Cluster-Randomized Trial Demonstrating Impact on Academic Achievement of Elementary Social-Emotional Learning

Final version October 20, 2014

David J. Schonfeld [a], Ryan E. Adams [b], Bridget K. Fredstrom [b], Roger P. Weissberg [c], Richard Gilman [b], Charlene Voyce [d], Ricarda Tomlin [d], and Dee Speese-Linehan [e]

[a] St. Christopher's Hospital for Children, 160 East Erie Avenue, Philadelphia, PA 19134; david.schonfeld@drexelmed.edu

[b] Cincinnati Children's Hospital Medical Center, MLC 4002, 3333 Burnet Avenue, Cincinnati, OH 45229-3039; ryan.adams@cchmc.org; bridget.fredstrom@cchmc.org; richard.gilman@cchmc.org

[c] University of Illinois at Chicago, Department of Psychology, 1007 West Harrison St, Chicago, IL 60607-7137; RPW@uic.edu

[d] Yale University School of Medicine, 333 Cedar Street – FMB 14, PO Box 208062, New Haven, CT 06520-8062; charlene.voyce@yale.edu; ricarda.tomlin@yale.edu

[e] Social Development Department, New Haven Public Schools, 80 Hamilton Street, New Haven, CT 06511; deecode1234@comcast.net

Acknowledgments

This study was funded by a research grant (RO1 HD045362: Comprehensive Elementary School Risk Prevention) from the National Institute of Child Health and Human Development (NICHD) awarded to David J Schonfeld, MD.

Abstract

This study evaluated the results of a social and emotional learning (SEL) program on academic achievement among students attending a large, urban, high-risk school district. Using a cluster-randomized design, 24 elementary schools were assigned to receive either the intervention curriculum (*Promoting Alternative Thinking Strategies*, or *PATHS*) or a curriculum that delivered few if any SEL topics (i.e., the control group). In addition to state mastery test scores, demographic data, school attendance, and dosage information were obtained from 705 students who remained in the same group from the 3rd to the 6th grade. Analyses of odds ratios revealed that students enrolled in the intervention schools demonstrated higher levels of basic proficiency in reading, writing, and math at some grade levels. Although these between-group differences held for race/ethnicity, gender, and socio-economic status, significant within-group differences also were noted across these variables. Collectively, these findings indicated that social development instruction may be a promising approach to promote acquisition of academic proficiency, especially among youth attending high-risk school settings. Implications of these findings with respect to SEL programs conclude the paper.

Keywords: Academic achievement; Social-emotional learning; Elementary school; Social development

Cluster-Randomized Trial Demonstrating Impact on Academic Achievement of Elementary Social-Emotional Learning

Learning, especially for children, generally occurs in a social context. As such, children's ability to understand and successfully negotiate interpersonal relationships and conflicts can impact not only their social and emotional adjustment but also their academic achievement (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Welsh, Parke, Widaman, & O'Neil, 2001). In addition, the classroom culture and overall school climate can positively or negatively affect children's readiness to learn, which in turn may influence their academic engagement, effort, and performance. Particularly among elementary school students living in disadvantaged urban communities, the quality of programming to strengthen social and emotional competence serves as a foundation for future developmental milestones, including the capacity to learn at optimal levels (Elias & Haynes, 2008; Hong, 2009; Wentzel, Russell, & Baker, 2014; Wiburg & Brown, 2007). Nevertheless, the actual application and evaluation of such programs in schools is infrequent (Ennett et al., 2003; Forman et al., 2013). As a result, there has been ongoing interest in social and emotional learning programs and the role they may play in promoting academic achievement.

Social and emotional learning (SEL) programs are noted as especially promising given their emphasis on teaching children to (a) identify, label, and understand the emotions that motivate their behaviors; (b) develop empathy and appreciate the interests and needs of others; (c) identify positive solutions to interpersonal conflicts through employing a series of social problem-solving cognitive strategies; and (d) use these social-emotional and cognitive skills to establish and maintain positive interpersonal relationships. Improvements in these competencies vis-à-vis SEL programs have been shown to improve cognitive, affective, and behavioral

competencies such has self-management, social awareness, relationship skills, and responsible decision-making (CASEL, 2008). Nevertheless, few published studies have extended their analyses to examine the impact of SEL programs on learning outcomes. For example, in a recent meta-analysis of 213 studies to investigate the impact of SEL programs on a number of outcomes, only 16% included academic measures (e.g., grades, standardized achievement scores) and many of these studies were based on unpublished and/or non-peer-reviewed data (see Durlak at al., 2011).

There are several plausible reasons why SEL programs may promote learning outcomes. First, SEL instruction teaches children core problem-solving skills; strategies taught to solve interpersonal problems – such as identifying a problem, setting goals to address the problem, generating an appropriate response, and evaluating the outcome – can be applied to evaluate and address academic problems. Second, SEL programs often provide structured classroom activities that enable teachers to improve student-teacher relationships, thus encouraging students to approach them for learning assistance as needed and increasing their interest and engagement in the educational process. Third, teachers who are able to instruct students in effective interpersonal skills and management of interpersonal conflict are, in turn, likely to be more adept at classroom management techniques and have classes whose students are less disruptive and more receptive to the instruction. As a result, these teachers may be able to deliver more complete and effective academic instruction. Thus, it is likely that some combination of enhanced social-emotional competence of students, improved teacher skills and classroom management, and a more supportive classroom environment all contribute to academic achievement (Rimm-Kaufman et al., 2014; Zins, Bloodworth, Weissberg, & Walberg, 2004).

All of these issues may be particularly critical in high-risk schools that are characterized by challenges in one or more of these areas.

Although a number of SEL curricula exist, we chose the *Promoting Alternative Thinking Strategies* (PATHS) (Kusché & Greenberg, 1994) for this study, given that the curriculum is one of the few SEL programs whose efficacy has been demonstrated using randomized, controlled designs (Conduct Problems Prevention Research Group, 1999; 2010; Kam, Greenberg, & Kusché, 2004; Riggs, Greenberg, Kusché, & Pentz, 2006). Relevant to this study, *PATHS* is one of only two programs recognized as a Model Program (the highest possible rating) for use with elementary school-age children by the National Registry of Effective Programs (Substance Abuse and Mental Health Services Administration [SAMHSA], 2003). Controlled trials using the *PATHS* curriculum have demonstrated considerable reductions in externalizing and internalizing problem behavior, peer aggression, conduct problems, and hyperactivity, and improvements in children's emotion regulation, planning and frustration tolerance (Conduct Problems Prevention Research Group, 1999; 2010; Greenberg & Kusché, 1998; Malti, Ribeaud, & Eisner, 2011). Further, these findings were particularly salient for students from at-risk backgrounds – i.e., those with higher baseline levels of problem behaviors and who attend economically disadvantaged schools (Conduct Problems Prevention Research Group, 2010). Finally, one trial of the *PATHS* curriculum with children who were deaf showed significant improvements in reading comprehension beyond that of matched controls (Greenberg & Kusché, 1998), providing preliminary evidence of the program's positive impact on academic achievement.

