School Psychology Quarterly
2015, Vol. 30, No. 1, 123–141

1045-3830/15/$12.00 http://dx.doi.org/10.1037/spq0000079

Efficacy of the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP) Primary Version

James Clyde DiPerna, Puiwa Lei, Jillian Bellinger, and Weiyi Cheng
The Pennsylvania State University

A multisite cluster randomized trial was conducted to examine the effects of the Social Skills Improvement System Classwide Intervention Program (SSIS-CIP; Elliott & Gresham, 2007) on students' classroom social behavior. The final sample included 432 students across 38 second grade classrooms. Social skills and problem behaviors were measured via the SSIS rating scale for all participants, and direct observations were completed for a subsample of participants within each classroom. Results indicated that the SSIS-CIP demonstrated positive effects on teacher ratings of participants' social skills and internalizing behaviors, with the greatest changes occurring in classrooms with students who exhibited lower skill proficiency prior to implementation. Statistically significant differences were not observed between treatment and control participants on teacher ratings of externalizing problem behaviors or direct observation.

Keywords: social skills, problem behavior, classroom intervention, universal prevention

During the past decade, researchers, policymakers, and practitioners have advocated for a public health approach to service delivery within school settings. Specifically, multitiered models have been developed to promote and remediate key student outcomes associated with the schooling process (Pyle & Vaughn, 2012). Although many initial school-based models were focused on early literacy and reading, educational stakeholders have begun to focus on models targeting other important student outcomes, such as mathematics or classroom behavior, for implementation within school settings (Froiland, 2011; Lembke, Hampton, & Beyers, 2012). Regardless of the target skill or behavior domain, a critical feature of such approaches to service delivery is the implementation of an effective universal (Tier 1) program that promotes the development of the key skill/behavior(s) of interest (Jones, Yssel, & Grant, 2012).

The Social Skills Improvement System Classwide Intervention Program (SSIS-CIP; Elliott & Gresham, 2007) is a universal program intended to facilitate the development of students' prosocial skills and reduce problem behaviors in the classroom. The SSIS-CIP was developed for teachers to use within the general education classroom, and the program utilizes 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 (Elliott & Gresham, 2007). Studies have shown, for example, that teachers who provide frequent reinforcement typically experience increased rates of student on task behavior (Sutherland, Wehby, & Copeland, 2000), increased rates of appropriate behavior from challenging students (Reinke, Lewis-Palmer, & Martin, 2007), and decreased rates of student interruptions (Reinke, Herman, & Stormont, 2013). Further, research suggests there is a positive relationship between teachers' use of praise and their sense of self-

This article was published Online First August 11, 2014.

James Clyde DiPerna, Puiwa Lei, Jillian Bellinger, and Weiyi Cheng, Department of Educational Psychology, Counseling, and Special Education, The Pennsylvania State University.

Jillian Bellinger is now at School of Social Work, University of Texas-Austin.

The research reported here 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 concerning this article should be addressed to James Clyde DiPerna, 105 Cedar Building, The Pennsylvania State University, University Park, PA, 16802. E-mail: jdiperna@psu.edu

efficacy with classroom management (Reinke et al., 2013).

Observation of behaviors modeled by others, in combination with reinforcement and feedback, also plays a significant role in promoting the development of social behaviors. A systematic review of social skills interventions conducted with young children with disabilities indicated that modeling interventions led to positive social outcomes (Cohen's $d = .32–.69$; Vaughn et al., 2003). Modeling may also be paired with role-playing which provides an opportunity to practice a target skill and receive feedback. This can further promote the acquisition and performance of new social behaviors (Elliott & Gresham, 2008). Ideally, practice should occur as often as possible to build fluency and automaticity as well as maintain the behavior over time (Gresham, 2002; Spence, 2003; Vaughn et al., 2003).

In the context of social skills training, cognitive–behavioral approaches emphasize a child's ability to problem solve and self-regulate her or his own behavior (Scarpa & Lorenzi, 2013). Daunic, Smith, Brank, and Penfield (2006) implemented a cognitive–behavioral social problem-solving curriculum with fourth and fifth grade students at risk for behavior problems. Results indicated that students made gains in their knowledge of social problem-solving strategies as well as decreases in reactive and proactive aggressive behaviors. Cognitive–behavioral interventions also have been found to lead to moderate increases in emotion identification among children in primary grades (Whitcomb & Merrell, 2012).

