Psychology in the Schools, Vol. 53(2), 2016
View this article online at wileyonlinelibrary.com/journal/pits

DOI: 10.1002/pits.21891

EFFECTS OF A UNIVERSAL POSITIVE CLASSROOM BEHAVIOR PROGRAM ON STUDENT LEARNING

JAMES CLYDE DIPERNA, PUIWA LEI, JILLIAN BELLINGER, AND WEIYI CHENG

The Pennsylvania State University

The purpose of this study was to examine the impact of a universal program to promote positive classroom behavior on students' approaches to learning and early academic skills. Second grade classrooms ($N = 39$) were randomly assigned to treatment and business-as-usual control conditions. Teachers in intervention classrooms implemented the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP) over a 12-week period. Participating students' ($N = 494$) engagement, motivation, and academic skills were assessed before and after treatment implementation. Results indicated that students with lower levels of engagement and motivation at pretest experienced significant improvement in these areas after exposure to the SSIS-CIP. Although no significant differences were observed in reading, students receiving supplemental instructional services demonstrated greater gains in mathematics than did their peers in the control condition.

Supporting students' social development and reducing behaviors that interfere with learning are critical outcomes of the schooling process (Greenberg et al., 2003). Evidence-based universal interventions that support students' social and academic competence represent one promising approach to promoting students' school success (Bradshaw, Zmuda, Kellam, & Ialongo, 2009). Such social–emotional learning programs often include instruction in processing, integrating, and applying social information in a variety of contexts. Through systematic instruction, these skills are introduced, modeled, practiced, and reinforced within classroom settings (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Promoting students' social–emotional development using such approaches has been shown to improve academic achievement and students' motivation to achieve (Greenberg et al., 2003).

A substantial and growing body of research has demonstrated that a moderate positive relationship exists between prosocial behaviors and academic achievement (e.g., Bandura, 1986; Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; DiPerna & Elliott, 2000; Malecki & Elliott, 2002). Socioemotional interventions implemented at the class- and school-wide levels have produced positive effects on students' academic competence. Meta-analyses of universal social–emotional and behavioral interventions have demonstrated small ($d = .33$; Durlak et al., 2011) to moderate ($d = .46$; Sklad, Diekstra, De Ritter, Ben, & Gravesteijn, 2012) effects on students' academic achievement.

Positive social and emotional skills also have been found to support students' capacity to create and sustain positive learning environments (Caprara et al., 2000). Children with interpersonal skills, the capacity for social problem solving, and the ability to enlist academic or social support are more likely to succeed in both the academic and social context of the classroom. In addition, such children are better able to pay attention, work with others, and respect others' views (Caprara et al., 2000). They also are more likely to be organized in their approach to learning, persevere in the face of challenges, and manage stress (Durlak et al., 2011). The promotion of social and emotional behaviors may therefore serve as a protective factor to foster academic learning (Linares et al., 2005).

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A090438 to The Pennsylvania State University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

Correspondence to: James DiPerna, 105 CEDAR Building, The Pennsylvania State University, University Park, PA, 16802. E-mail: jdiperna@psu.edu

Further, as social skills improve class-wide, students' engagement in instruction and learning tends to increase (Sklad et al., 2012).

Approaches to learning (e.g., persistence, emotion regulation, attentiveness, and motivation) reflect the ways in which children engage in classroom interactions and learning activities (Fantuzzo et al., 2007; Li-Grining, Votruba-Drzal, Maldonado-Carreno, & Haas, 2010). These behaviors have been shown to be related to, and potentially facilitate, academic success and socioemotional adjustment (Fantuzzo, Perry, & McDermott, 2004; McDermott, Leigh, & Perry, 2002). Denton and West (2002) found that children who exhibited positive approaches to learning at the start of kindergarten were more than twice as likely to score high on standardized reading and mathematics achievement tests in the spring of first grade. Similarly, teacher ratings of kindergarten students' approaches to learning were positively correlated with their achievement in first grade (George & Greenfield, 2005). Conversely, children who display inattentive behaviors also tend to demonstrate lower motivation and persistence to academic tasks (Fantuzzo, Bulotsky-Shearer, Fusco, & McWayne, 2005).

Several longitudinal studies have suggested that the relationship between approaches to learning and academic competence continues over time (Chen & McNamee, 2011). For example, Li-Grining et al. (2010) examined the relationship between approaches to learning at the start of kindergarten and trajectories of reading and math achievement in kindergarten and first, third, and fifth grade. Results indicated that children with better approaches to learning experienced greater rates of academic growth than did children with less adaptive approaches to learning. Further, early approaches to learning were equally beneficial for children regardless of their race, ethnicity, and socioeconomic status (Li-Grining et al., 2010). Duncan and colleagues (2007) used six longitudinal data sets to examine the relationship between approaches to learning at school entry and later academic achievement. Approaches to learning such as concentration, task persistence, and motivation to learn were consistently related to achievement outcomes.

Research suggests that approaches to learning, including motivation (i.e., approach, persistence, and level of interest regarding academic learning) and engagement (i.e., attention and active participation in classroom activities), also may mediate the relationship between social and academic behaviors (DiPerna, Volpe, & Elliott, 2005). Specifically, early approaches to learning may serve as the foundation for school readiness skills that facilitate later acquisition of more advanced academic skills and concepts (Cunningham & Stanovich, 1997). Children who display positive approaches to learning exhibit greater engagement in, and completion of, academic activities; whereas children who do not use positive approaches may become distracted and give up on tasks more quickly, thereby missing valuable learning opportunities. Guiding children toward positive approaches to learning may increase their self-regulation and executive functioning so that they are able to learn more effectively (Chen & McNamee, 2011).