Using a cluster-randomized, controlled longitudinal design, we tested whether a four-year *PATHS* curriculum would affect student academic achievement in at-risk schools, which

themselves are commonly associated with poor social and emotional skills, student-teacher relationships, classroom and school climate, and school engagement (Greenberg et al. 2003; van de Grift & Houtveen, 2007). The school district targeted in this study historically had established K-12 SEL programming. However, due to budgetary constraints and an increasing emphasis on core academic skills, the implementation of these pre-existing programs was very limited in schools throughout the district. Indeed, pre-intervention interviews with teachers revealed that at the time of the study most students were receiving little or none of these lessons. Thus, our study allowed us to compare the effects of SEL programming in intervention schools against those that essentially had little-to-no overlap with SEL content.

Although our study is one of the first to examine the impact of multi-year SEL programming on academic achievement among young students, several working hypotheses were formulated based on preliminary findings (e.g., Greenberg & Kusché, 1998). First, we hypothesized that those students assigned to the intervention group would be more likely to achieve at least basic proficiency on state mastery tests in reading, writing, and math. Second, we hypothesized that these improvements would be influenced by a dosage effect, that is, the total number of *PATHS* lessons that individual students in the intervention group received over the course of the study. We chose to focus on the impact of advancing academic proficiency at the lowest level (i.e., below basic proficiency) given that the program has been demonstrated to have the greatest impact on behavior and emotion for students most at-risk (Conduct Problems Prevention Research Group, 2010) and the belief that students below basic proficiency might be most vulnerable to the negative impacts of suboptimal social and emotional skills, classroom and school climate, and school engagement. We also examined if any within-intervention group

findings were influenced by select demographic variables, namely race/ethnicity, gender, and socio-economic status.

Method

Participants

The school district in this study was located in a large, urban city in the Northeast. The district serves a predominantly minority population, for whom the median household income is approximately $36,000. All 24 elementary schools within the district were included in the study and divided into two equivalent clusters designed to balance racial/ethnic distribution, percent of students qualifying for a free or reduced price lunch, and school size. Using block randomization procedures (Efird, 2011) to achieve balance in the number of schools and participants per group, we randomly assigned one cluster to the *PATHS* intervention group (n = 692 students; n = 12 schools) and the other to the control group (n = 702 students; n = 12 schools). Sociodemographic data for the whole sample are presented in Table 1. Analyses revealed no significant differences between intervention and control schools for any of these sociodemographic characteristics (all $\chi^2 < 3.605$, p = ns; see Table 1). A comparison between intervention and control schools of state achievement test scores in math, reading, and writing in the year prior to the project was conducted using the percentage of those reaching basic proficiency status at the school during the fourth grade as the outcome variable. An independent sample t-test revealed no differences between intervention and control schools (reading t-value (22) = 0.19, p = .85, d = .08; writing t-value (22) = 0.28, p = .78, d = .12; math t-value (22) = 0.64, p = .53, d = .27).

Our analyses in this study focused on the 705 students who remained in the same group (intervention or control) across all years of the study. This smaller sample, which we will call the study sample, comprised 51% (344 intervention and 361 control) of the original sample. Gender was relatively even across the study sample, with 347 females (49%) and 358 males (51%). We obtained ethnicity information for each participant from the school district, which combined race and ethnicity into one variable (it did not distinguish between non-Hispanic/Latino and Hispanic/Latino Caucasians and other groups). Based on this information, the study sample included 339 African Americans (48%), 286 Hispanic/Latino Americans (41%), 67 Caucasians (9%), and 13 Others (2%). In addition, 68% of the study sample qualified for free or reduced lunch status. The average age at the end of the first year (i.e., 3rd grade) for the study sample was 8.90 (SD = 0.69) years.

The loss of half of the original cohort sample was due to a number of factors, including students leaving the school district, being retained and/or promoted, changing schools to one with a different classification or a non-participating school, or being in a special education classroom in a participating school (see Figure 1). Nevertheless, analyses revealed no significant differences by gender, free or reduced lunch status, or ethnic composition between those students in the study sample versus the larger sample (all $\chi^2 < 4.621$, p = ns), nor was there an age difference (t-value = 0.44, p = ns). Furthermore, there were no significant differences between intervention and control schools in the number of participants lost over the course of the study ($\chi^2 = .409$, p = ns; see Table 2). Finally, there were no significant differences between intervention and control schools for the reasons why participants did not stay in the same school across the study's four years ($\chi^2 = 9.28$, p = ns).

Procedure

Data was collected beginning in the Fall of 2004 (grade 3) through the Spring of 2008 (grade 6). Students attending the control schools received the pre-existing SEL curriculum that was comprised of independent commercially available curricula, such as *Second Step* (Committee on Children, 1992) and *Project Charlie* (*Chemical Abuse Resolution Lies in Education*; Charest, Gwinn, Reinisch, Terrien, & Strabridge, 1987), which are intended to prevent and address an array of problem behaviors and to meet state-mandated instruction in a variety of topics (e.g., substance abuse, violence prevention, bullying, AIDS education, and pregnancy prevention) without additional teacher training or support (see Schonfeld, Adams, Fredstrom, Tomlin, Voyce, & Vaughn, 2012, Table 1 for details of curricula in the comparison SEL curriculum). As noted previously, baseline interviews revealed that few if any of these SEL lessons were being delivered to these children. For the intervention schools, the SEL curriculum was replaced with the *PATHS* curriculum. We collected multiple types of student data, along with bi-weekly teacher reports of how many *PATHS* lessons they taught. Students took the State Mastery Tests in March of each year in which they are reported.

This study was deemed exempt by the first author's institutional review board because it involved research in an educational setting in which comparisons among curricula (social development instruction was already a required curriculum) were being examined, utilizing outcome measures routinely obtained by the school system to monitor the impact of the educational program.