Evidence indicates that multifaceted approaches to social skills training, rather than interventions focused on one teaching modality, lead to the greatest gains in social behavior (Bierman et al., 2010; Spence, 2003). However, many social skill programs fail because they do not allocate adequate time for children to practice newly learned skills (Beelmann, Pfingsten, & Losel, 1994; Bierman et al., 2010). Similarly, promoting the generalization of target behaviors is an important consideration given that social skills demonstrated in a single intervention setting have limited benefit for an individual. As such, instruction and rehearsal within natural contexts are necessary to promote skill generalization (DuPaul & Eckert, 1994). In addition, January, Casey, and Paulson (2011) concluded that classroom social skills programs are most effective when they incorporate structured and active lessons.

There are a number of popular evidence-based programs focused on the promotion of positive classroom behavior during the elementary grades. These include programs such as the Good Behavior Game (GBG; Barrish, Saunders, & Wolf, 1969), Positive Action Program (PA; Flay, Allred, & Ordway, 2001), and Promoting Alternative Thinking Strategies (PATHS; Kusche & Greenberg, 1994). Target outcomes for each of these programs range from primarily reducing disruption and aggression (GBG) to promoting both social and emotional awareness (e.g., PATHS). The SSIS-CIP differs from these programs in that it targets common social behaviors identified by a nationally representative sample of teachers as important for classroom success (Elliott & Gresham, 2007). Similarly, the theoretical foundations for these programs range along a continuum from behavioral (GBG) to social learning (SSIS-CIP) to cognitive–behavioral-affective-dynamic (PATHS). In addition, although four of the five programs utilize structured lessons, the number of lessons varies widely across programs with the SSIS-CIP utilizing significantly fewer lessons (30) than the other three programs (72–180).

Perhaps most importantly, unlike the GBG (e.g., Bradshaw, Zmuda, Kellam, & Ialongo, 2009; Cappella et al., 2012; Lannie & McCurdy, 2007), PATHS (Conduct Problems Prevention Research Group, 1999; Domitrovich, Cortes, & Greenberg, 2007; Greenberg, Kusche, Cook, & Quamma, 1995), and PA (e.g., Flay et al., 2001; Li et al., 2011; Washburn et al., 2011), no experimental studies have been completed to date regarding the efficacy of the SSIS-CIP. As such, the purpose of this study was to evaluate the efficacy of the SSIS-CIP using a multisite cluster randomized trial (CRT). Consistent with this design, classrooms are randomly assigned to intervention and control conditions within schools. As a result, the CRT design is replicated across multiple schools, and schools serve as a blocking variable that allows for modeling potential differences across sites (schools). The primary hypothesis was that children in classrooms implementing the SSIS-CIP demonstrate improved social skills compared to children in nonimplementing (business-as-usual) classrooms. A secondary hypothesis was that children in the SSIS-CIP condition

demonstrate fewer problem behaviors than their peers in comparison classrooms.

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 U.S. 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 20–25 students, and all students were invited to participate in the project. Although approximately 52% of students received parental permission to participate in the study, 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 a slightly higher (but nonsignificant) percentage of students who were males, received special education services, and had been retained. The intervention classrooms, however, enrolled a significantly higher percentage of students of minority status. In addition to the student participants, 39 teachers (one per classroom) also participated in the study. All of these teachers were White, and 79% were female. Most of the teachers reported significant classroom experience ($M = 14.4$ years of experience, $SD = 9$). All participants were treated in accord with the ethical principles of the American Psychological Association.

Table 1
Student Demographic Characteristics by Condition

	SSIS-CIP $N = 228$	Control $N = 204$
Age (in years)	7.37 (.38)	7.34 (.38)
Male	46.49	44.61
White[a]	66.67	79.41
Black/African American[a]	21.59	14.22
Asian	1.76	1.96
Hispanic or Latino	7.05	2.94
Other race	3.08	0.98
Special education consideration	5.70	6.37
Special education	11.40	6.37
Supplementary services	21.49	26.96
Retained in grade in prior year	6.14	3.43
Promoted to next grade	98.68	100.0

Note. Mean (*SD*) are reported for age; % reported for all other variables. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program.
[a] Chi-square test is significant at the .05 level.

