a study. The 2005 study by Dietsch, Bayha, and Zheng was a randomized controlled trial that included 372 fourth-grade students from eight schools located in rural, poor, and ethnically diverse communities. Results indicated no treatment effects in terms of behavior, knowledge, or attitudes; however, the study did find positive, statistically significant treatment effects for math grades (ES = 0.46) and absences (ES = -1.13), but not for reading grades (ES = 0.31). These results indicate no treatment effect on social skills but a positive effect on math grades only which seems anomalous, making implications difficult to draw.

The third program had strong evidence of a positive program effects and the extent of that evidence was determined to be medium to large. This program was designed to promote character development *and* academic achievement as it included daily oral language development and weekly writing assignments. Twelve studies reviewed the effects of this program, but only two met WWC evidence screening. One study met the WWC standards (Flay, Acock, Vuchinich, & Beets, 2006), and the second study met standards with reservations (Flay & Allred, 2003). The 2006 study by Flay et al. was a randomized controlled trial that examined the outcomes at the end of the third year of program implementation of 2,666 third and fourth grade students in 20 elementary schools. Behavioral outcomes included reports of substance use, and academic outcomes included reading and math standardized test scores. Findings indicated significant effects favoring the treatment group for behavioral outcomes

(average ES = 0.52), and positive treatment effects for reading (ES = 0.19) and math (ES = 0.17) test scores, but these results were not statistically significant.

The 2003 study by Flay and Allred used a quasi-experimental design and included 36 elementary schools. Behavioral outcomes included violence and suspensions rates, and academic outcomes included reading and math standardized test scores and grade retention rates. Findings indicated significant treatment effects on behavioral outcomes (average ES = 0.50) as well as significant positive treatment effects for reading and math standardized test scores at the elementary, middle, and high school levels. Because standard deviations were not reported, the WWC could not compute the student-level effect sizes, however, it was reported that the statistical significance for the achievement outcome was comparable to other studies.

Thus, a search of the WWC for publications that evaluate school-based character education interventions to improve student outcomes related to positive character development, prosocial behavior, and academic performance yielded meager results. Only three programs had effectiveness ratings, and only one of these demonstrated medium/large effects of character education on academic achievement. But, 10 of the 12 studies examining this program did not meet IES standards, suggesting that research in this area needs to be more rigorous. The study that found significant prosocial program effects also found significant achievement program effects, suggesting that to increase student achievement the intervention must also increase social skills.

The literature outside of the WWC engine also yields mixed academic achievement results for universal social competency programs. In 2000, Solomon,

Battistich, Watson, Schaps, and Lewis conducted a quasi-experimental study, examining a comprehensive elementary school character development program designed to create a strong school community through classroom, school-wide, and parent involvement components. Two schools in each of six school districts across the U.S. were selected as treatment schools and were each matched with a control school based on school size, student poverty level and ethnic distribution, percent of limited-English speakers, and achievement test results. Reading and math achievement were assessed by standardized achievement tests administered at each grade level in each district. Program effects were assessed with a series of univariate ANOVA and ANCOVA analyses, comparing change in achievement scores across time for students in treatment and comparison schools. There were no consistent overall program effects for the entire sample, but variations in levels of implementation affected the results. In the five program schools that made significant progress in implementation, positive effects were found in every domain except academic achievement. The results of standardized achievement tests among these five schools and their respective comparison schools varied. In one treatment school, there was a negative program effect on math achievement test scores ( $ES = -0.29$ ), in another, there was a positive program effect on standardized performance tests of reading in year two ( $ES = 0.61$ ), and math in year three ( $ES = 0.39$ ), and in another school, there was a positive program effect on standardized tests of math in year three ( $ES = 0.54$ ), but no significant difference in reading performance in any year. These results do not indicate a clear pattern of program effects on standardized test scores. One reason this program may have produced positive achievement effects is that of

its three major classroom facets (i.e., developmental discipline, cooperative learning, and literature-based language arts curriculum), two are academically motivated. Thus, this program may be considered more of a comprehensive universal intervention, involving both social and academic skill training, as opposed to just social skills.

Catalano et al. (2003) examined a comprehensive, multifaceted prevention program focused on enhancing protective factors and reducing risk factors and problem behaviors. Teacher-reported student academic outcomes were analyzed using growth curve analyses (HLM) controlling for income, gender, and baseline scores; this analysis provided good statistical conclusion validity. Results indicated a significant decrease in teacher ratings of antisocial behavior and a significant increase in teacher ratings of social competency for treatment students over control students. The results also indicated significantly higher teacher and parent reports of academic performance for treatment students. Although students in the treatment program showed an increase in parent and teacher ratings of academic performance, the construct validity of the achievement measures was limited as the study used a three-question teacher report and two-question parent report of achievement, rather than report card grades and standardized test scores.

In 2004, Battistich, Schaps, and Wilson examined the effects of a comprehensive elementary school universal intervention designed to promote positive development. The overall results indicated significant positive treatment effects for student-rated, school-related attitudes of positive teacher-student relations ( $ES = 0.11$ ), and liking for school ( $ES = 0.13$ ). The full sample of program and comparison students did not differ with respect to academic achievement, however results

indicated that students in high-implementation elementary schools had higher academic achievement than their matched comparison students (GPA ES = 0.37, standardized test score ES = 0.37). These findings support the idea that the fidelity and quality of implementation may positively affect treatment outcomes; however, data were analyzed using ANCOVA. Hierarchical modeling would have better accounted for the clustering of students within schools, as ignoring this clustering leads to standard errors that are too small and a higher probability of finding differences between treatment and comparison students.

There is a substantial amount of evidence supporting comprehensive prevention programs targeting maladaptive social-emotional, behavioral, and violence problems. Much of the research also supports the positive relation between universal social competency programs and academic achievement. The evidence for some of these positive findings, however, has methodological limitations, including a lack of a comparison group, a lack of implementation fidelity measures, and inadequate measures of academic achievement (i.e., teacher and parent reports of achievement as opposed to standardized test scores and report card grades). For example, in a meta-analysis of violence prevention programs, 78% of the studies reviewed used indirect measures of academic achievement such as teacher, parent, or student reports (Wilson and Lipsey, 2005). Randomized, controlled longitudinal experiments are needed to better understand the effects that social skills interventions have on academic achievement, as measured by report card grades and standardized test scores.

When achievement outcomes are evaluated, research generally reports significant prosocial program effects along with positive achievement effects, and it

seems that gains in social skills must be reached and maintained in order to engender a positive educational environment and thus gains in achievement. But if the intervention reduces academic instruction time by substituting a social competency curriculum, particularly in a school where no behavior problems exist, academic achievement scores may be negatively affected. It is important to understand the academic effects of universal social skills programming that is ineffective in promoting social skills or is delivered to population with low rates of social-behavioral problem. Perhaps school-based social skills programming can positively affect academic outcomes only to the extent to which it increases students' prosocial behaviors and schools' social climate. This study examined an ineffective delivery of Second Step, a prevention program which purposes to increase students' social competency skills and to indirectly improve student academic achievement.

### *Second Step*

Second Step is a universal violence prevention program designed to reduce the development of social, emotional, and behavioral problems, and promote development of social competencies. Second Step is grounded in the theoretical foundations of social learning theory (Bandura, 1986), social information-processing (Dodge, Pettit, McClaskey, & Brown, 1986), cognitive-behavioral therapy, and Luria's 1961 model of self-regulation. The program is widely used in the United States, has been adapted for use in Australia, Canada, Germany, New Zealand, Norway, and the United Kingdom, and has been translated into Spanish (Frey, Hirschstein, & Guzzo, 2000).

Second Step provides classroom lessons and activities that teach social skills such as empathy, emotion management, problem solving, and cooperation (Committee for Children, 2007). The curriculum is organized around three areas of social-emotional competency: empathy, social problem solving, and anger management. The empathy unit focuses on identifying feelings in self and others, considering others' perspectives, and responding appropriately to others' emotions. The social problem solving unit focuses on identifying problems, brainstorming solutions, evaluating the appropriateness of a solution, and strategizing on what do next. The anger management unit focuses on strategies to help students recognize anger cues in the self, use positive statements and stress-reduction techniques, practice reflection, and react appropriately to domain specific situations (e.g., receiving criticism and being left out).

The curriculum is intended to be taught once or twice a week in the classroom. The lessons are structured around cards depicting social-emotional situations, and facilitators are given a guided lesson script to lead discussions, conduct role play exercises, practice decision making, and rehearse behavioral skills (Frey, Hirschstein, & Guzzo, 2000). Lessons are accompanied by notes regarding child development and transfer of training ideas, and are supplemented by homework, books, videos, and extension activities. The curriculum also calls for opportunities to be created by the facilitator in which student skill development and behaviors may be practiced and then positively reinforced (Frey, Hirschstein, & Guzzo, 2000).

### *Second Step Research*

Second Step is promoted by the Committee for Children (2007) as a research-based violence prevention program. It has been the subject of multiple evaluations, including pilot studies, formative evaluations, quasi-experiments, and randomized controlled trials, but the literature examining the efficacy of Second Step is limited in rigor and scope (Appendix). A review of this research can be found in Harak (2006, 2008), and Carey, Dimmitt, Hatch, Lapan, Lee, and Whiston (2008).

Several formative studies were conducted to assist in program development. These studies found significantly higher social skills knowledge in treatment versus comparison groups, and also high teacher ratings of the program, lesson format, and the instructor guide (Beland, 1988, 1989, 1991; Frey & Sylvester, 1997; Moore & Beland, 1992). These formative evaluations, however, examined knowledge specific to the Second Step curriculum, not explicit student behavior, and did not randomly assign students to groups. The difference between knowledge- and skill-acquisition is important; the former assesses students' learning of the program's curriculum (e.g., vocabulary and knowledge of social skills lessons), while the latter assesses students' observed positive and negative behaviors.