Intervention Group: The *PATHS* Curriculum

Enhanced SEL Curriculum. The *PATHS* curriculum (Kusché & Greenberg, 1994) is a SEL program that provides children in grades K-6 with skills to find positive, nonviolent solutions to social problems. Four broad areas are covered. *Emotional Awareness and*

Understanding topics focus on teaching children to recognize emotions in themselves and others and to take the perspective of others to understand how their behaviors affect others. *Self-Control* teaches youth appropriate self-calming strategies and better ways to express negative emotions. *Interpersonal Problem-Solving Skills* teach youth to identify the problem, set goals, generate responses, and evaluate outcomes. The final curriculum area focuses on *Developing Peer Relations and Enhancing Self-Esteem and Social Responsibility*.

In this study, the *PATHS* curriculum consisted of a varying number of lessons across the project years (45 lessons in grade 3; 42 lessons in grade 4; 39 lessons in grade 5; and 31-44 lessons in grade 6). Because instruction started when the study cohort was in the 3rd grade, the authors of the *PATHS* curriculum modified the lesson sequence to compensate for the students' lack of exposure to the curriculum in grades K-2; essentially, the study curriculum used lessons designed for students one grade lower.

Enhanced teacher training and support. Teachers in the intervention group received annual in-service training averaging 16-20 contact hours per year and in-classroom support (i.e., coaching) throughout the four-year project period. Group training sessions were usually held at the end of the summer just prior to the start of the school year. These training sessions included instruction in background information and curriculum content, participation in curriculum activities, opportunities to practice teaching with peer feedback, and personal reflection. In some situations, teachers were able to practice implementing some of the new curricular material within their class in the year prior to the cohort. In-classroom support and coaching sessions were also conducted annually. These annual in-classroom support visits and coaching sessions consisted of either modeling or co-teaching lessons by a veteran social development facilitator hired by the school system to oversee the project. The amount of support provided to teachers

varied based on their need and receptiveness, and ranged from monthly check-in visits to modeling or co-teaching of lessons in the curriculum (see Table 3).

Level of Programmatic Implementation across Intervention and Control Schools

As expected, there was a large difference between the amount of SEL instruction in the intervention and control schools. During the four-year project period, control school classrooms received an average of 3.5 to 15.9 SEL lessons per year. In general, SEL instruction in the control schools -- if delivered at all -- were provided in a non-uniform fashion across schools and school years. In contrast, the number of *PATHS* lessons taught in the intervention school classes ranged from an average of 25.3 to 31.0 lessons per year.

Due to budgetary constraints during the last decade and an increasing emphasis on core academic skills, training and support related to SEL programming for teachers, and therefore the implementation of the pre-existing SEL curricula in control schools, was very limited throughout the district. In contrast, intervention teachers received substantial training prior to implementing the curriculum, and project staff ensured that all teachers in intervention schools received a make-up training session if they missed the primary session. Intervention schools received extensive monitoring and support from the SEL facilitator responsible for supporting only the teachers in the cohort's grade level (see Table 3); research program staff also provided yearly observations to assess quality of curriculum implementation. This level of monitoring helped to substantiate the accuracy of the teacher checklists of lessons taught.

The pre-existing SEL instruction in control schools was not a single, sequenced curriculum and was implemented in a very limited manner in these schools. In other words, there were various types of instruction provided in a non-uniform fashion across schools and school years. Because of the lack of consistency in the programing across the control schools, we are

unable to compare the impact of the *PATHS* curriculum to a comparison SEL curriculum; rather, the control schools should be viewed more as a non-intervention control group that had limited SEL instruction (see Schonfeld, Adams, Fredstrom, Tomlin, Voyce, & Vaughn, 2012, Table 1 for details of curricula in the comparison SEL curriculum). While this issue was beyond the control of the study, the presence of a district policy supporting SEL programming (in the absence of effective implementation) allowed the inclusion of the full district cohort of students, including the most high-risk students the intervention hoped to impact.

Measures

The primary outcome measure of this study was the State Mastery Test scores which served as an independently administered measure of academic achievement. Control variables included school lunch status as a proxy measure of socio-economic status, as well as race/ethnicity, gender, and attendance. Finally, we created a dosage variable to capture exposure to the intervention for each student in the intervention group.

State Mastery Test (MT). We assessed academic performance using the State Mastery Test (MT). This is a statewide achievement test administered annually in the spring in grades 4-8 to measure students' ability to use problem-solving skills for academic tasks in math, reading, and writing. The MT was created using test development and validation strategies. Several advisory committees were involved in identifying the content, standards, and items for the MT, including members from the State Department of Education and a representative sample of educators from the state's school districts. In addition, an independent company determined that the test items were in agreement with the content and content strands and standards. Correlations between the Metropolitan Achievement Test (MAT) and the MT showed strong evidence for convergent and concurrent validity (reading: 0.79-0.83; math: 0.78-0.82; writing: 0.83-0.86).

Standards for items were also set, including reliability coefficient of at least 0.90, item difficulties greater than or equal to 0.25 and item-total score correlations greater than or equal to 0.20 (Hendrawan & Wibowo, 2013).

The MT report several different types of scores, including raw and scaled scores as well as performance level scores. Scaled scores range from 100-400 for each of the content areas and are transformed into performance level scores for each academic content area (ranging from Below Basic, Basic, Proficient, Goal, and Advanced). We elected to examine these performance level scores, rather than the raw data, since these performance levels reflect common vernacular and understanding of the academic progress of each student. As we sought to determine whether SEL would affect student academic achievement in high-risk students most likely to be vulnerable to the negative effects of poor social and emotional skills, namely those at the lowest level of academic proficiency, participants' performance level scores were divided into two categories: those who had reached at least basic proficiency status and those who were below basic proficiency status (see Table 4).

Gender. We coded gender so that females were compared to males (referent).

Race/Ethnicity. Since the majority of participants were African-American, we coded race/ethnicity into a single variable with African-American as the referent group. Students' race/ethnicity was coded as Hispanic/Latino, Caucasians, Other, and African-American (referent). In preliminary analyses, the small number of those participants in the Other group made it difficult for the models to converge. Thus, those in the Other group had their data for race/ethnicity set to missing.

Lunch Status. For each year of the project, schools were required to gather information on the percentage of students who qualified for a free or reduced price lunch based on federal

guidelines of household income. Participants were coded as qualifying for paid lunch or free/reduced price lunch (referent). This qualification status was used as a marker of lower socio-economic status for each student.