There are a number of popular evidence-based interventions focused on the promotion of positive classroom behavior during the elementary grades. Such programs include the Promoting Alternative Thinking Strategies (PATHS; Kusche & Greenberg, 1994) program, Good Behavior Game (GBG; Barrish, Saunders, & Wolf, 1969), and Positive Action Program (PA; Flay, Allred, & Ordway, 2001). These universal programs have focused on enhancing protective factors by teaching students an array of social–emotional competencies, such as problem-solving, self-control, emotional awareness, stress management, peer-related social skills, and appropriate school behavior (Conduct Problems Prevention Research Group, 2010; Linares et al., 2005). Implementation of the GBG in conjunction with an enhanced academic curriculum in first grade resulted in higher reading and mathematics scores on a standardized achievement test, reduced utilization of special education services, higher rates of high school graduation, and higher rates of college attendance at age 19 (Bradshaw et al., 2009). The PA program has demonstrated positive effects on students' behavior, school involvement, and academic achievement. Furthermore, behavioral effects were as large or larger in higher risk schools (Flay & Allred, 2003). PATHS also has demonstrated small to moderate effects on children's cognitive concentration (Conduct Problems Prevention Research Group, 2010)

The Social Skills Improvement System Classwide Intervention Program (SSIS-CIP; Elliott & Gresham, 2007) is a universal (Tier 1) program intended to facilitate the development of prosocial behavior that positively impacts learning in the classroom. The SSIS-CIP is a 10-unit curriculum for teachers to use within the general education classroom, and the program uses instructional strategies (e.g., reinforcement, modeling, role-playing, problem-solving) grounded in several established theories of student learning and behavior, such as operant, social learning, and cognitive–behavioral. Unlike other interventions focused primarily on promoting social–emotional competencies (e.g., PATHS, PA), the SSIS focuses on promoting specific social skills related to academic success (e.g., getting along with others, paying attention, asking for help, following rules). The SSIS also takes advantage of situated learning, positive peer models, and systematic instruction, while including components for ongoing monitoring and feedback. (Additional details regarding the SSIS-CIP are provided in the Method section.) Although the SSIS-CIP draws on each of these empirically supported instructional strategies and theories of behavior change, the effect of this universal program on students' classroom learning has not been examined to date. In addition, there are few studies of the academic outcomes associated with universal interventions to promote students' social competence. As such, the primary purpose of this study was to test the hypothesis that the SSIS-CIP improves students' engagement, motivation, and early academic skills in reading and mathematics. In addition, we tested for interactions among student variables (i.e., baseline skill levels, age, sex, race, supplementary services, and special education) and SSIS-CIP exposure on student outcomes (i.e., academic motivation and engagement, on-task behavior, reading, and math).

Method

Participants

Thirty-nine second-grade classrooms participated in the study. Nineteen (49%) of these classrooms were from four elementary schools in a small urban district located in the Mid-Atlantic region of the United States. The remaining classrooms were from two other elementary schools in a small rural district also located in the Mid-Atlantic region. The total classroom sample represented 95% of all second-grade classrooms across the participating schools.

Participating classrooms enrolled 20 to 25 students, and all students were invited to participate in the data collection associated with this study. Approximately 52% of students received parental permission to participate, and the demographic characteristics of the student sample were consistent with the second-grade student population across the six participating elementary schools. As shown in Table 1, participants from the classrooms randomly assigned to the intervention condition included slightly older students and a significantly higher percentage of minority students. Conversely, a larger percentage of participants in the control sample received supplemental services (e.g. Title I support). In addition to the students, 39 teachers (1 per classroom) also participated in the study. All of these teachers were Caucasian, and 79% were female. Most of the teachers reported significant classroom experience ($M = 14.4$ years, $SD = 9$ years). All participants were treated in accord with the ethical principles of the American Psychological Association.

Measures

Academic Competence Evaluation Scales. The Academic Competence Evaluation Scales (ACES; DiPerna & Elliott, 2000) Academic Motivation and Academic Engagement subscales were used as measures of participants' approaches to learning. The Academic Motivation subscale includes 11 items that measure a student's approach, persistence, and level of interest regarding

Table 1
Student Demographic Characteristics by Condition

	SSIS $n = 210$	Control $n = 192$
Age, years[a]	7.42 (.40)	7.29 (.36)
Male	45.24	44.79
White[b]	68.57	81.77
Black/African American[b]	21.05	12.04
Asian	1.91	2.09
Hispanic or Latino[b]	6.70	2.09
Other race	1.44	1.05
Special education consideration	5.24	4.69
Special education	9.52	4.69
Supplementary services[b]	18.57	27.08
Retained in grade from prior year	5.71	3.13
Promoted to next grade	98.57	100

Note. Mean (*SD*) are reported for age. Percentages are reported for all other variables.
[a]Two-level analysis of variance model result significant at .05 level.[b]Chi-square result significant at .05 level.

Table 2
Reliability and Intra-Class Correlation for Measures

	Reliability Index		ICC (School)		ICC (Class)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Teacher Rating[a]						
Academic Motivation	.98	.98	.01	.01	.11	.13
Academic Engagement	.96	.96	.04	.06	.15	.19
Direct Observation[b]						
Active Engaged Time	.94	.92	<.0001	.07	.27	.18
Passive Engaged Time	.90	.90	.16	.12	.19	.10
Academic Achievement						
Math Scaled Score	–	–	.07	.08	.14	.13
Reading Scaled Score	–	–	.15	.13	.21	.19

Note. ICC = intra-class correlation.
[a]Cronbach's alpha; [b]Kappa agreement index.

academic learning. The Academic Engagement subscale includes eight items that reflect attention and active participation in classroom activities. Items are rated using a 5-point format, ranging from *Never* to *Almost Always*. Psychometric evidence for scores from the ACES is strong and consistent with its intended purpose (DiPerna & Elliott, 2000). Reliability estimates based on data from the current sample likewise are strong (Table 2).

Cooperative Learning Observation Code for Kids. The Cooperative Learning Observation Code for Kids (CLOCK; Volpe & DiPerna, 2010) observation protocol was used to conduct independent observations of students' on-task behavior in the classroom setting. The CLOCK includes two categories of student behavior relevant to the primary research questions for this study: Active Engaged Time and Passive Engaged Time. Active Engaged Time reflects when a child is actively attending to an assigned task (e.g., raising hand, asking teacher a relevant question, using a finger to guide reading). Passive Engaged Time refers to times when a child is passively attending to an assigned task (e.g., listening to a teacher talk, looking at the whiteboard, looking at a worksheet). Each of these behaviors is observed using a partial interval format, with each interval lasting 15 seconds. The CLOCK is based on a compilation of codes similar to other empirically supported classroom observation systems, such as the Behavioral Observation System of Young Students (Volpe & Missall, 2007) and the Behavior Observation System for Students (Shapiro, 1996).