Kelder et al. (1996) developed a three-year, comprehensive, school-based violence prevention program designed to prevent violence among middle school students in an urban school district. The intervention, based on Social Learning Theory (Bandura, 1986), included four main components: modification of the school environment, a violence-prevention curriculum, peer leadership, and parent education. The Second Step curriculum was used as the violence-prevention

curriculum component; the four-component design of the intervention prevented the assessment of the effect of an individual component. Four treatment and four control schools were matched on ethnicity, size, and baseline violence and randomly assigned to condition. A nested, cross-sectional and cohort design was used in which the school was the unit of analysis. A first year assessment indicated a need for a violence-prevention intervention in these schools (Kelder, et al., 1996). An outcome evaluation at the conclusion of two years of program implementation included 2,246 students who had baseline data and at least one follow-up survey completed (Orpinas, et al., 2000). Multilevel general linear models were developed to examine within- and between-school differences. The nested cohort design model included baseline aggression scores, race/ethnicity, and academic performance as covariates. No significant treatment effect was found between treatment and control groups for the cohort evaluation, and the authors reported that a general trend of the intervention was towards a negative treatment effect, although these effects were not statistically significant. It should be noted that implementation measures indicated that implementation fell far short of expectations and that Second Step was used as a single component of a larger school intervention and was therefore not analyzed in isolation, so this study provides a poor test of the efficacy of Second Step.

Grossman et al. (1997) conducted a school randomized study of Second Step to determine whether the program led to a reduction in aggressive behavior and an increase in prosocial behavior. Six matched pairs of schools in four districts were randomly assigned to treatment or control conditions. The schools were paired based on the proportion of students receiving free or reduced meals and the proportion of

minority enrollment. Results were analyzed at the school-level using Generalized Estimating Equation while adjusting for individual level covariates (i.e., baseline behavior score, academic performance, behavioral problems, grade, and sex). There were no significant group differences found in parent-reports of student behavior or in teacher-reports of classroom behavior. But behavioral observations recorded on the playground and in the cafeteria revealed negative treatment effects for physically aggressive behavior and positive treatment effects for neutral/prosocial behavior. These findings were limited by a lack of data on implementation fidelity as well as the fact that Second Step was not implemented school-wide, as intended. This study is cited by the program creators as evidence that Second Step decreases aggression and increases positive behavior (Committee for Children, 2007). But a more complete summary would be that behavioral observations indicated Second Step reduced aggressive behaviors and increased prosocial behaviors outside of the classroom (i.e., on the playground and in the cafeteria), but no effects were found among parent- and teacher-reports of student behavior.

Second Step has also been examined among young inner-city children. McMahon, Washburn, Felix, Yakin, and Childrey (2000) studied the effect of the program over a period of 28 sessions among preschool and kindergarten students living in Chicago public-housing developments. The sample was drawn from one preschool and one kindergarten at two separate sites, and consisted of 109 students, the majority of whom were African American children, all of whom qualified for free and reduced meals (FARM). The authors assessed program knowledge, behavior problems, and social skills at pre- and posttest intervals as measured by child

interviews, teacher ratings of social skills, and behavioral observations at the classroom level (i.e., disruptive behavior, verbally aggressive behavior, and physically aggressive behavior). According to the results of the child interviews, there was a significant effect in knowledge gained about the identification of feelings and how to respond to conflict situations. The behavioral observations indicated a significant decrease over time for all three behavioral categories (i.e., disruptive behavior, verbal aggression, and physical aggression). The teacher ratings revealed no significant effects. The significant reduction of problem behavior according to the observations and the lack of effects on teacher ratings are consistent with studies by Grossman et al. (1997) and Gottfredson et al. (in preparation). The study is cited by the program's developers as evidence that Second Step increases knowledge of social skills (Committee for Children, 2007).

In a similar study, McMahon and Washburn (2003) implemented Second Step in two inner-city middle schools with 156 African American students in grades five through eight. Analyses of variance (MANOVA and ANOVA) controlling for sex, school, and grade were conducted for prosocial behavior, aggression, empathy, impulsivity, sense of school membership, and program knowledge. The study found that treatment participants gained knowledge about violence prevention skills and showed an increase in prosocial behaviors and empathy, but no effect was found for aggressive behavior. School membership had the most influence on changes in outcomes. This study is cited by the Committee for Children (2007) as evidence that increased empathy skills (a core component of Second Step) are related to lower

levels of aggression, even though no program effect was found for aggressive behavior.

McMahon and Washburn (2003) have attributed the inconsistent effects of Second Step found in the research to poor implementation and low community acceptance; they have consequently called for research to examine the effects of Second Step implemented on a school-wide level, as it is designed, in order to capitalize on possible effects and appropriately assess program efficacy. In response, Cooke, Ford, Levine, Bourke, Newell, and Lapidus (2007) used a non-experimental design (pre-test post-test, no comparison group) to evaluate Second Step implemented as a city-wide intervention in all elementary schools of a small city. The consented participants were 741 third and fourth grade students from five elementary schools. Students in grades one and two were excluded because a self-report was not deemed appropriate for this age group, and students in grade five were excluded due to the potential confounding caused by their participation in another violence prevention program. Aggressive antisocial and prosocial behavior was assessed at the beginning and end of the school year, before and after implementation, using student self-reports, archival data (discipline referrals), and behavioral observations replicating those used by Grossman et al. (1997). The student self-reports yielded significant improvements in coping, cooperative behavior, suppression of aggression, and consideration of others. The authors reported non-significant results, including a decrease in impulse control and an increase in aggressive behavior. Baseline rates of violence and aggression were low, and thus, no significant results were found for the behavioral observations, and not enough data were present to analyze discipline

referrals. The limitations of this study include a lack of a control group and an inadequate method in which simple *t*-tests were used to analyze the data with no consideration for multi-level or school effects. Although this study is not capable of providing information about the efficacy of Second Step, it does lend support to the idea that low baseline levels of aggression and discipline referrals make it difficult to analyze the effects of a social/behavioral intervention.

Frey, Nolen, Van Schoiack-Edstrom, and Hirschstein (2005) examined the effects of Second Step on social cognitions and prosocial/antisocial behaviors. Eleven schools were randomly assigned to treatment or control conditions, and after the first year, four additional schools were non-randomly assigned to the control condition ( $n = 1,253$  students). The effects of the program were measured by teachers' ratings of students' social behaviors, students' self-reports of hostile attributions, and behavioral observations in two contrived conflict situations. Omnibus multivariate analyses of covariance (MANCOVA) were used as the primary analyses, controlling for sex, grade, and baseline ratings of social competence. When this analysis revealed an interaction of a covariate with group, confirmatory multilevel modeling (HLM) was used to model nested data, indicating good statistical conclusion validity. The results indicated that students who received Second Step were more likely to prefer prosocial goals, were associated with significant benefits in prosocial behavior, and were more likely to espouse prosocial goals and social reasoning than those in the control groups. In addition, teachers in the treatment schools in the first year reported greater increases in social competence and greater decreases in antisocial behavior than did teachers in control schools. The results of the naturalistic behavioral observations

indicated that students in the treatment group displayed less aggression than those in the control group. These results correspond with similar studies that reported higher observational ratings of prosocial behaviors in treatment groups (Gottfredson et al., in preparation; Grossman et al., 1997; McMahon et al., 2000) and are cited as evidence that Second Step leads to improved social competence and less adult conflict intervention (Committee for Children, 2007).

The literature examining the efficacy of Second Step is limited by, among other things, inadequate methods, a lack of control group, and improper implementation. And most importantly, only one study evaluating the efficacy of Second Step has examined program effects on academic achievement (Gottfredson et al., in preparation). In an evaluation of the scientific research evidence supporting the effectiveness of Second Step, the National Panel for School Counseling Evidence-Based Practice concluded that future studies of Second Step need to determine whether the social/behavioral skill development affects academic outcomes (Carey, Dimmitt, Hatch, Lapan, Lee, & Whiston, 2008). The Committee for Children (2007), the creator of Second Step, claims that the program integrates academics with social and emotional learning by aligning with state academic content standards. The curriculum at every level of the program is reported to include activities that extend the lessons into academic content areas such as language arts, math, and science (Committee for Children, 2007). Quite apart from the potential conflict of interests of the source, the plausibility of such a statement is dubious, considering the variation in state standards and differences between academic subjects. Furthermore, research has commonly supported the correlation between social/emotional skills and academic

achievement, but there remain inadequate empirical results about the effect of school-based prevention programs on academic outcomes.

Recent recommendations suggest that in order to enhance the evaluations of school-based prevention programs, multiyear analyses should be employed, a broader range of outcomes should be incorporated, and implementation should be assessed to ensure fidelity (Flay et al., 2005; Greenberg et al., 2003).

A recent study by Gottfredson et al. (in preparation) incorporated these recommendations in an examination of the efficacy of Second Step. This was a school-randomized, controlled trial in 12 elementary schools, with data collected on consented students in grades one through five. Outcomes were measured by parent, teacher, and student ratings of student social competency, student report card grades and standardized test scores after one, two, and three years of implementation, and behavioral observations after two years of implementation. Implementation data were also collected and reported, and summaries of the data suggest that the program was well-implemented (Harak, 2008; Gottfredson et al., in preparation). To model potential treatment effects, data were explored using hierarchical linear modeling with the individual at Level 1 and the school at Level 2. The results showed no consistent pattern of positive effects across any outcome provided no support for the use of Second Step to prevent problem behavior or promote social competency. The authors also report a major limitation was that baseline data did not reveal high rates of problem behaviors and speculate that there was no problem with students' social skills or behavior that necessitated an intervention. Across three years, the results also

yielded no program effects on report card grades (coefficients<sup>1</sup> ranging from –0.34 to 0.06), but did indicate a non-significant, negative trend of program effects on standardized achievement test scores (coefficients ranging from -0.30 to -0.18). This trend did not indicate a statistically significant negative program effect on academic achievement as the confidence intervals of the effect estimates were broad and included many positive values. To replicate the achievement results found by Gottfredson et al. with the sample in this study (combined cohorts rather than separate analyses per grade level), the student covariates sex, ethnicity, FARM status, and baseline achievement were included in the Level 1 model, and treatment was tested at the school-level (Level 2). Table 1 shows the Second Step treatment effect parameters for report card grades and standardized test scores.

The present study uses the data gathered by Gottfredson et al. (in preparation) to further explore the academic achievement effects of a universal, school-based social skills program that had no effects on social skills or problem behaviors. A qualitative analysis of the implementation of Second Step found that teachers implementing the program experienced tension between teaching academic content and teaching the program curriculum, and the curriculum was not administered to part of the sample because of the need to focus on academic subjects (Larsen & Samdal, 2007). The goal of this study is to examine the effect of Second Step on *time-spent* in formal social skills instruction and academic instruction time, and also to investigate the effect of *time-spent* in formal social skills instruction on students' academic achievement.