Percent free lunch. For each participant, data were available for each year which reflected the annual average percentage of students eligible for free or reduced-price lunch within the student's school.

Attendance. Attendance data were provided by the school district. Data were collected indicating number of days absent, as well as the total number of days each participating student was enrolled in school. This information was transformed into a yearly assessment of the proportion of days each student was absent. We then created a new variable to reflect the average proportion of days absent from school across all four years.

Percent minority status. At the request of the state's department of education, the school district gathered annual data for each school that included the average percentage of minority students. For each participant, data were available for each year which reflected the average percentage of minority students within their school.

Exposure to the intervention (Dosage). To measure the number of *PATHS* lessons each student experienced in the intervention schools, curriculum checklists were completed by all teachers on a bi-weekly basis in the intervention schools that recorded the number of *PATHS* lessons and activities taught. The total number of lessons taught by each student's teacher was calculated yearly during each year of the intervention. Students' annual number of lessons were then summed and averaged over the course of the project to create an average number of lessons experienced.

Plan of Analysis

To examine the effects of the program, we performed multilevel logistic regressions to account for nested data: i.e., student-level information nested within schools. Thus, each analysis accounted for variability at the level of individual student at level-one. At level-two, individual school was used as the grouping variable to account for variability at the level of the school. For each model, random effects were estimated for the intercepts. Analyses examined intervention effects for MT scores (i.e., below basic proficiency status vs. at least basic proficiency status, given that the PATHS curriculum is anticipated to have the most impact on students who are functioning at the lowest level of academic proficiency which is below basic proficiency) for math, reading, and writing separately during the 4^{th}, 5^{th}, and 6^{th} grades as outcome variables. For each regression, race/ethnicity, gender, student-level lunch status, and attendance were entered as level-one (student-level) predictor variables. These regressions examined group status (intervention vs. control; referent is control) as a level-two, school-level predictor, to test the effectiveness of the intervention. We used group status as a level-two predictor because the intervention was implemented at the level of the school and not the student. In addition, percent minority status and percent free lunch were entered as additional level-two predictors (school-level) variables.

The second set of analyses tested dosage effects of the program by examining the effects of the average number of *PATHS* lessons taught over the course of the intervention as an individual-level predictor. As was the case in the first set of analyses, this set also accounted for the same individual- and school-level control measures listed above. For this set of analyses, only data from those in the intervention group were utilized because the control group was not exposed to any *PATHS* lessons. Only 6^{th}-grade outcome measures (i.e., the final year of the

project) were used since the dosage variable was the cumulative number of lessons taught over the course of the four years of intervention.

All analyses were conducted with Mplus 5.21 (Muthén & Muthén, 1998) and used a full information maximum likelihood (FIML) estimation method, which estimates a likelihood function for each missing value based on the variables that are present. Because the analyses were conducted for only those that remained in their intervention classification over the course of the study, the number of participants included in the analyses across each grade was the same and there were only two participants with missing data across all the variables. The number of participants included in all analyses reported below is 703 for the analyses examining differences between control and intervention schools and 342 for the analyses examining dosage effects. Possible increases in Type I error rates due to multiple analyses were addressed by using false discovery rate (FDR) controlling procedures (Benjamini & Hochberg, 1995). FDR controlling procedures reduce the possible Type II errors that occur when dividing the alpha by the number of analyses as other procedures do when addressing the concern of Type I error resulting from multiple analyses by addressing the expected proportion of false positives across the family of analyses rather than addressing the possibility of one false positive. The FDR procedure provides a q-value, similar to a p-value, but accounts for the number of analyses conducted in a family of analyses, for each individual analysis. The current study reports both *p*- and *q*-values and utilized the standard of $< .05$ for the minimum level at which an individual test may be called significant. Type of outcome for the intervention effects (i.e., reading scores, writing scores, math scores) was used as the grouping for which the q-values were calculated. Finally, all effects are reported as odds ratios (OR). OR represent an association between specific conditions and an outcome, which in the current study is achieving MT proficiency status. Specifically, it describes the odds

that an outcome happens in a given specific condition (e.g. intervention) compared to the odds that the outcome will happen in another separate condition (e.g. control). In terms of the effects of the intervention in the current study, the OR coefficient is the estimated increase in the log odds of reaching proficiency MT status for those exposed to the intervention compared to the odds of those in the control group.

Results

Intervention Effects (Intervention vs. Control)

Reading. For reading MT scores in the 4th grade, the probability of those in the intervention group attaining basic proficiency status was 1.72 times higher than the probability of those in the control group attaining basic proficiency status (63.2% vs. 54.5%, respectively, see Table 4; OR = 1.72, $p < .05$, see Table 5; OR 95% CI = 1.49-1.95). There were no significant differences for 5th or 6th grade scores.

Writing. For writing MT scores, there were no significant differences between the groups during the 4th grade. However, for 5th grade writing scores, the probability of those in the intervention group attaining basic proficiency status was 1.52 times higher than the probability of those in the control group attaining basic proficiency status (91.7% vs. 89.1%, respectively; OR = 1.52, p and $q < .05$; OR 95% CI = 1.26-1.78). In regards to the 6th grade scores, results revealed that the probability of those in the intervention group attaining basic proficiency status was 1.51 times higher than the probability of those in the control group attaining basic proficiency status (91.9% vs. 89.4%, respectively, see Table 4; OR = 1.51, p and $q < .05$, see Table 5; OR 95% CI = 1.24-1.78).

Math. For math MT scores in the 4th grade, the probability of those in the intervention group attaining basic proficiency status was 1.63 times higher than the probability of those in the control group attaining basic proficiency status (87.9% vs. 82.2%, respectively, see Table 4; OR = 1.91, $p < .01$ $q < .05$, see Table 5; OR 95% CI = 1.63-2.19), but there were no significant differences for 5th or 6th grade scores.

Intervention Effects (Dosage)

Reading. For reading mastery status, the number of lessons was a significant predictor of 6th grade basic proficiency. Specifically, the probability of attaining basic proficiency status are increased 1.37 times for each additional lesson taught (OR = 1.37, p and $q < .05$, 95% CI = 1.10-1.64).

Writing. For writing mastery status, the number of lessons did not predict 6th grade basic proficiency (OR = 1.03, p and $q > .49$, 95% CI = .72-1.34).

Math. For math mastery status, the number of lessons was a significant predictor of 6th grade basic proficiency. Specifically, the probability of attaining basic proficiency status is increased 1.29 times for each additional lesson taught (OR = 1.29, p and $q < .05$, 95% CI = 1.02-1.30).