Due to the resources required to complete multiple direct observations of student behavior and the minimum sample size necessary to detect a significant difference between conditions, 6 participants (3 boys, 3 girls) were randomly identified as target students for the CLOCK observations within each classroom. Each of these students was observed on three separate occasions during each data collection period. To standardize the observation context and assess social skills in the target setting (i.e., classroom), all observations were completed during mathematics instruction. (Both participating districts used Everyday Math curriculum, which features collaborative learning and discussion.) Each observation was 12 minutes long, and one paired-observation (two raters) was completed per target student and data collection period. Inter-observer agreement for the paired CLOCK observations was high across all target behavior domains and paired observations (Table 2).

Academic Skills. The STAR Reading and Math computerized adaptive tests (Renaissance Learning, 2009, 2010) were used to assess changes in students' reading and math skills. Through vocabulary-in-context test items, STAR Reading requires students to rely on background information, apply vocabulary knowledge, and use active strategies to construct meaning from the assessment text. STAR Math is composed of a series of multiple-choice mathematical problems that assess proficiency with numeration and computation objectives. The STAR Reading and Math assessments were administered before and after intervention implementation. Each assessment took approximately 10 minutes to complete. Overall reliability of the STAR Reading and Math scores is high (.95; Renaissance Learning, 2010). STAR Reading and Math scores demonstrate strong positive correlations with scores on other standardized achievement tests (e.g., California Achievement Test, Pennsylvania System of School Assessment, and Stanford Achievement Test), as well as with teacher ratings of students' proficiency in reading and math skills (Renaissance Learning, 2009, 2010).

Classroom Assessment Scoring System: Kindergarten–Third Grade. The Classroom Assessment Scoring System: Kindergarten–Third Grade (CLASS K-3; Pianta, La Paro, & Hamre, 2008) is a structured observation system developed to assess the overall quality of the classroom instructional environment in the primary grades. Specifically, CLASS K-3 yields scores in three domains: Emotional Support, Classroom Organization, and Instructional Support. These broad domains are further differentiated across 10 dimensions (Positive Climate, Negative Climate, Teacher Sensitivity, Regard for Student Perspective, Behavior Management, Productivity, Instructional Learning Formats, Concept Development, Quality of Feedback, & Language Modeling). Each dimension is rated on a 7-point scale, ranging from *Low* to *High*. Ratings are assigned after an observer completes an observation "cycle" (20 minutes of observation followed by 10 minutes of assigning ratings to dimensions/domains). According to the authors, a minimum of two observation cycles should be completed to yield valid dimension and domain scores.

Psychometric evidence for the CLASS is sound (Mashburn et al., 2008) and provides support for its intended purpose. In the current study, each classroom was observed once (two cycles) during the first data collection window to determine whether there were significant differences in instructional environments across the participating classrooms. Observers were formally trained by a CLASS-certified instructor and achieved the CLASS-mastery criterion (>80% accuracy) before completing observations. Domain scores demonstrated acceptable levels of internal consistency (.81–.93). In addition, paired observations were completed for approximately 50% of the classrooms ($n = 18$), and interrater correlation indices across these paired observations were moderate to high (.56–.76) for the CLASS domain scores.

Procedure

Data Collection. Data were collected as part of a multiyear project, including two separate efficacy trials of the SSIS-CIP in primary classrooms. Data for the present study were drawn from the initial efficacy trial, which featured second grade classrooms. After obtaining approval from the superintendent of each district and principals at each of the participating elementary schools, all second-grade teachers were invited to participate in the project. Letters requesting parental consent for their child's participation in the data collection process were distributed to the parents (or guardians) of each student in the participating second-grade classrooms. Both the business-as-usual and treatment classrooms followed the same data collection schedule. Child-level data were collected during 4-week periods before (November-December) and after (March-April) SSIS-CIP implementation in the classrooms randomly assigned to the "treatment" condition. Specifically, teachers completed the ACES for all participating children from their classrooms. Research data collectors administered the STAR Reading and Math academic assessments to all participating children from their assigned classrooms. Research staff also completed CLOCK observations for the aforementioned randomly selected subsample of participating students (3 boys & 3 girls) from each classroom. Each of these students was observed during mathematics instruction on three separate occasions within each of the pre- and post-data collection periods.

All data collectors ($N = 27$) completed formal training (approximately 12 hours of didactic instruction, practice observations, and individualized feedback) regarding the application and use of the STAR academic assessments and CLOCK observation system. Each data collector also had to meet a mastery criterion (80% accuracy when observing a video of students in an elementary classroom) before they were allowed to conduct classroom observations as part of the project. One third of the CLOCK observations were completed by pairs of observers to ensure reliability (Table 2).

Intervention Implementation. The SSIS-CIP is a brief curriculum intended to improve children's social skills and reduce problem behavior that negatively impacts learning in the classroom (Elliott & Gresham, 2007). The SSIS-CIP includes instructional units focused on 10 key classroom social behaviors that have been identified by teachers as important for classroom success (e.g., following directions, asking for help, ignoring distractions). Each unit focuses on a single skill and includes three scripted lessons, brief video vignettes (30-90s), and practice exercises (student booklets). Each lesson requires approximately 20 to 25 min to complete and relies on six instructional strategies (describe, model, role-play, practice, monitor progress, and generalize) to help children learn the target skill for that unit. Additional information regarding the SSIS-CIP is available in the Instructor's Handbook (Elliott & Gresham, 2008).

Teachers in classrooms randomly assigned to the SSIS-CIP implementation condition ($N = 20$) were formally trained in advance of curriculum implementation. Specifically, the lead author conducted a 1-day workshop with teachers from the implementation condition. During the first half of the workshop, the facilitator provided a detailed overview of the SSIS-CIP curricular materials, including lesson plans, student booklets, and video vignettes. During the second half, teachers then practiced teaching each lesson from the first SSIS-CIP unit in small groups. As teachers practiced, the workshop facilitator provided structured feedback regarding fidelity of their role-play lessons. In addition, teachers had the opportunity to ask questions regarding curricular implementation. After completion of the formal training, implementing teachers were expected to teach one SSIS-CIP unit (three lessons) per week.