---

<sup>1</sup> A coefficient here is approximately equal to an effect size. All dependent variables were standardized so that a coefficient represents a percent of a standard deviation change in the dependent variable for a unit change in the independent variable.

### *Importance of Instruction Time*

A review of models and theories of student classroom learning described variables important to the performance of individual learners and determined that “quantity of instruction” was a construct common to all learning theories (Haertel, Walberg, & Weinstein, 1983). In other words, academic instruction time is an essential component of student classroom learning and achievement. Research has consistently shown that both increased instructional time and the proportion of time students spend actively engaged in learning are positively related to academic achievement (Fisher & Berniler, 1985; Brown & Saks, 1986; Hossler, Stage, & Gallagher, 1988). Research has also shown that the amount of time that teachers allocate to instruction in a particular curriculum content area is positively associated with student learning in that content area (Fisher, Berliner, Filby, Marliave, Cahen, & Dishaw, 1981; Bodovski, & Farkas, 2007). The positive correlation between academic engagement and achievement-related outcomes (i.e., standardized test scores and report card grades) has been replicated across various samples and all school age-groups (Connell, Spencer, & Aber, 1994; Marks, 2000; Skinner, Wellborn, & Connell, 1990; Connell & Wellborn, 1991). Thus, increased classroom instruction time leads to greater student achievement.

### *Importance of Implementation*

Prevention programming must be implemented with integrity to produce optimal outcomes (Weissberg, Kumpfer, & Seligman, 2003; Collaborative for Academic, Social, & Emotional Learning, 2008). A social/behavioral prevention program must be based on theory and implemented with fidelity in order to reduce

problem behaviors and increase protective factors. G. D. Gottfredson (1984), outlined a program development evaluation (PDE) method to facilitate the design, implementation, and testing of quality programs. This method was based on organization development, field research practices, and evaluation research. The first principle of PDE is that an effective project must be guided by an explicit theory that is consistent with evidence from research and can be translated into practice. In addition to a foundation in theory and implementation integrity, a program must effectively reduce problem behaviors and increase protective factors in order to support student academic achievement.

In a review of prevention research, D. C. Gottfredson, Fink, Skroban, and Gottfredson (1997) noted that most prevention programs are not implemented according to the underlying program theory and are consequently not effective. They suggested that features of the school and ecological context are related to the strength and fidelity of implementation. The program curriculum must be explicit and not confusing or burdensome to those who deliver it, usually teachers. In addition, teacher participation and support is crucial to positive outcomes, and staff training must be effective and sufficient in length. The characteristics of school personnel and school climate are also relevant to prevention programs; these include such variables as teacher self-efficacy, teacher morale, teacher instructional skills, school leadership, school resources, and the organizational capacity of the school.

D. C. Gottfredson, Gottfredson, & Skroban (1998) demonstrated the implications of quality implementation in a case study examination of a universal prevention program designed to increase social competencies and reduce problem

behaviors. The program was implemented across five years in a middle school serving a predominantly residential, lower middle-class area. The research question was not about effectiveness, but about how well-designed, well-implemented prevention programs generalize to the realistic conditions found in schools. The prevention program was not implemented according to expectation and results showed that the program failed to reduce any form of problem behavior or any measured predictor of problem behaviors.

In a meta-analysis, Wilson and Lipsey (2007) examined the effectiveness of school-based psychosocial prevention programs for reducing aggressive and disruptive behavior. The authors concluded that effects were larger for better-implemented programs, and more specifically, programs with greater treatment dosage (i.e., duration or frequency) tended to produce larger reductions in aggressive behavior (Wilson & Lipsey, 2007).

### *Research Questions*

Prevention programs with the greatest effects are those that were implemented with greater integrity, which included both the quality and quantity of program delivery. Large-scale interventions succeed according to the amount and quality of assistance the population receives (Durlak, 1995). As universal prevention programming typically requires delivery at the classroom-level, the time necessary for implementation to produce program results may decrease time available for classroom academic instruction. As academic instruction time is positively related to student achievement, decreasing classroom academic instruction may have a negative effect on achievement. The research questions of this study are:

1. Does the treatment program (Second Step) result in more time-spent in formal social skills instruction and less academic instruction time in the classroom than the programming in the control schools?
2. Does time-spent in formal social skills program instruction decrease academic achievement?

## Methods

### *Sample*

The sample included 12 public elementary schools in a Maryland school district (Table 2). The teacher sample was 113 third, fourth, and fifth grade teachers in both treatment and control schools. The student sample was 1,724 students in the three cohorts that were to receive treatment for three years; 48% of the sample was female, and 8% of the sample qualified for free or reduced meals. Because of the lack of ethnic diversity in the sample, the ethnicity variable was collapsed into two categories: historically high achieving students, which made up 92% of the sample and included White and Asian students; and historically low achieving students, which included Black, Latino, and American Indian/Alaskan Native students.

### *Random Assignment*

The intervention was a school-randomized, controlled trial in 12 suburban elementary schools. Schools were paired based on similarity of ethnic composition and FARMS participation, and one school of each pair was randomly assigned to the intervention condition. The treatment prevention program was delivered in grades one

through five in the intervention schools, and the control schools continued their regular activities but agreed not to implement the treatment program, Second Step.

By definition, randomization rules out sampling bias as a threat to internal validity. But in order to account for sampling error, a hierarchical analysis of the baseline sample data was conducted separately for the covariates as dependent variables, as Bernoulli distributions, and treatment at Level 2 to determine whether treatment status significantly predicted each of the covariates. Table 3 shows that no post-randomization treatment-control group differences existed in terms of sex, ethnicity, FARM status, or report card grades.

### *Measures*

#### *Student Variables*

*Student demographics.* The following student demographic covariates were included in the individual level model: sex, ethnicity, free and reduced meal status (FARMS), which were obtained from school records. Females were coded one. Ethnicity was a dichotomous variable, and the historically higher performing students were coded one. Free or reduced meal participation was the best estimate of socioeconomic status and was also a dichotomous variable, with those qualifying for free or reduced meals coded one. Table 4 displays the correlations of all student-level variables.

*Report card grades outcome.* Report card grades were taken from school records. In this school district, the format and items of the report cards for grades one and two are different from grades three, four, and five. Cohort 1 began the first year of implementation in grade one and finished in grade three (34.5% of the sample);

Cohort 2 began the first year of implementation in grade two and finished in grade four (34.5% of the sample); and Cohort 3 began the first year of implementation in grade three and finished in grade five (31% of the sample).

Standardized test score baseline data were not available for all students, so the fourth quarter report card grades of the 2005-2006 school year were used as academic baseline for all analyses. A factor analysis was conducted for the 15 baseline report card items from the areas of language arts, math, science, social studies, art, music, and technology, and seven items had high loadings on academic achievement (Table 7). After conducting an internal consistency item analysis, the mean of the seven items were used to form the baseline report card achievement measure (Cohort 1 alpha = .88; Cohort 2 alpha = .90; Cohort 3 alpha = .92).

The school district changed the format of report cards for all grades levels for the 2006-2007 school year, so the items on the outcome measure were different than the items on the baseline measure. The outcome report card academic measure was the “final” report card grades of the 2006-2007 school year. A factor analysis was conducted for the 12 report card items and six items had high loading on academic achievement (Table 8). After conducting an internal consistency item analysis, the mean of the six items were used to form the outcome report card achievement measure (Cohort 1 alpha = .91; Cohort 2 alpha = .92; Cohort 3 alpha = .92).

Due to skewness in the data, the inverse log transformed report card scores were reflected so that high values reflected high grades. The measures were then standardized as z-scores separately by cohort (controlling for cohort variance in the final sample) and then standardized again among all three cohorts.

*Standardized test scores outcome.* Standardized reading and math test scores were taken from the 2007 Maryland State Assessment (MSA), the standardized state test in Maryland administered to students in grades three through five (Maryland State Department of Education, 2003). The means of the reading and math scores were used to form the composite standardized test achievement measure. The variable was standardized as  $z$ -scores separately within each cohort because the test for each grade had a different metric. That is, the cutoff scores for the performance levels differed by grade level, and the same standardized score across grade levels were not equal. The measure was then standardized again among the pooled cohort data.

#### *Classroom Variables*

At the conclusion of the three years of program implementation, teachers in both treatment and control schools were given a questionnaire regarding the social and character education materials used and activities delivered in the classroom, including frequency and duration of implementation (Gottfredson et al., in preparation). The questionnaire completion rate for the sample teachers was higher in treatment schools (59 of 61, 97%) than in control schools (47 of 52, 90%).

An HLM analysis was conducted in order to determine if there were different response rates for the treatment and control groups which could have influenced the portion of the research that treated time-spent in social skills instruction as a dependent variable. Data were missing from seven teachers, five from control schools and two from treatment schools. A dichotomous variable was created indicating the presence of teacher data with the control group as the reference and coded zero, thus treated as a Bernouli case. Treatment was entered at Level 2 to determine whether

treatment status significantly predicted the presence or absence of data. The results indicated that a teacher in a treatment school was about three times more likely to complete the questionnaire than a teacher in a control school, however this difference was not significant and the confidence interval for the log odds was broad and included one (Table 5).

In order to adjust for the missing data in the final analyses, EM imputation was used to impute missing estimates for the seven teachers, a method which resulted in conservative estimates of missing data. In addition, four teachers in the treatment schools seemed to have misinterpreted the intent of the items and were deleted from the analysis: all four teachers reported delivering formal social skills instruction once a day or more, which did not seem like reasonable schedule of formal, manualized social skills instruction in the classroom. The final sample included 113 teachers, 61 of whom were in the treatment schools.

*Frequency.* The responses to the frequency item, “How often do students in your classroom typically receive formal instruction in social skills?”, (Table 6) were used to produce an estimate of the yearly frequency of program implementation, based on a 36 week school year. “Once a day or more” was scaled to 180 times a school year; “About two or three times per week” was scaled to 72 times a school year; “Once per week” was scaled to 36 times a school year; “About every other week” was scaled to 18 times a school year; “About once per month” was scaled to nine times a school year; and “Less than once a month” was scaled to five times a school year.

*Duration.* The responses to the duration item, “When the formal instruction in social skills is delivered, how long is a typical single session?”, (Table 6) were used to produce an estimate of session length: “15 minutes or less” was assessed as eight minutes a session; “16-30 minutes” was assessed as 22 minutes a session; “31-59 minutes” was assessed as 45 minutes a session; and there were no responses in the “60 minutes or more” range. The measure was divided by 60 for an estimate of session duration, in hours.