Group Differences in MT Mastery Status

Effects for control variables. As seen in Table 5, there were significant effects on proficiency across the three areas of achievement for the individual-level control variables of race/ethnicity and gender. For reading, writing, and math, Caucasian students were more likely to have basic proficiency status compared to African American students during the 4th, 5th, and 6th grades. Hispanic/Latino students were more likely to have basic proficiency than African American students for writing during the 4th grade and for math during the 6th grade. For reading, females

were more likely to have basic proficiency status than males during the 4th and 5th grade. Females were more likely to have basic proficiency status than males for writing in all three grades. For the school-level control variables, only percentage of those in the school with free or reduced lunch was associated with basic proficiency status for math in the 4th grade. As the percentage of students receiving free or reduced school lunches decreased, the likelihood of having students with basic proficiency status increased. Finally, none of the control measures were found to moderate any of the intervention effects.

Discussion

Research is just beginning to examine the potential contribution of SEL instruction to academic achievement (Durlak et al., 2011). We took advantage of a project (CITATION omitted for masked review) that was designed to explore the efficacy of an evidence-based SEL curriculum on preventing the onset of risk behaviors to demonstrate its impact on academic achievement. Overall, we noted positive intervention effects of the curriculum in at least some grade levels for all three academic content areas. Specifically, the intervention group showed greater basic proficiency in 4th grade reading and math, as well as 5th and 6th grade writing compared to the control group, with the analyses for the dosage effects providing additional support for the intervention effects for reading and math. These findings are consistent with the results of the recent meta-analysis conducted by Durlak and colleagues (2011) that demonstrated a positive impact of SEL instruction in grades K-12. Study findings thus add to a growing body of literature demonstrating that SEL enhances students' connection to school, classroom behavior, and academic achievement (Snyder et al., 2010; Rimm-Kaufman et al., 2014; Zins et al, 2004).

The intervention included not only an evidence-based SEL curriculum, but it also incorporated other intervention elements, including additional training and support to the teachers, that may have contributed to the positive impact on academic achievement seen. The study design did not permit determination of which components of the intervention were responsible for the intervention effects noted.

Overall, the effects were relatively small, but it is our opinion that the effect sizes do not diminish the meaningfulness of the findings. First, considering that the intervention does not directly address the topics of reading, writing, or math as a subject but rather is a Social-Emotional Learning Program that was designed to directly address risky behaviors, the fact that the effects of program was found not only to reduce early sexual behavior (Schonfeld, Adams, Fredstrom, Tomlin, Voyce, & Vaughn, 2012) but also carried over to academic test performance is quite meaningful. Additionally, the outcome here is not at the level of the degree of change in test scores but rather a categorical difference between achieving or not achieving basic proficiency on the test, an outcome that is quite important to school districts, schools, and students in light of the high stakes that these academic tests hold.

There are several other limitations to the current study worth noting. First, in order to increase the likelihood of equality across groups, analyses included only those students who remained in the same classification (intervention vs. control) across all years of the study. In so doing, nearly one-half of the students were lost to attrition over the 4-year intervention period, primarily because of withdrawal from the school district, changes in enrollment to schools not involved in the study, or being retained and failing to progress with the study cohort. There were no significant differences between intervention and control schools concerning reasons why participants did not stay in the same group across the four years, but the attrition rate is still of

concern. While this attrition rate is high, it is comparable to, or even less than, rates that have been seen in other studies of similarly high-risk urban school districts (Conduct Problems Prevention Research Group, 2010). For example, in one analysis involving New York City Schools, another urban school system in the Northeast where the majority of students are Black or Hispanic/Latino, fewer than four out of ten students were sequentially promoted from first to eighth grade; standard academic progress was the exception rather than the rule (Weinstein, Pakes, Donis-Keller, & Schwartz, 2008). In a study involving the *PATHS* curriculum that was conducted in another urban school system in the Northeast (Harrisburg, PA), high family mobility and school district reorganization generated student mobility of approximately 35-40% during just one school year (Kam, Greenberg, & Walls, 2003). For the current study, we required students not only to be continuously and sequentially promoted and remain within the school district, but they also could not change schools during the 4-year time period. Given that prior studies have demonstrated that the greatest impact of the SEL is for those students who were most socially disadvantaged, and the high-risk behaviors these SEL interventions sought to curtail are most prevalent in such high-risk schools, it is critical to conduct studies of the effects of SEL curricula in high-risk school districts in spite of the high mobility characteristic of such systems.

Second, because this study was conducted in one high-risk urban school system in the Northeast and assessed academic proficiency through the state's mastery test results, caution should be used in generalizing these findings to lower risk populations and even to other high-risk urban communities. Future replication with other widely-used measures of academic proficiency is important. Third, we had no true baseline data to control for equivalence between the control and intervention group since the participants did not complete academic testing

during the 2nd (or 3rd) grades. While the groups were randomized, there is no guarantee that they were truly equivalent when the study started.

Given that this study was a first step towards identifying how SEL may influence academic achievement, we acknowledge that there are additional processes that may be involved, including possible mediators to help us explain results, such as specific components of the curriculum including emotion regulation, teacher skills, and classroom management which we were unable to examine separately. Additional limitations that may limit the generalizability of our results include how motivation, on the part of teachers as well as students, and organizational climate of the school are related to the implementation and effectiveness of the SEL curricula. In addition, teachers in the intervention classes on average taught only about two-thirds of available *PATHS* lessons in any given year.

Finally, the study design lacked a non-intervention control group and even the comparison condition received some SEL instruction. The inclusion of all teachers within the intervention schools independent of their personal interest, ability, willingness to teach the SEL curriculum, quality of instructional technique, or degree of completion of the curriculum provides a high degree of external validity, but also produces a conservative bias (Conduct Problems Prevention Research Group, 1999). Furthermore, most of these teachers were only introduced to the *PATHS* curriculum just prior to teaching the lessons to the study cohort. Teachers who were more experienced in teaching the *PATHS* curriculum would likely be better able to model the SEL skills and establish the appropriate classroom environment, rather than to teach just the didactic component of the lessons plans. For this reason, it is likely that the assessment of the intervention during the first year of implementation, although the standard for most evaluation research of this nature, is a very conservative test of its potential impact. To the

extent possible, we attempted to introduce teachers in the intervention schools to the intervention curricula and to encourage them to begin teaching sample lessons one year prior to when they would be teaching the study cohort, but, limited access to teachers for professional development time and very high teacher mobility both compromised our ability to fully achieve this goal. However, the single cohort design of the study was an advantage in that students had a new teacher each year of the project, so high teacher mobility did not affect implementation of the curriculum, but was limited to the early training.