Fidelity of SSIS-CIP Implementation. Two complementary methods were used to evaluate and ensure fidelity of implementation of the SSIS-CIP lessons. First, implementing teachers completed weekly standardized checklists indicating their level of implementation (using a 4-point scale, ranging from *Not Implemented* to *Full Implementation*) for the five core components (introduce, define, discuss, identify & practice steps, model/role-play) of each lesson within the unit. In addition, independent observers completed direct observations for approximately 20% of the SSIS-CIP lessons taught by each teacher. Specifically, staff observed the entire lesson and then completed a structured report form that included 20 specific instructional actions/activities. Observers recorded whether or not each was completed during the observed lesson and provided a summative judgment regarding the overall implementation of the five core lesson components using a 4-point scale, ranging from *Not Implemented* (1) to *Full Implementation* (4).

During the implementation period, project staff monitored fidelity (both self-report and independent observations) to ensure that teachers demonstrated at least 90% fidelity in their implementation of the lessons within an instructional unit. If a teacher's implementation fell below the criterion threshold for a unit, a member of the research team contacted the teacher to discuss the area(s) of difficulty, reasons for the difficulty, and what needed to be done differently to achieve the curricular implementation standard. In addition, the research team periodically checked with all teachers (approximately every other week) to see whether they had any implementation questions, make sure no unexpected barriers/difficulties had arisen that would adversely impact their ability to implement the SSIS-CIP lessons, and thank them for their ongoing efforts. As a result of the scripted format of the SSIS-CIP lessons and these monitoring efforts, implementation fidelity was high across all lessons, units, and implementing classrooms based on summative ratings by teachers (98%) and independent observers (97%).

Design and Data Analyses

This study used a multisite cluster randomized trial design to test the efficacy of SSIS-CIP on each of the key outcome variables. Classrooms were randomly assigned to experimental conditions (SSIS-CIP and business-as-usual control) within schools. Multilevel modeling was used to evaluate the effects of SSIS-CIP to account for the nested data structure of students within classrooms within schools. As a result of the modeling complexity attributed to three-level structures, we initially tested the degree to which the schools differ with respect to each of the outcomes of this investigation. These unconditional models yielded intraclass correlation (ICC) coefficients that determined the degree to which the assumption of independence was violated due to the clustering of students in classes in schools (Raudenbush, 1997).

In evaluating effects of SSIS-CIP on each of the outcome measures, we included both student- and class-level predictors to adjust for their effects. Student-level predictors included pretest scores of the respective outcome measure (group-mean centered), students' age (grand-mean centered), reading pretest score (grand-mean centered except for reading outcome), students' sex (1 = male, 0 = female), race (1 = White, 0 = others), whether students received supplementary services (1 = yes, 0 = no), and whether students received some form of special education (1 = yes, 0 = no). The dummy variable predictors were grand-mean centered. Reading pretest score was included as a covariate because it was related to missing data status (i.e., students with missing data had lower average reading pretest scores than did students without missing data). Class-level predictors included grand-mean centered class average of pretest scores of the respective outcome measure. Treatment efficacy was tested using dummy codes for experimental conditions (1 = SSIS-CIP, 0 = control).

Moreover, interaction effects between treatment and pretest scores (both class and student levels), as well as student demographic variables, were tested by adding product terms (between SSIS-CIP and each of these variables) to the model. When the product term was statistically significant at the .05 level, the pattern of interaction was further examined by plotting the adjusted means. Otherwise, nonsignificant product terms were dropped from the final model for parsimony. We estimated multilevel models using the Mixed procedure of SAS (version 9.3) for all outcomes (e.g., Singer, 1998).

In addition, we estimated effect sizes of SSIS-CIP compared with the control (business-as-usual) condition. Specifically, we computed the effect size as a standardized mean difference by dividing the adjusted (for pretest scores and other student- and class-level covariates) group mean difference by the unadjusted pooled within-group student-level standard deviation of the pretest outcome measure. This effect size computation (i.e., using student-level standard deviation to standardize the adjusted difference for Hedges' g) followed the guidelines of What Works Clearinghouse (WWC, n.d., p. 45) for "ES [effect size] computation based on results from HLM [hierarchical linear modeling] analyses in studies with cluster-level assignment." Pooled within-group standard deviation of pretest scores was used because pretest scores were not affected by treatment.

Results

Figure 1 depicts the flow of classroom and student participants throughout the study. Some cases were excluded from the final analyses due to missing demographic data. There were no statistically significant differences on all but one (reading) of the pretest measures between these cases and those retained in the analyses. Reading pretest score was therefore included as a covariate for all models. Students missing age data were substituted with the mean age because all participants were in the same grade level. Given that the percentage of missing data on the outcome variables (1.3%–4.6%) was low and random, cases were deleted listwise for analysis.

Table 2 presents ICCs at both the class and school levels for all outcome measures. Class-level ICCs for posttest outcome measures ranged from moderate (.10 for classroom observations of passive engaged time) to large (.19 for academic engagement and reading scaled score). These levels of ICCs suggested that standard errors might be underestimated if the nested data structure was not taken into account. Therefore, at a minimum, a two-level model was used for each outcome to provide proper standard error estimates.

School-level variances of all posttest outcome measures were statistically nonsignificant based on z-tests (two-tailed ps $>$ .05). However, school-level ICCs for posttest reading (.13) and passive engaged time (.12) were considered medium-sized[1] (Raudenbush, Spybrook, Liu, & Congdon, 2005). We examined both two- and three-level models for these outcome measures and tested their deviance difference based on the final model with the same fixed effects. Because deviance change was statistically nonsignificant for both reading (Δ deviance $= 0.1$, two-tailed $p = .36$) and passive engaged time (Δ deviance $= 0.3$, two-tailed $p = .29$), and the fixed effect estimates and test results were similar between the two- and three-level counterparts, we decided to report the two-level model results for parsimony.