*Total time.* To construct a variable for total classroom time-spent on social skills instruction, the measure of *frequency* of social skills instruction was multiplied by the measure of *duration* of instruction sessions for a measure of total time, in hours, devoted to program instruction for the school year (i.e., an interaction effect of the *frequency* and *duration* measures).

*Lost academic time.* The responses to item, “Social skills instruction takes time away from academic instruction,” were recoded and standardized, and used as a measure of perceived academic instruction time lost to social skills programming (Table 6). This variable is a measure of teachers’ perception of academic time lost to social skills instruction, not an objective measure of lost academic time.

Although the social skill program instruction *time-spent* variables were discrete and ordinal in nature, an analysis of these variables generally showed that they had monotonic relations with the dependent variables and thus were treated as continuous.

*Aggregate achievement.* Students’ baseline report card achievement measure was aggregated to the classroom-level and standardized, and used as an indicator of

the average academic achievement of the students in the classroom from the year prior. Average classroom achievement will be included at Level 2 for the purpose of statistical control and to increase power. The inclusion of this aggregate variable controls the estimate of *time-spent* effects for average classroom achievement.

### *School Variables*

*Treatment.* Treatment status of the school was a dichotomous variable; control schools were the reference group and coded zero.

### *Analyses*

Students were nested within classrooms and schools and consequently were not independent of one another. Therefore, a multilevel data analysis method, Hierarchical Linear Modeling (HLM; Raudenbush & Bryk, 2002), was used to explore the individual-school and individual-classroom relations.

### *First Research Question*

A two-level model was used to determine whether time-spent on formal social skills program instruction differed as a condition of treatment. To address this research question, four separate equations with different measures of program instruction time as the dependent variable were modeled. The four dependent variables were: *frequency* of social skills program instruction in the classroom (“How often do students in your classroom typically receive formal instruction in social skills?”); *duration* of a typical session of social skills program instruction (“When the formal instruction in social skills is delivered, how long is a typical single session?”); the *total time* of formal program instruction (the interaction effect of *frequency* and

*duration*); and a teacher rating of *lost academic time* to social skills instruction (“Social skills instruction takes time away from academic instruction”). The two-level model included classroom *i* at Level 1 and school *j* at Level 2. Treatment was included at Level 2, as schools were assigned to treatment.

A fully unconditional model was created with no predictors at the classroom or school level. This model was used to partition total variance in the outcome measures into their within- and between-school components. These parameters were estimated in a mixed-level model in which each of the four specific *time-spent* outcome measures of program instruction were specified for separate within-school models.

The four *time-spent* measures were the Level 1 dependent variables; thus, four separate models were created. The Level 1 (classroom) model was as follows:

$$Y_{ij} = \beta_{0j} + r_{ij} \tag{1}$$

Where:

$Y_{ij}$  represents the classroom outcome measure of program instruction for teacher *i* in school *j*:

Model 1: *frequency*,

Model 2: *duration*,

Model 3: *total time*,

Model 4: *lost academic time*;

$\beta_{0j}$  represents the mean program instruction for classrooms in school *j* for:

Model 1: *frequency*,

Model 2: *duration*,

Model 3: *total time*,

Model 4: *lost academic time*;

$r_{ij}$  represents the classroom effect, or the deviation of classroom  $i$ 's instruction time mean from the school mean.

Treatment was added as a school-level variable to create the between-school model. As a dichotomous variable, treatment was uncentered at Level 2, making the control schools the reference group. The Level 2 treatment coefficient in the final estimation of fixed effects of each model indicated the treatment effect for each measure of *time-spent* in social skills instruction. The Level 2 (school) model was as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1ij} + u_{0j} \quad (2)$$

Where:

$\gamma_{01}$  represents the treatment effect, or mean instruction time difference between treatment and control schools for:

Model 1: *frequency*,

Model 2: *duration*,

Model 3: *total time*,

Model 4: *lost academic time*;

$W_{1ij}$  is the treatment indicator for each measure of instruction time;

$u_{0j}$  represents the random school effect, or the deviation of school  $j$ 's mean from the grand mean.

There was a problem in the match of two schools in this study; the school with the largest percent of ethnic minority students (75%) was matched with the school

with the next largest percent of minority students (30%), making the sample non-equivalent despite randomization (Harak, 2008). As a result, historically advantaged students (White and Asian students), were more likely to be assigned to treatment. In order to address the validity threat of sample non-equivalence, a sensitivity analysis was conducted excluding these two unevenly matched schools. It is predicted that sensitivity analyses including only ten schools will yield similar findings to those of the primary analyses.

### *Second Research Question*

The second research question asks whether formalized social skills instruction may result in decreased academic achievement. To answer the second research question, a two-level model was used to determine whether increased *time-spent* on formal program instruction in the classroom led to decreased academic achievement. There were two measures of academic achievement: a variable measuring report card grades and a variable measuring standardized test scores. Separate but similar analyses were conducted for each of these dependent variables. The models represent academic achievement for the student  $i$  (Level 1) within a classroom  $j$  (Level 2).

A fully unconditional model was created with no predictors at any level to partition total variance in the outcome measure of academic achievement across the two levels (classroom and school).

The student-level model was created by adding Level 1 control variables: students' sex, ethnicity, FARM status, and the baseline measure of academic achievement (prior year). The model represents student academic achievement for student  $i$  within classroom  $j$  as a function of the covariates. Student sex, ethnicity,

FARM status, and the prior year achievement variables were considered statistical controls and grand-mean centered. Level 1 slopes were tested to determine if their effects on achievement could be assumed to be fixed or random; those variables with significantly varying slopes were treated as random across classrooms. The student-level model was as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{1..}) + \beta_{2j}(X_{2ij} - \bar{X}_{2..}) + \beta_{3j}(X_{3ij} - \bar{X}_{3..}) + \beta_{4j}(X_{4ij} - \bar{X}_{4..}) + r_{ij} \quad (3)$$

Where:

$Y_{ij}$  represents the student outcome academic achievement measure;

$\beta_{0j}$  represents the covariate-adjusted mean academic achievement for classroom  $j$ ;

$(X_{1ij} - \bar{X}_{1..})$  represents grand-mean centered *sex*;

$(X_{2ij} - \bar{X}_{2..})$  represents grand-mean centered *ethnicity*;

$(X_{3ij} - \bar{X}_{3..})$  represents grand-mean centered *FARM status*;

$(X_{4ij} - \bar{X}_{4..})$  represents grand-mean centered *prior year achievement* for student  $i$  in classroom  $j$ ;

$r_{ij}$ , represents the deviation of student  $ij$ 's score from the predicted score based on the student-level model.

In the classroom-level model, the aggregate baseline achievement measure and the program instruction *time-spent* variables were grand-mean centered. It should be noted that school-aggregated ethnicity and FARMS measures were included in preliminary analyses as school social composition variables, but were deleted from final analyses for parsimony as they did not have significant achievement effects. The

final estimation of the fixed effects of each of the *time-spent* variables indicated the size effect that each variable had on student achievement. The classroom-level equations were as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01j}(W_{1j} - \bar{W}_{1..}) + \gamma_{02j}(W_{2j} - \bar{W}_{2..}) + \gamma_{03j}(W_{3j} - \bar{W}_{3..}) + \gamma_{04j}(W_{(F \times D)j} - \bar{W}_{(F \times D)..}) + \gamma_{05j}(W_{5j} - \bar{W}_{5..}) + u_{0j} \quad (4)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (5)$$

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (6)$$

$$\beta_{3j} = \gamma_{30} + u_{3j} \quad (7)$$

$$\beta_{4j} = \gamma_{40} + u_{4j} \quad (8)$$

$$\beta_{5j} = \gamma_{50} + u_{5j} \quad (9)$$

Where:

$\gamma_{01j}$  represents the effect of mean *prior year achievement* of students in current classrooms;

$(W_{1j} - \bar{W}_{1..})$  represents the grand-mean centered prior year achievement of students in current classrooms;

$\gamma_{02j}$  represents the effect of *frequency* on academic achievement, or mean achievement difference between classrooms;

$(W_{2j} - \bar{W}_{2..})$  represents grand-mean centered frequency;

$\gamma_{03j}$  represents the effect of *duration* on academic achievement, or mean achievement difference between classrooms;

$(W_{3j} - \bar{W}_{3..})$  represents grand-mean centered duration;

$\gamma_{04j}$  represents the effect of *total time* (i.e., the interaction effect of *frequency* and *duration*) on academic achievement, or mean achievement difference between classrooms;

$(W_{(FXD)j} - \bar{W}_{(FXD)..\})$  represents grand-mean centered total time;

$\gamma_{05j}$  represents the effect of mean *lost academic time* on academic achievement, or mean achievement between classrooms;

$(W_{5j} - \bar{W}_{5..})$  represents grand-mean centered lost academic time;

$u_{0j}$  represents the deviation of classroom  $j$ 's Level 1 coefficient,  $\beta_{0j}$ , from its predicted value based on the classroom-level model;

$u_{1j}$ ,  $u_{2j}$ ,  $u_{3j}$ ,  $u_{4j}$ ,  $u_{5j}$  are the error terms for the randomly varying slopes, which were set to zero if the corresponding classroom effect,  $\beta_{pj}$ , was specified as fixed or nonrandomly varying.

## Results

### First Research Question

The intraclass correlations (ICC) for the *time-spent* outcome variables in the unconditional models can be seen in Table 9. This implies that enough of the variance in the *frequency* (56%), the *duration* (14%), the *total time* (65%), and the *lost academic time* measures (9%) lies between schools that it is potentially useful to explore the sources of this variability.

Table 10 displays the summary of treatment effects for each of the *time-spent* variables. Across one year, students in treatment schools received a significant 18 more sessions of formal social skills instruction than students in control schools

(Coefficient = 18.42, SE = 6.98,  $p < .05$ ). The *duration* of formal social skills instruction in treatment schools was about five minutes longer than instruction in control schools, but the difference was not statistically significant (Coefficient = 0.08, SE = 0.05,  $p = .12$ ). The *total time* measure was used as an estimate of total formal skills instruction time across the year, and students in treatment schools had significantly more hours in social skills instruction, about 12, than did students in control schools (Coefficient = 12.45, SE = 5.35,  $p < .05$ ). Ratings by classroom teachers in treatment schools of the extent to which social skills instruction takes time away from academic instruction were about a third of a standard deviation higher than ratings by teachers in control schools and not significant (Coefficient = 0.34, SE = 0.24,  $p = .18$ ).