While our study had several identifiable limitations, there were also a number of strengths worth noting. Because we were able to enroll all students in regular or bilingual education throughout the school system, the study does not suffer from enrollment bias. By assessing academic outcomes in a study framed to educators as focusing on reducing risk behaviors, the design also minimized potential expectancy bias, which may have characterized prior studies that relied on course grades or teacher assessments. The outcome measure (the state mastery test) was developed independently of the study and administered independently of the study team. Additionally, the outcome measure was not tied directly to the intervention, thus avoiding any instances in which teachers were directing specific instruction to the outcome measure. Furthermore, the curriculum was not designed specifically to advance academic achievement. Unlike claims from school systems of effective interventions to enhance academic achievement, this curriculum most certainly did not “teach to the test” by such means as rote memorization.

Several additional characteristics set this study apart from others that have examined this topic. The intervention was delivered in an inner-city high-risk school setting where curricula of this nature are most needed and potentially most impactful. The intervention utilized an

independent evaluation of the PATHS curriculum, which has shown promise in such high-risk settings, but has not prior been shown to have a positive impact on academic achievement. Our study utilized a four-year cluster-randomized, controlled longitudinal design, which provides stronger causal assumptions for the effectiveness of the program in both the short- and long-term than could be made in previous studies. In addition, findings above and beyond control variables, as well as analyses that took into account the nested sample design and explored group differences (i.e., intervention versus control) provided a more nuanced and robust picture of the effects of the curriculum.

Many schools are actively restricting classroom time devoted to any subjects or activities that do not appear to directly prepare children for high-stakes testing in reading, writing and arithmetic and teachers are increasingly finding their job performance linked to the degree to which their students demonstrate achievement in these subject areas (Seifer, Gouley, Miller, & Zakriski, 2004). As a result, many important components of children's education, including SEL, are being seriously compromised or eliminated entirely (Durlak et al., 2011; Snyder et al., 2010). Even within the school system where this study was conducted – one that has a long-standing required SEL program – teachers often confided that there were administrative pressures to limit the time spent on non-academic subjects, including the SEL curriculum that was the focus of this study. It is therefore particularly relevant that we have demonstrated that SEL instruction not only does not lower such test scores, but in fact has a significant positive impact on students' achieving basic academic proficiency in all three academic areas studied, at least at some grade levels.

It is becoming increasingly clear that the degree of impact on academic proficiency demonstrated by SEL programs is far from trivial. In a recent meta-analysis, Durlak and

associates (Durlak et al., 2011, p. 24) demonstrated that the effect size for SEL interventions on academic achievement tests "are similar to or, in some cases, higher than those achieved by other types of universal interventions…the post-mean ES [effect size] is comparable to the results of 76 meta-analyses of strictly education interventions (Hill et al., 2007)." If school administrators and educators become aware of the reality that SEL programs may in fact constitute effective educational interventions that not only enhance students' social and emotional skills, classroom behavior, and school attachment (all worthwhile achievements in and of themselves), but also academic proficiency especially for the most disadvantaged students (Rimm-Kaufman et al., 2014), we believe that SEL programs will be more readily implemented and supported.

References

Benjamini, Y.; & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, Series B 57 (1): 289–300.

Charest, P., Gwinn, T., Reinisch, N., Terrien, J., & Strawbridge, C. (1987). *Project Charlie: Chemical Abuse Resolution Lies in Education.* Edina, MN: Storefront/Youth Action.

Collaborative for Academic, Social, and Emotional Learning and National Center for Mental Health Promotion and Youth Violence Prevention (CASEL). (2008). *Social and Emotional Learning (SEL) and Student Benefits: Research Implications for the Safe Schools/Healthy Students Core Elements.* Boston: Education Development Center.

Committee for Children. (1992). *Second Step*. Seattle, WA: Committee for Children.

Conduct Problems Prevention Research Group. (1999). Initial impact of the fast track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology, 67*(5), 648-657. doi: 10.1037/0022-006X.67.5.631.

Conduct Problems Prevention Research Group. (2010). The effects of a multi-year universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology, 78,* 156-168. doi: 10.1037/a0018607.

Durlak, J.A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., & Schellinger, K.B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82,* 474-501. doi: 10.1111/j.1467-8624.2010.01564.x.

Efird, J. (2011). Blocked randomization with randomly selected block sizes. *International Journal of Environmental Research and Public Health, 8*, 15-20. doi: 10.3390/ijerph8010015.

Elias, M. J., & Haynes, N. M. (2008). Social competence, social support, and academic achievement in minority, low-income, urban elementary school children. *School Psychology Quarterly, 23*, 474-495. doi: 10.1037/1045-3830.23.4.474.

Ennett, S. T., Ringwalt, C. L., Thorne, J., Rohrbach, L. A., Vincus, A., Simons-Rudolph, A., et al., (2003). A comparison of current practice in school-based substance use prevention programs with meta-analytic findings. *Prevention Science, 4*, 1-14. doi: 10.1023/A:1021777109369.

Forman, S. G., Shapiro, E. S., Codding, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., et al. (2013). Implementation science and school psychology. *School Psychology Quarterly, 28*, 77-100. doi: 10.1037/spq0000019.

Greenberg, M.T., & Kusché, C.A. (1998). Preventive intervention for school-aged deaf children: The PATHS Curriculum. *Journal of Deaf Studies and Deaf Education, 3,* 49-63. doi: 10.1093/oxfordjournals.deafed.a014340.

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. doi: 10.1037/0003-066X.58.6-7.466.

Hendrawan, I. & Wibowo, A. (2013). The Connecticut Mastery Test: Technical Report. Hartford, CT: Connecticut State Board of Education. Available: http://www.sde.ct.gov/sde/lib/sde/pdf/student_assessment/research_and_technical/Public_2013_CMT_Tech_Report.pdf.

Hong, J. S. (2008). Feasibility of the Olweus Bullying Prevention Program in low-income schools. *Journal of School Violence, 8*, 81-97. doi: 10.1080/15388220802067953.

Kam, C., Greenberg, M.T., & Kusché, C.A. (2004). Sustained effects of the PATHS curriculum on the social and psychological adjustment of children in special education. *Journal of Emotional and Behavioral Disorders, 12*, 66-78. doi: 10.1177/10634266040120020101.