As expected from random assignment at the classroom level, there were no statistically significant differences in CLASS scores between treatment and control classrooms. (CLASS variables initially were included as covariates in the multilevel models but were statistically nonsignificant and therefore removed from the models for parsimony.) Similarly, there were no statistically significant differences between treatment and control conditions (based on two-level models) on any of the pretest measures (Tables 3 and 4).

Student- and class-level means for engagement (teacher ratings & direct observation), academic motivation, reading, and mathematics are reported in Tables 3 and 4. Parameter estimates for the final multilevel model for these variables are presented in Table 5. As expected, student-level pretest

[1] We also conducted a deviance difference test between two- and three-level unconditional models for each outcome measure. All but passive engaged time were statistically nonsignificant at the .05 level.

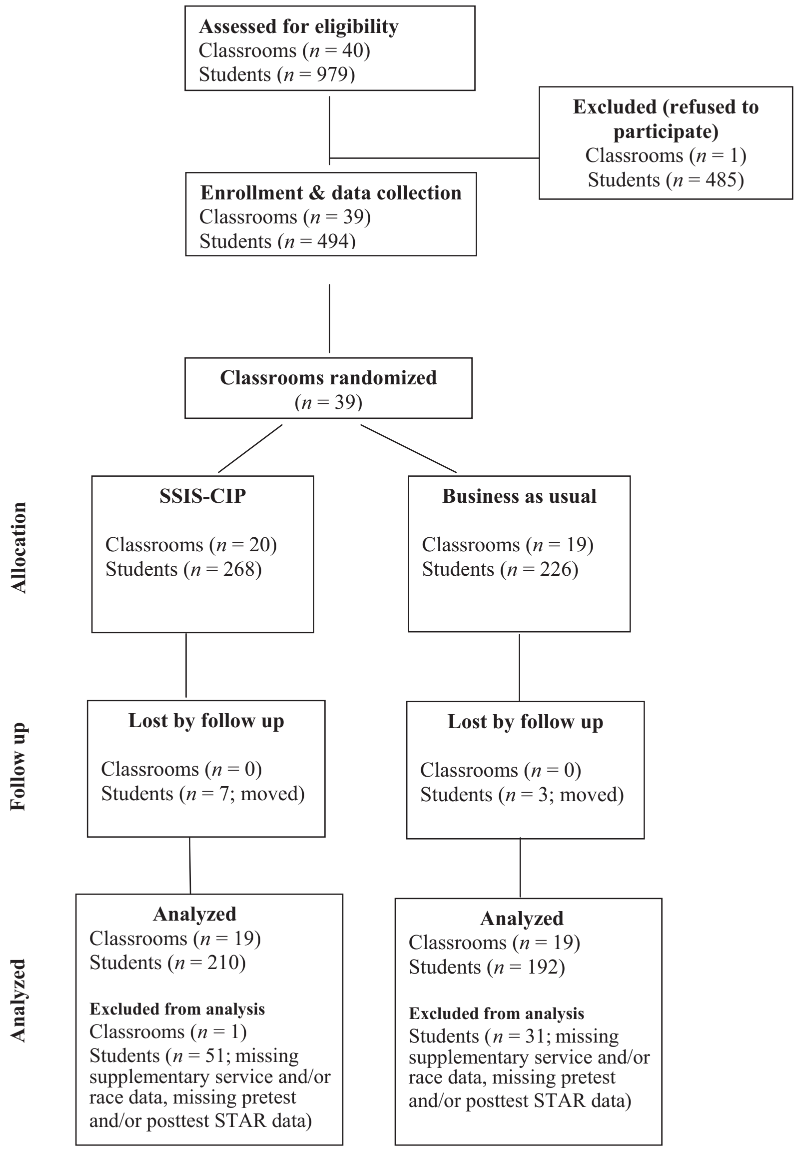

FIGURE 1. SSIS-CIP Cluster-Randomized Trial Participant Flow Chart.

scores were statistically significant predictors for most of the corresponding posttest outcome scores (except for classroom observations of active and passive engaged time). Class-level pretest scores were likewise statistically significant predictors for all the posttest outcome scores. After adjustment for pretest differences, students identified as White received higher average teacher ratings on motivation and reading posttest scores than did students identified as non-White. Male students were observed to spend less active engaged time than did female students, holding pretest differences and the other variables constant. After adjustment for pretest differences, students who received at least one form of supplementary service (e.g., Title I) received lower average teacher ratings on motivation and engagement and scored lower on the reading posttest than did students who did not

Table 3
Student-Level Means (Standard Deviations) for Outcome Measures by Time and Condition

	Pretest		Posttest	
	SSIS-CIP	Control	SSIS-CIP	Control
Teacher Rating[a]				
Academic Motivation	3.58 (1.01)	3.40 (1.05)	3.82 (.92)	3.39 (1.09)
Academic Engagement	3.93 (.89)	3.63 (.93)	4.22 (.80)	3.66 (.99)
Direct Observation[b]				
Active Engaged Time	2.42 (.77)	2.25 (.82)	2.35 (.83)	2.18 (.91)
Passive Engaged Time	2.09 (.76)	2.07 (.78)	2.13 (.82)	2.16 (.83)
Academic Skills				
Math Scaled Score	457.55 (95.89)	439.50 (82.72)	489.84 (105.88)	475.70 (96.30)
Reading Scaled Score	257.80 (135.03)	223.04 (102.73)	301.14 (144.94)	276.59 (124.85)

[a]SSIS-CIP, $N = 210$; control $N = 192$; [b]SSIS-CIP, $N = 102$, control $N = 96$.