The sensitivity analysis excluding the two unevenly matched schools, which addressed the validity threat of sample non-equivalence, yielded similar findings to those of the primary analyses (Table 10).

Thus, to answer the first research question, students in the treatment program received significantly and substantially more formal instruction in social skills than students in control schools. Teachers in treatment schools also tended more than did control school teachers to endorse the opinion that social skills instruction took time away from academic instruction, but this difference was not statistically significant.

### *Second Research Question*

In order to determine the most appropriate and parsimonious model to answer the second research question, the classroom-level variables were added sequentially to the second level in order to understand the proportion reduction in the variance of

the student achievement measures between the unconditional model and these separate models. Table 11 displays the proportion of the variance explained between classrooms in the achievement outcome measures by the successive models. Each model explained significantly more of the between-classroom variance of the outcome measures than the previous model (except when *duration* was added to the standardized test scores model). But when the *total time* measure, i.e., the interaction term, was added, the model explained significantly more between-classroom variance of the outcome measures than the previous models. The measure of *academic time lost* did not significantly change the proportion of explained variance between classrooms in the achievement outcomes, regardless of when it was entered into the model, so it was not included in the final model. In other words, the teacher rating of *academic time lost* to social skills instruction did not have a significant incremental effect on student achievement. Thus, the final model used to answer the second research question included the aggregate baseline measure of achievement, *frequency*, *duration*, and *total time* at Level 2. This model explained 52% of the variance between classrooms in the report card grades measure and 58% of the variance between classrooms in the standardized test scores measure (Table 11).

This model indicated an interaction effect of the *frequency* and *duration* measures of time-spent in formal social skills instruction on student achievement; that is, the *total time* measure helped to explain a meaningful portion of the variance between classrooms in the achievement outcomes. Due to multicollinearity between the classroom-level variables *frequency*, *duration*, and *total time*, this model could not meaningfully estimate the effect parameters of these *time-spent* variables on student

achievement. To examine this interaction and to account for the multicollinearity between the *time-spent* measures, the sample was split between teachers who spent 30 minutes or *less* in duration on formal social skills lessons and teachers whom spent *more* than 30 minutes in duration on formal social skills lessons. The following paragraphs describe the results illuminating these interactions, first for report card grades and then for standardized test scores.

*Report card grades outcome.* The intraclass correlations (ICC) for the report card grade achievement outcome in the unconditional model was 0.14 ( $\sigma^2 = 0.86$ ,  $\tau_{00} = 0.14$ ), implying that a significant amount of the variance in this achievement outcome (14%) lies between schools that it is potentially useful to explore the sources of this variability ( $\chi^2 = 398.83$ ,  $df = 112$ ). The effect of the Level 1 covariate baseline achievement on the report card grades outcome was significantly varying and thus the baseline achievement-outcome achievement slope was treated as random across classrooms.

Table 12 presents the results of the within-classroom model for the full sample. Females scored approximately 6% of a standard deviation higher than males ( $p < .05$ ). Historically high achieving students scored approximately 11% of a standard deviation higher than historically low achieving students but the difference was not considered statistically significant. Student's qualifying for free or reduced meals scored approximately 24% of a standard deviation lower than student's not qualifying ( $p < .01$ ). Prior achievement had an effect of approximately three-quarters of a standard deviation on report card grade achievement ( $p < .01$ ).

The results of the final report card grades model are presented in Table 13. To answer the second research question, there was no significant effect of the frequency of social skills lessons on students' report card grades when the lessons were longer than 30 minutes. But when lessons were 30 minutes or less, students receiving 10 more sessions of social skills program instruction scored a significant but very small 6% of a standard deviation lower on the report card grades measure (Coefficient = -0.006,  $SE = 0.003$ ,  $p = .03$ ; the unit of the coefficient is one lesson, 10 lessons are slightly less than one standard deviation).

Thus, controlling for students' sex, ethnicity, FARMS, and baseline achievement, the frequency of social skills program instruction had a very small, negative effect on students' report card grades when lessons were 30 minutes or less, but no significant effect for longer lessons.

*Standardized test score outcome.* The intraclass correlations (ICC) for the standardized test score achievement outcome in the unconditional model was 0.19 ( $\sigma^2 = 0.81$ ,  $\tau_{00} = 0.19$ ), implying that a significant amount of the variance in the achievement outcome (19%) lies between schools that it is potentially useful to explore the sources of this variability ( $\chi^2 = 536.31$ ,  $df = 112$ ). None of the Level 1 covariate-outcome slopes were significantly varying so their effects on standardized test score achievement were treated as fixed across classrooms.

Table 14 presents the results of the within-classroom model for the full sample. Females scored approximately 3% of a standard deviation lower than males and the difference was not statistically significant. Historically high achieving students scored approximately 23% of a standard deviation higher than historically

low achieving students ( $p < .01$ ). Student's qualifying for free or reduced meals scored approximately 22% of a standard deviation lower than student's not qualifying ( $p < .01$ ). Prior achievement had an effect of approximately two-thirds of a standard deviation on report card grade achievement ( $p < .01$ ).

The results of the final standardized test score model are presented in Table 15. To answer the second research question, the frequency of social skills instruction had a very small negative effect on students' standardized test scores, regardless of session duration. When the lessons were longer than 30 minutes, students receiving 10 more sessions of social skills program instruction scored a significant 6% of a standard deviation lower on the standardized test (Coefficient = -0.006,  $SE = 0.002$ ,  $p = .02$ ; the unit of the coefficient is one lesson, 10 lessons are slightly less than one standard deviation). When lessons were 30 minutes or less, students receiving 10 more sessions of social skills program instruction scored a significant 8% of a standard deviation lower on the standardized test score measure (Coefficient = -0.008,  $SE = 0.002$ ,  $p = .00$ ; the unit of the coefficient is one lesson, 10 lessons are slightly less than one standard deviation).

Thus, controlling for students' sex, ethnicity, FARMS, and baseline achievement, the frequency of social skills program instruction had a very small, negative effect on students' standardized test scores.

## *Discussion*

### *Limitations*

A limitation of this research with respect to the first research question was that the confidence intervals for the effect size estimates are moderate in size (Table 10).

Nevertheless, the confidence intervals for most of the effect sizes did not contain the zero value, indicating that the treatment schools spent more time on social skills instruction.

There was a ceiling effect exhibited in the report card grades achievement measure, as the sample was generally high-achieving. In order to adjust for skewness, the report card grades achievement measures were reflected, normalized using a natural log transformation, and standardized, bringing the skewness closer to zero in every case. Although the skewness problem was addressed, there remains the problem of modeling academic change for students with high report card grades as the outcome measure may not adequately capture academic improvement, limiting the variance in the dependent variables and the ability to capture the differential effects of the conditions. The report card grade achievement measure however, seemed to be reliable, with expected bivariate relations with other outcome variables and covariates (Table 4).

There may have been a problem with the measurement of the *duration* and *lost academic time* variables. The correlations between the classroom-aggregated measures of achievement and the measures of the *duration* of social skills sessions, and the *academic time lost* to social skills instruction were all quite small and positive, which does not represent the expected relations given the hypothesis that more time in social skills instruction will decrease achievement (Table 16). The measures of *duration* and *lost academic time* were both trichotomous variables and may not have been sensitive enough to appropriately represent the variance across classrooms. Also, the correlations between the *lost academic time* measure and the

measures of *frequency*, *duration*, and the *total time* were quite small (Table 16), and teachers' ratings of *academic time lost* to social skills programming had no significant effect on student achievement (Table 11). These results could be taken at face value, or a skeptic could attribute the results to the potentially poor construct validity of this measure, as the wording for the *lost academic time* item could either elicit teachers' perceptions of academic time lost, or could elicit teachers' attitudes about social skills instruction, and was not an actual measure of time (Table 6). Thus, future studies would be strengthened by using a more direct measure of instructional time, such as classroom observation.

### *Implications*

The first purpose of this study was to examine the effects of a well-implemented, school-wide social skills intervention program (Second Step) on *time-spent* in formal social skills instruction and academic instruction time for a sample with low rates of social-behavioral problems. The results of this study indicated that students in the treatment program received significantly more formal instruction in social skills than students in control schools. According to teachers' perceptions, social skills instruction took time away from academic instruction, but the treatment-control difference was not statistically significant.

The second purpose of this study was to examine the effects time-spent in social skills programming in the classroom on student report card grades and state-mandated standardized test scores. The results of this study indicated that the *frequency* of social skills lessons had a very small, negative effect on student achievement. It may be that teachers who delivered shorter lessons tended to deliver

more lessons across the school year, and many brief social skills lessons presented in the classroom could result in more transitions between tasks, more classroom disruptions, more intrusions upon academic content, less time-spent in engaged learning, and thus lower student achievement. But the effects of this study were not consistent across measures of lesson duration or measures of achievement, and were so small as to doubt any theoretical and practical significance or implications. Thus, contrary to the hypothesis, increased time spent in social skills programming did not decrease student achievement.

The conclusions of this study are useful for future research and interventions, and have policy implications. Implementation fidelity is important in achieving the goals of universal social competency programs, but fidelity may also translate into more time-spent in social skills instruction. Social skills programming and academic instruction time are variables that can be directly manipulated by policy makers and administrators. Researchers and policy makers should consider how the implementation of school-based social skills programming will affect student achievement, particularly given the current emphasis placed on standardized test scores.

The student sample of this study had low baseline rates of social-behavioral problems across multiple measures and was not necessarily in need of a universal social competency program. There is an interesting conjecture that can be made about the population to which this sample can be generalized; mainly that a universal social skills program that is delivered to a sample that may not need such an intervention may be unsuccessful in further promoting social competence, decreasing problem

behaviors, and increasing achievement. Perhaps there is a threshold of social competence within a school beyond which further increases have no positive effect on student academic achievement.

This is not to say that the school district was mistaken for implementing a social skills program; low baseline rates of problem behaviors do not suggest no action is needed. If the district thought there was a problem, or thought that implementing a social competency program was part of their mission, perhaps it was worthwhile. But decisions about school-based prevention/interventions must consider the goals of such programming, identify the population in need, and determine the level of treatment appropriate for the identified populations (i.e., universal, targeted, intensive). Although social skills instruction did not indirectly decrease student achievement through available academic instruction time, future research should continue to monitor this issue when evaluating the efficacy of school-based social competency interventions.