Kam, C., Greenberg, M.T., & Walls, C.T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science*, *4*(1), 55-63. doi: 10.1023/A:1021786811186.

Kusché, C. A., & Greenberg, M. T. (1994). *The PATHS Curriculum.* Seattle, WA: Developmental Research and Programs.

Kusché, C. A., & Greenberg, M. T. (2006). Brain development and social-emotional learning: An introduction for educators. In M. Elias, & H. Arnold (Eds.). *The Educator's Guide to*

Emotional Intelligence and Academic Achievement: Social-Emotional Learning in the Classroom (pp.15-34*).* Thousand Oaks, CA: Corwin.

Malti, T., Ribeaud, D., & Eisner, M. P. (2011). The effectiveness of two universal preventive interventions in reducing children's externalizing behavior: A cluster randomized controlled trial. *Journal of Child Clinical and Adolescent Psychology, 40,* 677-692. doi: 10.1080/15374416.2011.597084.

Muthén, L.K., & Muthén, B.O. (1998). *Mplus User's Guide 6th edition.* Los Angeles, CA: Muthén and Muthén.

Riggs, N.R., Greenberg, M.T., Kusché, C.A., & Pentz, M.A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS Curriculum. *Prevention Science, 7,* 91-102. doi: 10.1007/s11121-005-0022-1.

Rimm-Kaufman, S.E., Larsen, R.A., Baroody, A.E., Curby, T.W., Ko, M., Thomas, J.B., et al. (2014). Efficacy of the *Responsive Classroom* approach: results from a 3-year, longitudinal randomized controlled trial. *American Educational Research Journal,* published online first March 5, 2014. doi: 10.3102/0002831214523821

Schonfeld, D.J., Adams, R.E., Fredstrom, B.K., Tomlin, R., Voyce, C., & Vaughn, L. (2012). Social-emotional learning in grades 3 to 6 and the early onset of sexual behavior. *Sexuality Research and Social Policy,* 9, 178-186. doi: 10.1007/s13178-011-0077-7.

Seifer, R., Gouley, K., Miller, A.L., & Zakriski, A. (2004). Implementation of the PATHS curriculum in an urban elementary school. *Early Education & Development, 15*(4), 471-485. doi: 10.1207/s15566935eed1504_6

Snyder, F.J., Flay, B.R., Vuchinich, S., Acock, A., Washburn, I.J., Beets, M.W., & Li, K. (2010). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: a matched-pair, cluster randomized, controlled trial. *Journal of Research on Educational Effectiveness, 3*(1), 26-55. doi: 10.1080/19345740903353436.

Substance Abuse and Mental Health Administration. (2003). National Registry of Effective Programs. Retrieved November 28, 2011 from http://nrepp.samhsa.gov/ViewIntervention.aspx?id=20.

van de Grift, W., & Houtveen, T. (2007). Weaknesses in underperforming schools. *Journal of Education for Students Placed at Risk, 12,* 383-403. doi: 10.1080/10824660701758942.

Weinstein, M., Pakes, J., Donis-Keller, C., & Schwartz, A.E. (2008). *IESP Policy Brief: From One to Eight: A Longitudinal Portrait of the First Grade Class of 1995-1996.* IESP Brief No. 01-08 (March). New York: New York University.

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. doi: 10.1016/S0022-4405(01)00084-X.

Wentzel, K., Russell, S., & Baker, S. (2014). Peer relationships and positive adjustment at school. In M. J. Furlong, R. Gilman, & E. S. Huebner (Eds.), *Handbook of positive psychology in schools* (2nd Ed.)(pp. 260-277). New York, NY: Routledge.

Wiburg, K., & Brown, K. (2006). *Lesson study communities: Increasing achievement with diverse students*. Thousand Oaks, CA: Corwin Press.

Zins, J.E., Bloodworth, M.R., Weissberg, R.P., & Walberg H.J. (2004). The scientific base linking social and emotional learning to school success. In Zins, J.E., Weissberg, R.P.,

Wang, M.C., & Walberg, H.J. (Eds.), *Building Academic Success on Social and Emotional Learning: What Does the Research Say?* (pp. 3-22). New York, NY: Teachers College Press.