Table 4
Class-Level Means (Standard Deviations) for Outcome Measures by Time and Condition

	Pretest		Posttest		Adjusted standardized
Measures	SSIS-CIP	Control	SSIS-CIP	Control	difference[a]
Teacher Rating					
Academic Motivation	3.59 (.44)	3.42 (.53)	3.85 (.35)	3.41 (.51)	.35
Academic Engagement	3.95 (.45)	3.68 (.56)	4.21 (.36)	3.69 (.54)	.35
Direct Observation					
Active Engaged Time	2.39 (.43)	2.14 (.66)	2.33 (.47)	2.15 (.62)	.03
Passive Engaged Time	2.08 (.46)	2.23 (.78)	2.10 (.44)	2.26 (.62)	.01
Academic Skills					
Math scaled score	458.96 (52.49)	438.65 (42.02)	491.96 (61.36)	469.43 (47.11)	<.01
Reading scaled score	260.69 (89.11)	212.71 (63.47)	305.80 (92.28)	264.23 (78.53)	<.01

Note. SSIS-CIP, $N = 19$; control, $N = 19$.
[a]Adjusted for pretest scores and other student- and class-level covariates.

receive any such services. Moreover, students who scored higher on the reading pretest tended to score higher on the mathematics posttest, holding other variables constant.

There was a statistically significant interaction between SSIS-CIP and class-level pretest on teacher ratings of motivation and engagement (Figure 2). The adjusted differences between SSIS-CIP and control classrooms were larger for classes that had lower average pretest scores on these measures. For classes that had high average pretest scores, SSIS-CIP demonstrated no significant difference in improving their average scores when holding other variables constant. Moreover, there was a statistically significant interaction between SSIS-CIP and student-level pretest on teacher ratings of motivation and engagement. Students in the SSIS-CIP condition had slightly higher adjusted posttest means on teacher-rated motivation and engagement than did their peers in the control condition, and that the difference was larger for students who had lower pretest scores (Figure 2). In addition, there was a statistically significant interaction between SSIS-CIP and receipt of supplementary services on mathematics posttest. Holding other variables constant, students who received supplementary services and completed the SSIS-CIP had higher mathematics posttest scores

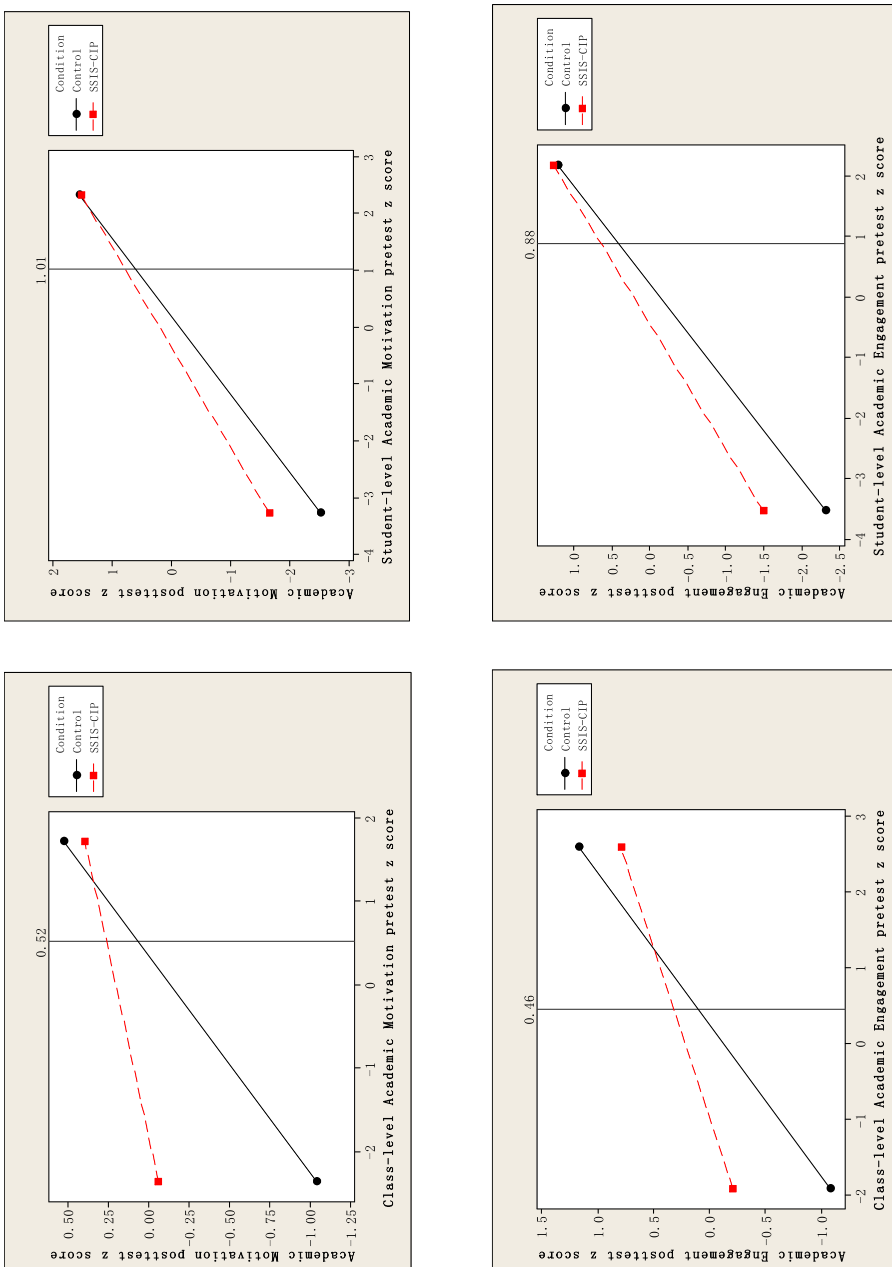

FIGURE 2. Graphs of interactions between SSIS-CIP and pretest score (student- or class-level) on post-test scores. (Area to the left of the vertical line in each graph indicates the region of significant differences between intervention conditions.)