## Author Note

Joseph F. Nese, Counseling and Personnel Services, University of Maryland.

This research is supported in part by grant number R305L030002A from the Institute of Education Sciences, U.S. Department of Education. Opinions expressed do not necessarily reflect the opinions or policies of the U.S. Department of Education.

The author would like to thank: the School Psychology Research Team – friends and colleagues Elise T. Harak and Allison J. Nebbergall; the Anne Arundel County School System Administration – Lucia Martin and Debbie Wooleyhand; and counselors in each of the research schools. Most special thanks are given to the author's advisor, Gary D. Gottfredson.

The author would also like to thank his partner, Rhonda N. Torke, for her incredible love and energy, and his family, James, Mary Ann, Gina, and Christina Nese – for their love and support.

Correspondence concerning this manuscript should be addressed to Joseph F. Nese, Counseling and Personnel Services, University of Maryland at College Park, College Park, MD 20742; email: [jnese@umd.edu](mailto:jnese@umd.edu).

## Appendix

### Second Step Research Summary

Source	Randomization [Unit]	Second Step Intervention	Time to Measure	Outcome	Measurement	Evaluation
Kelder, et al. (1996)	Randomly assigned schools to condition; N = 8 schools, N = 2,246 students; [school]	Four components: modification of the school environment, a violence-prevention curriculum (Second Step), peer leadership, and parent education.	Pretest, posttest at 2 years	Aggressive behaviors	Teacher Report; Student Report	No significant treatment effect was found between treatment and control groups; a general trend towards a negative treatment effect, Second Step was not analyzed in isolation, so this study provides a poor test of the efficacy of Second Step.
Grossman, et al. (1997)	After matching randomly assigned to each condition; N = 12 schools, N = 790 [school]	Thirty, 35-minute lessons, empathy training, impulse control and anger management, taught by teacher	Pretest, posttest, 6-months after implementation ended	Social Skills; Aggressive behavior	Teacher Report; Parent Report; observation	Scores did not differ significantly between groups; Observations revealed a decrease in physical aggression, increase in prosocial behavior
McMahon, et al. (2000)	1 preschool and 1 kindergarten; N = 109 students [individual]	28 sessions, delivered by teacher	Pretest, posttest	Program knowledge; Behavior problems; Social Skills	Child interviews; Teacher Report; behavioral observations	Child interviews revealed a significant gain of knowledge; observations indicated a significant decrease in disruptive behavior and verbal & physical aggression; the

Source	Randomization [Unit]	Second Step Intervention	Time to Measure	Outcome	Measurement	Evaluation
						teacher ratings revealed no significant effects
McMahon & Washburn (2003)	2 middle schools; N = 156 students, grades 5-8 [individual]	15 sessions, co-delivered by teachers and research team	Pretest, posttest	Aggression ; Prosocial behavior; Empathy, Impulsivity , and Sense of School Membershi p; Program knowledge	Teacher Report; Student Reports	Significant gain in knowledge; increase in prosocial behaviors and empathy; no effect for aggressive behavior
Cooke et al. (2007)	5 schools; N = 741 students, grades 3-4 [individual]	Full program curriculum, delivered by teachers	Pretest, posttest; no comparison group	Aggressive antisocial and prosocial behavior	Student Report; Behavioral observations; discipline referrals	The student reports indicated significant gains in some prosocial behaviors; observations yielded no significant results; not enough data were present to analyze discipline referrals
Frey et al. (2005)	Randomly assigned schools to condition; N = 11 schools; after first year, 4 additional schools were non-randomly	Full program curriculum, delivered by teachers	Pretest, posttest	Prosocial/a ntisocial behavior; Social goals/attrib utions	Teachers Reports; Students Reports; Behavioral observations	Treatment students more likely to prefer prosocial goals; treatment teachers reported greater increases in social competence and greater decreases in antisocial behavior; treatment displayed less aggression than those in

Source	Randomization [Unit] assigned to control; N = 1,253 students; [school]	Second Step Intervention	Time to Measure	Outcome	Measurement	Evaluation the control group
Gottfredson et al (in preparation)	Matched, randomly assigned schools to treatment; N = 12 schools; N = 1,976 students in grades 1-5 [school]	Full program curriculum, delivered by teachers	Pretest, posttest at 3 years; yearly reports	Social competenc y; academic achieveme nt	Parent, Teacher, Student Reports; Behavioral observations; archival records	No support found for prevention of problem behavior or promotion of social competency; positive results found by observations; non-significant, negative program effect on standardized achievement test scores

Table 1.

*Treatment Effects for Second Step on Report Card Grades and Standardized Test Scores*

Variable	Coefficient	SE	df	p
Report Card Grades	-0.12	0.18	10	.52
Standardized Test Scores	-0.20	0.18	10	.30

Table 2.

*Descriptive Statistics*

Variables	Treatment		Control	
	Mean	SD	Mean	SD
<i>Student-level</i>	<i>(n = 939)<sup>a</sup></i>		<i>(n = 785)<sup>b</sup></i>	
Female	0.49	-	0.46	-
Historically High Achieving <sup>c</sup>	0.94	-	0.91	-
FARM Status <sup>d</sup>	0.08	-	0.08	-
Baseline Achievement <sup>e</sup>	2.72	0.95	2.78	0.97
Outcome Report Card Grades <sup>e</sup>	2.99	0.61	3.09	0.60
Outcome Standardized Test Scores	438.26	32.82	445.51	33.05
<i>Classroom-level</i>	<i>(n = 61)</i>		<i>(n = 52)</i>	
Frequency of social skills instruction <sup>f</sup>	37.70	14.97	18.73	13.45
Duration of social skills instruction <sup>g</sup>	0.50	0.21	0.42	0.15
Total time <sup>h</sup>	19.64	13.59	7.63	5.57
Academic instruction time lost <sup>i</sup>	1.17	0.73	0.89	0.83
Baseline Achievement <sup>e</sup>	2.65	0.84	2.81	0.81

<sup>a</sup> Cohort 1 = 33%; Cohort 2 = 34%; Cohort 3 = 33%.

<sup>b</sup> Cohort 1 = 36%; Cohort 2 = 34%; Cohort 3 = 30%.

<sup>c</sup> Historically high achieving students coded 1 (White and Asian); historically low achieving students coded 0 (Black, Latino, and American Indian/Alaskan Native).

<sup>d</sup> Students not qualifying for free/reduced meals coded 0; students qualifying for free of reduced meals coded 1.

<sup>e</sup> Range: 0 – 4.

<sup>f</sup> Range: 5 – 72 sessions across one year.

<sup>g</sup> Range: 0.13 - 0.75 hours per session.

<sup>h</sup> Range: 1.83 – 54 hours across one year.

<sup>i</sup> Range 0 – 2.

Table 3.

*Student Sample Baseline Differences Between Treatment and Control Schools*

Variable	Odds Ratio	CI	df	p
Female <sup>a</sup>	1.08	0.79 – 1.47	10	0.60
Historically High Achieving <sup>b</sup>	1.95	0.28 – 13.77	10	0.47
FARM Status <sup>c</sup>	0.58	0.10 – 3.36	10	0.51
	Coefficient	SE	df	p
Baseline Achievement <sup>d</sup>	-0.15	0.17	10	0.41

<sup>a</sup> Females coded 1, males coded 0.

<sup>b</sup> Historically high achieving students coded 1 (White and Asian); historically low achieving students coded 0 (Black, Latino, and American Indian/Alaskan Native).

<sup>c</sup> Students not qualifying for free/reduced meals coded 0; students qualifying for free or reduced meals coded 1.

<sup>d</sup> Standardized variable.

Table 4.

*Correlations of Student-Level Variable*

	Female	Historically High Achieving	FARMS	Baseline Achievement	Report Card Grades Outcome
Female	-				
Historically High Achieving	-.03	-			
FARMS	-.00	-.43	-		
Baseline Achievement	.08	.18	-.22	-	
Report Card Grades Outcome	.10	.20	-.25	.76	-
Outcome Standardized Test Scores	.03	.22	-.26	.70	.80

*Note.*  $N = 1,774$ ; correlations equal to and above absolute .08 are significantly different from 0 at the .01 level.

Table 5.

*Differential Response Rates for Teacher Questionnaire Between Treatment and Control Groups*

Variable	Odds Ratio	CI	df	p
Teacher data availability	3.14	0.47 – 20.96	10	.21

Table 6.

*Teacher Questionnaire Items about the Time-spent of in Formal Social Skills Instruction in the Classroom*

Measure	Item
Frequency	“How often do students in your classroom typically receive formal instruction in social skills?” <sup>a</sup>
Duration	“When the formal instruction in social skills is delivered, how long is a typical single session?” <sup>b</sup>
Academic time lost	“Social skills instruction takes time away from academic instruction.” <sup>c</sup>

<sup>a</sup> Range: 0 – Once a day or more (recoded 5); 1 – About two or three times per week (recoded 4);

2 – Once per week (recoded 3); 3 – About every other week (recoded 2); 4 – About once per month (recoded 1); and 5 – Less than once a month (recoded 0).

<sup>b</sup> Range: 0 – 15 minutes or less (coded 8); 1 – 16-30 minutes (coded 22); 2 – 31-59 minutes (coded 45); and 3 – 60 minutes or more (coded 75).

<sup>c</sup> Range: 0 – Very true (recoded 2); 1 – Somewhat true (coded 1); 2 – Not true (recoded 0).

Table 7.

*Report Card Items for the Composite Baseline Measure of Academic Achievement*

Subject	Item	Factor Loadings		
		Cohort 1	Cohort 2	Cohort 3
Mathematics	Understands and applies problem-solving processes	.72	.75	.82
Mathematics	Understands and applies concepts	.75	.77	.84
Mathematics	Understands and applies computational processes	.67	.73	.81
Mathematics	Communicates mathematical concepts	.72	.76	.77
Language Arts	Uses reading strategies	.69	.75	.78
Language Arts	Comprehends what is read	.74	.78	.80
Language Arts	Writes effectively	.75	.78	.77

*Note.* Cohort 1 Range: Needs Improvement (coded 0); Progressing in the development of the skill (coded 1); Consistently demonstrating (coded 2). Cohorts 2 and 3 Range: E – Unsatisfactory (recoded 0); D – Below Average (recoded 1); C – Average (recoded 2); B – Good (recoded 3); A – Excellent (recoded 4).