Figure 1.
CONSORT diagram for cluster-randomized controlled trial of SEL

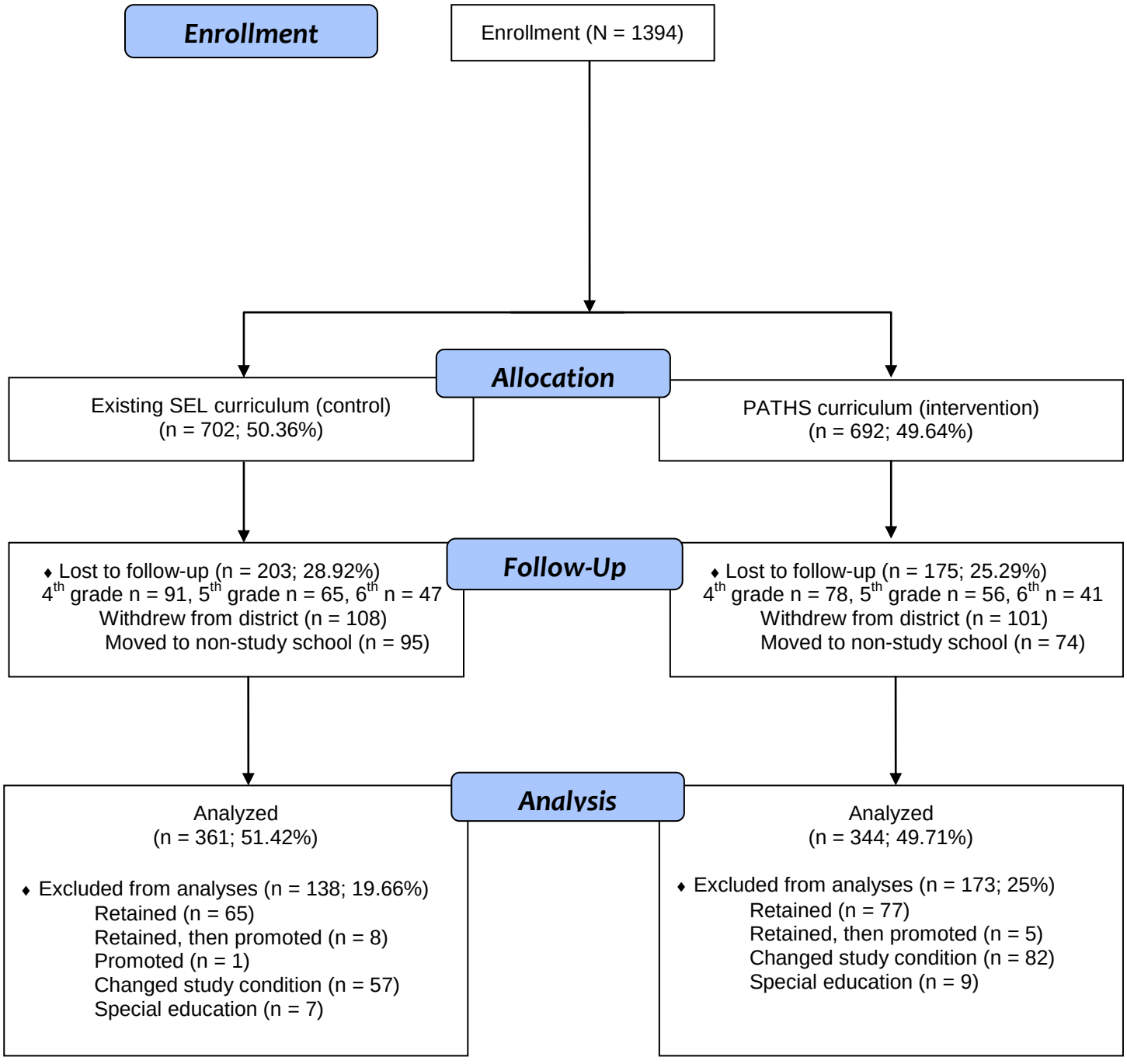

Table 1

Sociodemographics of Whole Sample (N = 1,394) in Year 1

Characteristic	Intervention (n = 692) *n*	%	Control (n = 702) *n*	%	Total (N = 1394) *n*	%
Gender						
Male	355	51%	384	55%	739	53%
Female	337	49%	318	45%	655	47%
Race/Ethnicity						
Black	386	56%	369	53%	755	54%
Hispanic/Latino	240	35%	265	38%	505	36%
Caucasian	55	8%	56	8%	111	8%
Other	11	1%	11	1%	22	2%
Missing			1	< 1%	1	<1%
Free or reduced price lunch	453	66%	466	66%	919	66%
Mean age (*SD)*	8.90	(0.69)	8.92	(0.69)	8.91	(0.71)

Table 2

Sociodemographics for Study Sample (n = 705) at Year 1

	Intervention (n = 344)		Control (n = 361)		Total (n = 705)	
Characteristic	*n*	%	*n*	%	*n*	%
Gender						
Male	161	47%	197	55%	358	51%
Female	183	53%	164	45%	347	49%
Race/Ethnicity						
Black	172	50%	167	46%	339	48%
Hispanic/Latino	132	38%	154	43%	286	41%
Caucasian	35	10%	32	9%	67	9%
Other	5	2%	8	2%	13	2%
Free or reduced price lunch	230	67%	252	70%	482	68%
Participants lost	348	50%	341	49%	689	49%
Mean age (*SD)*	8.90	(0.69)	8.90	(0.69)	8.90	(0.69)

Table 3

Social Development Facilitator Support to Teachers in PATHS Intervention Schools

Grade	# hours training offered	# hours training attended Range	Classroom visits M (Range)	In-depth coaching sessions M (Range)
3rd	16 (+2*)	(2-16)	6.3 (3-17)	1.3 (1-4)
4th	18 (+2*)	(2-18)	16.1 (11-27)	3.1 (1-12)
5th	20 (+7.5*)	(0-20)	21.8 (3-38)	11.8 (0-34)
6th	20 (+9*)	(2-20)	13.3 (5-24)	5.4 (0-18)

Note: *hours of make-up training offered

Table 4

Percentage of Students Reaching Basic Proficiency Status for each Mastery Test Category

	4th grade		5th grade		6th grade	
Category	n	%	n	%	n	%
Reading						
Intervention	215	63.2*	210	61.4	232	68.4
Control	195	54.5	218	60.9	251	68.4
Writing						
Intervention	303	89.9	310	91.7*	308	91.9*
Control	319	89.6	318	89.1	319	89.4
Math						
Intervention	297	87.9*	299	87.4	305	90.5
Control	295	82.2	306	85.7	312	87.6

Note. * indicates the comparison between intervention and controls is significantly different (p <.05); see Table 5 for odds ratios that correspond to the differences.

Table 5
Multilevel Logistic Regression Analyses on Mastery Test (MT) Scores

	4th Grade MT Basic proficiency	5th Grade MT Basic proficiency	6th Grade MT Basic proficiency
	OR	OR	OR
Reading MT			
Individual Level			
Race/Ethnicity (Ref = African American)			
Hispanic/Latino	0.95	1.02	1.15
Caucasian	6.35***	4.49***	10.01***
Gender (Ref = male)	1.37*	1.48**	1.22
Lunch Status (Ref = free lunch)	1.39	1.72	1.69
Attendance	1.01	1.02	1.02
School Level			
% Minority	1.04	1.01	1.00
% Free Lunch	0.94	0.98	0.98
Intervention (Ref = control school)	1.72*	1.06	1.13
Writing MT			
Individual Level			
Race/Ethnicity (Ref = African American)			
Hispanic/Latino	1.78*	0.89	1.38
Caucasian	3.98*	8.87*	3.22*
Gender (Ref = male)	2.17***	1.91**	2.83***
Lunch Status (Ref = free lunch)	1.15	1.71	0.82
Attendance	1.02	1.01	1.02
School Level			
% Minority	0.94	0.99	1.00
% Free Lunch	1.02	0.98	0.95
Intervention (Ref = control school)	0.91	1.52*	1.51*
Math MT			
Individual Level			
Race/Ethnicity (Ref = African American)			
Hispanic/Latino	1.45	1.24	2.12**
Caucasian	2.34*	3.50*	6.93***
Gender (Ref = male)	1.36	1.06	1.03
Lunch Status (Ref = free lunch)	0.88	1.25	1.68
Attendance	1.01	1.00	1.02
School Level			
% Minority	1.00	1.02	0.97
% Free Lunch	0.98*	0.99	1.00
Intervention (Ref = control school)	1.91**	1.21	1.38

Notes. * p and q < .05, ** p and q < .01, ***p and q < .001; OR = Odds ratio