Table 5
Mixed Model Estimates (Standard Errors) for Outcome Measures

	Teacher Rating		Direct Observation		Academic Skills	
	Motivation	Engagement	Active Engaged Time	Passive Engaged Time	Math	Reading
Intercept	3.47**	3.83**	2.27**	2.15**	485.91**	293.20**
	(.07)	(.07)	(.11)	(.10)	(4.79)	(5.97)
Student-levelpretest score	.81**	.73**	.01	–.01	.57**	.82**
	(.05)	(.05)	(.09)	(.10)	(.05)	(.04)
Class-levelpretest score	.86**	.99**	.61**	.59**	1.02**	.94**
	(.13)	(.13)	(.15)	(.13)	(.08)	(.06)
Condition	.35**	.32**	.02	.009	–3.07	–6.74
	(.10)	(.09)	(.16)	(.14)	(6.78)	(8.56)
Student-level pretest score* condition	–.18**	–.16*	NA	NA	NA	NA
	(.06)	(.07)				
Class-level pretest score* condition	–.61**	–.55**	NA	NA	NA	NA
	(.21)	(.20)				
Male	0.05	–0.03	–.26*	–.08	–6.53	–14.74*
	(.06)	(.06)	(.11)	(.11)	(6.50)	(7.14)
White	.18*	–0.0004	–.12	–.04	3.15	23.11*
	(.08)	(.07)	(.14)	(.14)	(7.78)	(8.83)
Supp. services	–.20*	–.20*	–.12	.25	–35.21**	–37.39**
	(.09)	(.08)	(.17)	(.16)	(11.09)	(9.68)
Supp. services* condition	NA	NA	NA	NA	57.37**	NA
					(15.83)	
Special education	–.08	–.15	–.34	.02	–28.78*	–18.16
	(.13)	(.12)	(.27)	(.27)	(13.27)	(14.62)
Age	.08	–.13	.08	–.22	–3.59	–4.85
	(.09)	(.09)	(.17)	(.17)	(10.09)	(10.97)
Reading pretest score	.001	.0004	–.0006	.0003	.22**	–
	(.0004)	(.0003)	(.0006)	(.0006)	(.04)	
Intercept variance	.05**	.04*	.11*	.07	<.0001	139.47
	(.02)	(.02)	(.05)	(.04)		(158.86)
Residual variance	.33**	.31**	.55**	.54**	4102.81**	4873.38**
	(.02)	(.02)	(.06)	(.06)	(293.43)	(365.15)

Note. NA = omitted because not statistically significant at the .05 level; supp. = supplemental.
*$p < .05$. **$p < .01$.

than students who only received supplementary services. However, none of the interaction effects on the observation data or reading posttest scores were statistically significant at the .05 level.

Effect sizes were calculated for the motivation and engagement measures (Table 4) at the mean of their respective pretest scores and controlling for student gender, age, race-ethnicity, supplementary service, and special education status. SSIS-CIP effect sizes on posttest teacher ratings of academic engagement and motivation were .35, which is small to medium, according to Cohen's (1988) criterion but "substantively important" based on WWC's (n.d., p. 60) .25 or greater criterion. However, this effect size should be interpreted with caution because the effect of SSIS-CIP depended on classes' and students' pretest levels of the respective outcomes. Effect sizes on classroom observations were small (≤ .03) for both active and passive engaged time. For academic achievement outcomes, there were no statistically significant SSIS-CIP effects after controlling for achievement pretest score, academic motivation, engagement, receipt of supplementary services, special education, age, gender, and race-ethnicity.

Discussion

The purpose of this study was to examine the impact of the SSIS-CIP on students' approaches to learning and early academic skills. Consistent with the primary hypothesis guiding the study, statistically significant positive differences were observed between SSIS-CIP and control participants on teacher ratings of students' motivation and engagement. Main effect sizes for these differences were in the small to moderate range; however, the observed interactions suggest that the SSIS-CIP demonstrates a more significant impact on students with lower initial levels of motivation and engagement. Similarly, students in classrooms with lower motivation and engagement across all students benefitted more from SSIS-CIP exposure than their peers did in classrooms with higher motivation and engagement.

Despite the observed significant differences between groups based on teacher judgments of engagement, there were no significant differences (main effect or interactions) between SSIS-CIP and control participants on direct observations of engaged time (active or passive). There are two plausible explanations for the differences between the direct observation and teacher report data. First, intervention teachers' perspectives regarding their students' skills may have changed, although their students' behavior actually did not. Second, the differences could be the result of teachers basing their judgments on the universe (all day, every day) of students' engagement behavior in their classroom, whereas the direct observation estimates were based on a much smaller sample of time (three 12-min observations during mathematics instruction per data collection period).

Although we are unable to determine which of these (or other) potential explanations are accurate, it is important to note that studies of universal interventions targeting social skills, behavior, or emotional functioning—even those featuring the most rigorous research designs such as randomized controlled trials (e.g., Cappella et al., 2012, Crean & Johnson, 2013)—often rely exclusively on teacher, student, or parent report of student behavior outcomes. As such, future studies should incorporate observation protocols that more broadly sample student behavior across the school day (and over time) to see if they yield data that are consistent with outcomes measured by behavior rating scales.

With regard to academic outcomes, there were no significant differences between groups in their posttest reading or mathematics skills after controlling for pretest differences. There was one interaction of note, however, relative to academic outcomes. Specifically, students who were receiving supplemental academic services demonstrated positive gains in mathematics relative to their peers who were not receiving such services. This finding is similar to those noted previously for teachers' judgments of engagement and motivation in that students with the greatest need or skill deficits appear to benefit the most from SSIS-CIP exposure.

Although we initially hypothesized that academic skills would change as a result of participation, an important consideration for future studies is how quickly change should be expected in a distal outcome variable (academic skills) for universal interventions targeting classroom behavior. The conceptual change model guiding this study is sequential (positive change in behavior improves engagement and motivation, which in turn allows students to further benefit from instruction and learn more). Based on this model, change in proximal variables must be established (i.e., sustained over time) before any change will be observed in distal variables. Thus, future research regarding academic outcomes from behavior interventions should include assessments of academic outcomes at a later point in time than their proximal outcomes.

Similarly, given that the focus of this study was on students in a primary (second) grade, the measures of academic outcomes focused on the core skill areas of reading and mathematics. Future studies are necessary to examine the effectiveness of the SSIS-CIP not only for older students (intermediate and middle school) but also academic outcomes in content areas, such as social studies, science, and so forth. As the focus of education switches from teaching the academic skills necessary to learn to teaching content, greater expectations are placed on students' ability to learn independently. Thus, constructs such as motivation and active engagement may assume an even more central and significant role in the acquisition of knowledge and skills as students' advance through the educational system, and as a result, universal interventions that promote these outcomes may have a stronger effect on learning.