Table 8.  
*Report Card Items for the Outcome Report Card Grades Measure*

Subject	Item	Factor Loadings		
		Cohort 1	Cohort 2	Cohort 3
Mathematics	Understands and applies concepts and computations <sup>a</sup>	.84	.83	.83
Mathematics	Understands and applies problem solving <sup>b</sup>	.85	.83	.83
Reading	Reading level <sup>b</sup>	.72	.77	.74
Reading	Uses reading strategies <sup>b</sup>	.82	.82	.84
Reading	Comprehends a variety of texts <sup>b</sup>	.82	.86	.86
Writing	Communicates effectively <sup>b</sup>	.80	.84	.82

<sup>a</sup> Range: 1 – Does not yet meet grade level expectations (recoded 0); 2 – Meets grade level expectations with assistance (recoded 1); 3 – Consistently meets grade level expectations (recoded 2); and 4 – Exceeds grade level expectations (recoded 3).

<sup>b</sup> Range: E – Unsatisfactory (recoded 0); D – Below Average (recoded 1); C – Average (recoded 2); B – Good (recoded 3); A – Excellent (recoded 4).

Table 9.

*The Intraclass Correlations for the Time-Spent in Formal Social Skills Instruction Outcome Variables*

Dependent Variable	Primary Analysis			Sensitivity Analysis (10 Schools)		
	$\sigma^2$	$\tau_{00}$	ICC	$\sigma^2$	$\tau_{00}$	ICC
Frequency	115.91	214.08	0.65	121.81	255.39	0.68
Duration	0.031	0.005	0.14	0.031	0.006	0.15
Total time	61.29	113.76	0.65	67.31	132.36	0.66
Academic instruction time lost	0.919	0.091	0.09	0.861	0.123	0.13

Table 10.

*Difference between Treatment and Control Schools on Teacher Reports of the Delivery of Formal Social Skills Instruction in the Classroom*

Variable	Primary Analysis				Sensitivity Analysis (10 Schools)			
	Coefficient	<i>SE</i>	<i>df</i>	<i>p</i>	Coefficient	<i>SE</i>	<i>df</i>	<i>p</i>
Frequency of social skills instruction	18.42	6.98	10	.03	20.08	8.33	8	.04
Duration of social skills instruction	0.08	0.05	10	.12	0.10	0.05	8	.09
Total time	12.45	5.35	10	.04	14.17	6.11	8	.05
Academic instruction time lost	0.34	0.24	10	.18	0.39	0.27	8	.19

Table 11.

*Proportion of the Variance Explained Between Classrooms in the Achievement Outcome Measures by the Separate Models*

Model	$\tau_{00}$ , null model between-classroom variance component	Final model between-classroom variance component	Proportion variance explained, between classrooms	$\chi^2$	<i>df</i>
<i>Report Card Grade Outcome</i>					
aggregate achievement	.140	.072	.49	428.84	110
aggregate achievement, <i>frequency</i>	.140	.069	.51	410.26	109
aggregate achievement, <i>frequency</i> , <i>duration</i>	.140	.068	.51	405.56	108
aggregate achievement, <i>frequency</i> , <i>duration</i> , <i>total time</i>	.140	.068	.52	401.64	107
aggregate achievement, <i>frequency</i> , <i>duration</i> , <i>total time</i> , <i>lost time</i>	.140	.068	.52	399.03	106
<i>Standardized Test Score Outcome</i>					
aggregate achievement	.193	.097	.49	520.96	111
aggregate achievement, <i>frequency</i>	.193	.081	.58	454.39	110
aggregate achievement, <i>frequency</i> , <i>duration</i>	.193	.082	.57	453.84	109
aggregate achievement, <i>frequency</i> , <i>duration</i> , <i>total time</i>	.193	.081	.58	447.70	108
aggregate achievement, <i>frequency</i> , <i>duration</i> , <i>total time</i> , <i>lost time</i>	.193	.082	.58	443.67	107

Table 12.

*Within-Classroom Fixed Effects Model for Report Card Grades Outcome*

<i>Fixed Effects</i>		Coefficient	SE
Outcome Achievement Intercept	$\gamma_{00}$	-0.02	0.03
Female-achievement slope	$\gamma_{10}$	0.06*	0.03
Ethnicity-achievement slope	$\gamma_{20}$	0.11	0.06
FARMS-achievement slope	$\gamma_{30}$	-0.24**	0.05
Baseline Achievement-achievement slope	$\gamma_{40}$	0.75**	0.02

\*  $p < .05$ .

\*\*  $p < .01$ .

Table 13.

*Effects of Time-spent in Social Skills Instruction on Report Card Grades*

Model	Coefficient	SE	df	p
<i>30 minutes or less duration</i>				
Frequency	-0.006	0.003	79	.03
Aggregate achievement	-0.09	0.09	79	.31
<i>More than 30 minutes duration</i>				
Frequency	-0.002	0.002	28	.44
Aggregate achievement	0.11	0.11	28	.35

Table 14.

*Within-Classroom Fixed Effects Model for Standardized Test Scores Outcome*

<i>Fixed Effects</i>		Coefficient	SE
Outcome Achievement Intercept	$\gamma_{00}$	-0.02	0.03
Female-achievement slope	$\gamma_{10}$	-0.03	0.03
Ethnicity-achievement slope	$\gamma_{20}$	0.23**	0.07
FARMS-achievement slope	$\gamma_{30}$	-0.22**	0.07
Baseline Achievement-achievement slope	$\gamma_{40}$	0.67**	0.02

\*  $p < .05$ .

\*\*  $p < .01$ .

Table 15.

*Effects of Time-spent in Social Skills Instruction on Standardized Test Scores*

Model	Coefficient	SE	df	p
<i>30 minutes or less duration</i>				
Frequency	-0.008	0.002	79	.00
Aggregate achievement	0.07	0.10	79	.50
<i>More than 30 minutes duration</i>				
Frequency	-0.006	0.002	28	.02
Aggregate achievement	0.04	0.13	28	.73

Table 16.

*Correlations of Classroom-Level Variables and Aggregated Achievement Outcome Measures*

	Frequency	Duration	Total Time	Lost Academic Time	Baseline Achievement	Aggregated Report Card Grades Outcome
Frequency	-					
Duration	.19	-				
Total Time	.84	.64	-			
Lost Academic Time	-.08	.07	-.06	-		
Baseline Achievement	-.06	.05	-.02	.03	-	
Aggregated Report Card Grades Outcome	-.14	.10	-.04	.02	.70	-
Aggregated Standardized Test Score Outcome	-.24	.03	-.15	.14	.68	.83

*Note.*  $N = 113$ ; correlations equal to and above absolute .64 are significantly different from 0 at the .01 level.

## References

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood, NJ: Prentice Hall.
- Battistich, V., Schaps, E., & Wilson, N. (2004). Effects of an elementary school intervention on students' "connectedness" to school and social adjustment during middle school. *The Journal of Primary Prevention, 24*, 243-262.
- Beland, K. (1988). *Second Step, grades 1-3: Summary report*. Seattle, WA: Committee for Children.
- Beland, K. (1989). *Second Step, grades 4-5: Summary report*. Seattle, WA: Committee for Children.
- Beland, K. (1991). *Second Step, preschool-kindergarten: Summary report*. Seattle, WA: Committee for Children.
- Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal, 108*, 115-130.
- Brown, B. W., & Saks, D. H. (1986). Measuring the effects of instructional time on student learning: Evidence from the Beginning Teacher Evaluation Study. *American Journal of Education, 94*, 480-500.
- Cantrell, V. L., & Prinz, R. J. (1985). Multiple perspectives of rejected, neglected, and accepted children: Relation between sociometric status and behavioral characteristics. *Journal of Consulting and Clinical Psychology, 53*, 884-889.
- Carey, J. C., Dimmitt, C., Hatch, T. A., Lapan, R. T., & Whiston, S. C. (2008).

Report of the national panel for evidence-based school counseling: Outcome research coding protocol and evaluation of student success skills and second step. *Professional School Counseling, 11*, 197-206.

Catalano, R. F., Mazza, J. J., Harachi, T. W., Abbott, R. D., Haggerty, K. P., & Fleming, C. B. (2003). Raising healthy children through enhancing social development in elementary school: Results after 1.5 years. *Journal of School Psychology, 41*, 143-164.

Cole, J. D., Lochman, J. E., Terry, R., & Hyman, C. (1992). Predicting early adolescent disorder from childhood aggression and peer rejection. *Journal of Consulting and Clinical Psychology, 60*, 783-792.

Collaborative for Academic, Social, and Emotional Learning (2003). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs.* <http://casel.org>

Collaborative for Academic, Social, and Emotional Learning (2008). *Social and Emotional Learning (SEL) and student benefits: Implications for the safe schools/healthy students core elements.* Chicago: Author.

Committee for Children (2007). Retrieved February 2, 2008 from <http://www.cfchildren.org/programs/ssp/overview/>

Conduct Problems Prevention Research Group (2000). Merging universal and indicated prevention programs: The Fast Track model. *Addictive Behaviors, 25*, 913-927.

Connell, J. P., Spencer, M. B., & Aber, J. L. (1994). Educational risk and resilience in

- African-American youth: Context, self, action, and outcomes in school. *Child Development, 65*, 493-506.
- Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes and development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Consortium on the School-Based Promotion of Social Competence (1994). The school-based promotion of social competence: Theory, research, practice, and policy. In R. J. Haggerty, L. R. Sherrod, N. Garmezy, & M. Rutter (Eds.), *Stress, risk, and resilience in children and adolescents: Processes, mechanisms, and interventions* (pp.268-316). New York, Cambridge University Press.
- Cooke, M. B., Ford, J., Levine, J., Bourke, C., Newell, L., & Lapidus, G. (2007). The effects of city-wide implementation of “Second Step” on elementary school students’ prosocial and aggressive behaviors. *The Journal of Primary Prevention, 28*, 93-115.
- DeVargas, R. C. (1998). A study of Lessons in Character: The effect of moral development curriculum upon moral judgment. Dissertation Abstracts International, 59 (11-A), 4042. (UMI No. 9913706)
- Dietsch, B., Bayha, J. L., & Zheng, H. (2005, April). Short-term effects of a character education program among fourth grade students. Paper presented at the American Educational Research Association annual meeting, Montreal, Quebec, Canada.

- DiPerna, C. J., & Elliott, S. N. (1999). The development and validation of the Academic Competence Evaluation Scales. *Journal of Psychoeducational Assessment, 17*, 207-225.
- Dodge, K. A., Pettit, G. S., McClaskey, C. L., & Brown, J. (1986). Social competence in children. *Monographs of the Society for Research in Child Development, 53*, 344-353.
- DuBois, D.L., Holloway, B.E., Valentine, J.C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*, 157-198.
- Durlak, J. A. (1995). *School-based prevention programs for children and adolescents*. Newbury Park, CA: Sage.
- Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., & Schellinger, K.B. (2008). Enhancing students' social and emotional learning promotes success in school: A meta-analysis. Manuscript submitted for publication. Durlak, J. A. & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25*, 115-152.
- Fisher, C. W., & Berliner, D. C. (1985). *Perspectives on instructional time*. White Plains, NY: Longman Inc.
- Fisher, C. W., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, L. S., & Dishaw, M. M. (1981). Teaching behaviors, academic learning time, and student achievement: An overview. *Journal of Classroom Interaction, 17*, 2-15.
- Flay, B., Acock, A., Vuchinich, S., & Beets, M. (2006). Progress report of the

randomized trial of Positive Action in Hawaii: End of third year of intervention. Available from Positive Action, Inc. 264 4th Avenue South, Twin Falls, ID 83301.

- Flay, B. R., & Allred, C. G. (2003). Long-term effects of the Positive Action program. *American Journal of Healthy Behavior*, 27, 6–21.
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6, 151-175.
- Frey, K. S., Hirschstein, M. K., & Guzzo, B. A. (2000). Second-Step: Preventing aggression by promoting social competence. *Journal of Emotional and Behavioral Disorders*, 8, 102-112.
- Frey, K. S., Nolen, S. B., Van Schoiack-Edstrom, L., & Hirschstein, M. K. (2005). Effects of a school-based social-emotional competence program: Linking children's goals, attributions, and behavior. *Applied Developmental Psychology*, 26, 171-200.
- Frey, K., & Sylvester, L. (1997). Research on the Second Step Program: Do student behaviors and attitudes improve? What do teachers think about the program? Seattle, WA: Committee for Children. (ERIC Document Reproduction Service No. ED426143)
- Gottfredson, G. D. (1981). Schooling and delinquency. In S. E. Martin, L. B. Sechrest, & R. Redner (Eds.), *New directions in the rehabilitation of criminal offenders* (pp. 424-469). Washington, DC: National Academy Press.
- Gottfredson, G. D. (1984). A theory-ridden approach to program evaluation: A

- method for stimulating researcher-implementer collaboration. *American Psychologist*, 39, 1101-1112.
- Gottfredson, D. C., Fink, C. M., Skroban, S., & Gottfredson, G. D. (1997). Making prevention work. In R. P. Weissberg (Ed.), *Issues in children's and families' lives (Vol. 4): Healthy children 2010: Establishing preventive services* (pp. 219-252). Thousand Oaks, CA: Sage.
- Gottfredson, D. C., Gottfredson, G. D., & Skroban, S. (1998). Can prevention work where it is needed most? *Evaluation Review*, 22, 315-340.
- Gottfredson, G. D., et al. (in preparation).
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, 58, 466-474.
- Grossman, D. C., Neckerman, H. J., Koepsell, T. D., Liu, P. Y., Asher, K. N., Beland, K., Frey, K., & Rivara, F. P. (1997). Effectiveness of a violence prevention curriculum among children in elementary school: A randomized controlled trial. *Journal of the American Medical Association*, 277, 1605-1611.
- Haertel, G. D., Walberg, H. J., & Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research*, 53, 75-91.
- Haney, P. & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytical review. *Journal of Clinical Child & Adolescent Psychology*, 27, 424-433.

- Harak, E. T. (2006). *First-year self-report outcomes of a character education experiment with elementary students*. Unpublished master's thesis, University of Maryland, College Park.
- Harak, E. (2008). *Outcomes of an elementary grades social competence experiment according to student self-report*. Unpublished doctoral dissertation, University of Maryland, College Park.
- Hawkins, J. D., Farrington, D. P., & Catalano, R. F. (1998). Reducing violence through the schools. In D. S. Elliot, B. A. Hamburg, & K. R. Williams (Eds.), *Violence in American schools: A new perspective* (pp. 188-216). Cambridge, England: Cambridge University Press.
- Hawkins, J. D., Von Cleve, E., & Catalano, R. F. (1991). Reducing early childhood aggression: Results of a primary prevention program. *Journal of the American Academy of Child and Adolescent Psychiatry*, 30, 208-217.
- Healy, W., & Bronner, A. F. (1936). *New light on delinquency and its treatment*. New Haven, CT: Yale University Press.
- Hill, J.C., Bloom, H.S., Black, A.R., & Lipsey, M.W. (2007). Empirical Benchmarks for Interpreting Effect Sizes in Research. Retrieved December 6, 2007, from <http://www.mdrc.org/publications/459/full.pdf>.
- Hossler, C., Stage, F., & Gallagher, K. (1988). The relationship of increased instructional time to student achievement. Bloomington, IN: Consortium on Educational Policy Studies, School of Education, Indiana University.
- Institute of Educational Sciences (2008). National Center for Educational Research:

Funding Opportunities. Retrieved April 18, 2008 from  
<http://ies.ed.gov/ncer/funding/>

Kelder, S. H., Orpinas, P., McAlister, A., Frankowski, R., Parcel, G. S., & 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.

Kellam, S. G., Brown, C. H., Rubin, B. R., & Ensminger, M. E. (1983). Paths leading to teenage psychiatric symptoms and substance abuse: Developmental epidemiological studies in Woodlawn. In S. B. Guze, F. J. Earls, & J. E. Barrett (Eds.), *Childhood Psychopathology and Development* (pp. 17-52). New York: Raven Press Books.

Kupersmidt, J. B., Coie, J. D., & Dodge, K. A. (1990). The role of poor peer relationships in the development of disorder. In S. R. Asher & J. D. Coie (Eds.), *Peer rejection in childhood. Cambridge studies in social and emotional development* (pp. 274-305). New York: Cambridge University Press.

Larsen, T., & Samdal, O. (2007). Implementing Second Step: Balancing fidelity and program adaptation. *Journal of Educational and Psychological Consultation, 17*, 1-29.

Luria, A. (1961). *The role of speech in the regulation of normal and abnormal behaviors*. New York: Liberright.

Maguin, E., & Loeber, R. (1996). Academic performance and delinquency. In M.

- Tonry (Ed.), *Crime and justice: A review of research* (Vol. 20, pp. 145-164). Chicago: University of Chicago Press.
- Malecki, C. K., & Elliot, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly*, *17*, 1-23.
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*, 153-184.
- Maryland State Department of Education (2003). Retrieved June 19, 2008, from <http://www.marylandpublicschools.org/MSDE/testing/msa/?WBCMODE=PresentationUnpublished%25%25%25%25%25>
- Mayer, G. (1995). Preventing antisocial behavior in the schools. *Journal of Applied Behavior Analysis*, *28*, 476-478.
- McMahon, S. D., & Washburn, J. J. (2003). Violence prevention: An evaluation of program effects with urban African American Students. *The Journal of Primary Prevention*, *24*, 43-62.
- McMahon, S. D., Washburn, J., Felix, E. D., Yakin, J., & Childrey, G. (2000). Violence prevention: Program effects on urban preschool and kindergarten children. *Applied and Preventive Psychology*, *9*, 271-281.
- Moore, B., & Beland, K. (1992). *Evaluation of Second Step, preschool-kindergarten: A violent prevention curriculum kit. Summary report*. Seattle, WA: Committee for Children.
- Munoz, R. F., Mrazek, P. J., & Haggerty, R. J. (1996). Institute of Medicine report on

- prevention of mental disorders: Summary and commentary. *American Psychologist*, 51, 1116–1122.
- National Association for School Psychologists (2006). *School Psychology: A Blueprint for Training and Practice III*. Bethesda, MD.
- Nelson, J. R., Martella, R. M., & Marchand-Martella, N. (2002). Maximizing student learning: The effects of a comprehensive school-based program for preventing problem behaviors. *Journal of Emotional and Behavioral Disorders*, 10, 136-148.
- Orpinas, P., Kelder, S., Frankowski, R., Murray, N., Zhang, Q., & 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.
- Payton, J., Weissberg, R.P., Durlak, J.A., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., & Pachan, M. (2008). The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.
- Raudenbush, S.W.' & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- Rones, M., & Hoagwood, K. (2000). School-based mental health services: a research review. *Clinical Child and Family Psychology Review*, 3, 223-241.
- Skinner, E.A., Wellborn, J.G., & Connell, J.P. (1990). What it takes to do well in

- school and whether I've got it: The role of perceived control in children's engagement and school achievement. *Journal of Educational Psychology*, 82, 22-32.
- Solomon, D., Battistich, V., Watson, M., Schaps, E., & Lewis, C. (2000). A six-district study of educational change: Direct and mediated effects of the child development projects. *Social Psychology of Education*, 4, 3-51.
- Verdugo, R., & Schneider, J. (1999). Quality schools, safe schools: A theoretical and empirical discussion. *Education and Urban Society*, 31, 286-308.
- Weissberg R. P., Kumpfer, K. L., & Seligman, M. E. P. (2003). Prevention that works for children and youth: An introduction. *The American Psychologist*, 58, 425-432.
- Welsh, M., Parke, R. D., Widaman, K., & O'Neil, R. (2001). Linkages between children's social and academic competence: A longitudinal analysis. *Journal of School Psychology*, 39, 463-482.
- Wentzel, K. R. (1991). Relations between social competence and academic achievement in early adolescence. *Child Development*, 62, 1066-1078.
- What Works Clearinghouse. (n.d.). Retrieved December 5, 2008, from <http://ies.ed.gov/ncee/wwc/> - 12/5/2008
- Wilson, S. J., & Lipsey, M. W. (2005). The effectiveness of school-based violence prevention programs for reducing disruptive and aggressive behavior. Revised report for the National Institute of Justice School Violence Prevention Research Planning Meeting, May 2005.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and

disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33, 130-143.