Based on the results of the current study, the SSIS-CIP yields positive outcomes (from teachers' perspectives) in the areas of academic motivation and engagement for those students with lower initial levels of skill. In addition, students who are most academically at risk (those receiving supplemental services) and participate in the SSIS-CIP demonstrate greater growth in mathematics skills than do their peers. Despite these positive findings, additional studies are necessary to clarify the impact of the SSIS-CIP on academic motivation, engagement, and skills. Such studies should include a broader sample of classrooms and schools from geographic regions other than the Mid-Atlantic. In addition, future studies should incorporate a broader range of observations to ensure that data best represent the universe of student behavior in the classroom. They also should include measures of both short- and long-term academic outcomes to better understand the direct and indirect impact (or lack thereof) of universal classroom behavior programs, such as the SSIS-CIP on students' academic outcomes. In addition, testing the efficacy of the SSIS-CIP version for intermediate and middle school students will provide insight regarding the generalizability of the outcomes from the current study to older students.

References

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. Journal of Social and Clinical Psychology, 4, 359–373.

Barrish, H., Saunders, M., & Wolf, M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. Journal of Applied Behavior Analysis, 2, 119–24.

Bradshaw, C. P., Zmuda, J. H., Kellam, S. G., & Ialongo, N. S. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. Journal of Educational Psychology, 101, 926–937.

Cappella, E., Hamre, B. K., Kim, H. Y., Henry, D. B., Frazier, S. L., Atkins, M. S., & Schoenwald, S. K. (2012). Teacher consultation and coaching within mental health practice: Classroom and child effects in urban elementary schools. Journal of Consulting and Clinical Psychology, 80, 597–610.

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. Psychological Science, 11, 302–306.

Chen, J-Q., & McNamee, G. D. (2011). Positive approaches to learning in the context of preschool classroom activities. Early Childhood Education, 39, 71–78.

Conduct Problems Prevention Research Group. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. Journal of Consulting and Clinical Psychology, 78, 156–168.

Crean, H. F., & Johnson, D. B. (2013). Promoting Alternative Thinking Strategies (PATHS) and elementary school aged children's aggression: Results from a cluster randomized trial. American Journal of Community Psychology, 52, 56–72.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. Developmental Psychology, 33, 934–945.

Denton, K., & West, J. (2002). Children's reading and mathematics achievement in kindergarten and first grade. NCES 2002–125. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

DiPerna, J. C., & Elliott, S. N. (2000). Academic Competence Evaluation Scales. San Antonio, TX: Psychological Corporation.

DiPerna, J. C., Volpe, R. J., & Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in the elementary grades. Journal of School Psychology, 43, 379–392.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. Developmental Psychology, 43, 1428–1446.

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, 405–432.

Elliott, S. N., & Gresham, F. M. (2007). Social Skills Improvement System: Classwide Intervention Program. Minneapolis, MN: Pearson, Assessments.

Elliott, S. N., & Gresham, F. M. (2008). Social Skills Improvement System: Intervention guide. Minneapolis, MN: Pearson, Assessments.

Fantuzzo, J. W., Bulotsky-Shearer, R., Fusco, R. A., & McWayne, C. (2005). An investigation of preschool classroom behavioral adjustment problems and social-emotional school readiness competencies. Early Childhood Research Quarterly, 20, 259–275.

Fantuzzo, J. W., Bulotsky-Shearer, R., McDermott, P., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. School Psychology Review, 36, 44–62.

Fantuzzo, J. W., Perry, M., & McDermott, P. (2004). Preschool approaches to learning and their relationship to other relevant classroom competencies for low-income children. School Psychology Quarterly, 19, 212–230.

Flay, B. R., & Allred, C. G. (2003). Long-term effects of the Positive Action Program. American Journal of Health Behavior, 27, S6–S21.

Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the Positive Action Program on achievement and discipline: Two matched-control comparisons. Prevention Science, 2, 71–89.

George, J., & Greenfield, D. B. (2005). Examination of a structured problem-solving flexibility task for assessing approaches to learning in young children: Relation to teacher ratings and children's achievement. Journal of Applied Developmental Psychology, 26, 69–84.

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.

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

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreno, C., & Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. Developmental Psychology, 46, 1062–1077.

Linares, L. O., Rosbruch, N., Stern, M. B., Edwards, M. E., Walker, G., Abikoff, H. B., & Alvir, J. M. (2005). Developing cognitive-social-emotional competencies to enhance academic learning. Psychology in the Schools, 42, 405–417.

Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. School Psychology Quarterly, 17, 1–23.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. Child Development, 79, 732–749.

McDermott, P. A., Leigh, N. M., & Perry, M. A. (2002). Development and validation of the preschool learning behaviors scale. Psychology in the Schools, 39, 353–365.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). The Classroom Assessment Scoring System Manual, K-3. Baltimore, MD: Brookes Publishing.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. Psychological Methods, 2, 173–185.

Raudenbush, S. W., Spybrook, J., Liu, X., & Congdon, R. (2005). Optimal design for longitudinal and multilevel research: Documentation for the "Optimal Design" software [Computer program] (Version March 09, 2005). Retrieved from http://wtgrantfoundation.org/FocusAreas#tools-for-group-randomized-trials

Renaissance Learning. (2009). STAR Math technical manual. Wisconsin Rapids, WI: Renaissance Learning.

Renaissance Learning. (2010). STAR Reading technical manual. Wisconsin Rapids, WI: Renaissance Learning.

Shapiro, E. S. (1996). Academic skills problems workbook. New York, NY: Guilford.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics, 25, 323–355.

Sklad, M., Diekstra, R., De Ritter, M., Ben, J., & Gravesteijn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? Psychology in the Schools, 49, 892–909.

Volpe, R. J., & DiPerna, J. C. (2010). Cooperative Learning Observation Code for Kids. Unpublished observation code.

Volpe, R. J., & Missall, K. (2007). Behavioral Observation System of Young Students. Unpublished observation code.

What Works Clearinghouse. (n.d.). What Works Clearinghouse: Procedures and standards handbook (version 2.1). Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_1_standards_handbook.pdf