Measures

Several measures were used to achieve the primary objectives for the study. Specifically, two measures were used to assess key outcome variables related to students' classroom behavior (Social Skills Improvement System Rating Scale, Cooperative Learning Observation Code for Kids). In addition, the Classroom Assessment Scoring System (CLASS, Pianta, La Paro, & Hamre, 2008) was used to assess the instructional environment in each participating classroom. Finally, structured observations were used to assess fidelity of implementation of the SSIS-CIP curriculum in classrooms randomly assigned to the intervention condition.

Social Skills Improvement System Rating Scales–Teacher Form (SSIS-RST; Gresham & Elliott, 2008). The SSIS-RST Social Skills and Problem Behaviors Scales were used to assess participants' behavior in the classroom setting. The Social Skills scale includes 46 items and yields seven subscales (communication, cooperation, assertion, responsibility, empathy, engagement, and self-control) in addition to a total composite. The Problem Behaviors scale includes 24 items, five subscales (externalizing, bullying, hyperactive-inattentive, internalizing, and autistic behavior), and a composite. Each item on the Social Skills and Problem Behaviors scales are rated using a 4-point format ranging from *Never* to *Almost Always*. Psychometric evidence for scores from the SSIS-RST is strong and consistent with its intended purpose (Gresham & Elliott, 2008). Reliability estimates based on data from the current sample likewise are strong (see Table 2).

Cooperative Learning Observation Code for Kids (CLOCK; Volpe & DiPerna, 2010). The CLOCK is a structured observation format that was used to facilitate independent observations of student prosocial and problem behavior in the classroom setting. The CLOCK features

Table 2
Reliability and Intraclass Correlation for Social Skills and Problem Behaviors

	Reliability index		ICC (school)		ICC (class)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Teacher rating[a]						
Social skills (composite)	.98	.98	.08	.04	.18	.23
Communication	.91	.92	.10	.05	.22	.27
Cooperation	.92	.93	.05	.02	.06	.14
Assertion	.85	.87	.04	.02	.33	.29
Responsibility	.92	.92	.08	.02	.10	.16
Empathy	.94	.94	.08	.09	.12	.13
Engagement	.93	.94	.07	.02	.21	.24
Self-control	.93	.95	.06	.03	.15	.20
Classroom observation[b]						
Positive social	.89	.89	.02	.02	.31	.32
Teacher rating[a]						
Problem behavior (composite)	.94	.95	.03	.00	.17	.24
Externalizing	.93	.94	.02	.01	.09	.16
Bullying	.90	.91	.02	.00	.03	.19
Hyperactive-inattentive	.89	.90	.03	.00	.10	.15
Internalizing	.87	.89	.03	.00	.27	.28
Classroom observation[b]						
Interference	.88	.83	.24	.14	.07	.09

Note. ICC = intraclass correlation.
[a] Cronbach's alpha. [b] Kappa agreement index.

two categories of student behavior relevant to the primary research questions for this study: positive social and interference. Positive social encompasses any appropriate social behavior that is permitted during the observation interval. Interference is intended to measure instances where a child's behaviors are distracting others or disruptive to the functioning of the classroom. Each of these behaviors is observed using a partial interval format with each interval lasting 15 s. 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 (BOYS; Volpe & Missal, 2007) and the Behavior Observation System for Students (BOSS; Shapiro, 1996).

Six participants (three of each gender) were randomly identified as target students for the CLOCK observations within a classroom, and each of these students were 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), each observation was completed during mathematics instruction. (Both participating districts used Everyday Math curriculum, which features collaborative learning and discussion.) Each observation was 12 min in length, and one paired-observation (two raters) was completed per target student and data collection period. Interobserver agreement for the paired CLOCK observations was high across all target behavior domains and paired observations (see Table 2). Consistent with research regarding the relationship between direct observation and teacher ratings (e.g., Abidin & Robinson, 2002; Childs, 1997; Hinshaw, Han, Erhardt, & Huber, 1992), the CLOCK observation data demonstrated small to moderate relationships with corresponding subscales on the SSIS-RST in the current sample. Specifically, positive social was positively related with the SSIS-RST social skills composite ($r = .19$) and negatively related with the SSIS-RST problem behaviors composite ($r = -.14$). Conversely, interference was negatively related with the SSIS-RST social skills composite ($r = -.16$) and positively related with the SSIS-RST problem behavior composite ($r = .30$).

Classroom Assessment Scoring System: Kindergarten–Third Grade (CLASS K-3; Pianta, La Paro, & Hamre, 2008). The CLASS K-3 is a structured observation system developed to assess the overall quality of the classroom instructional environment in the primary grades.

Specifically, the 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, and 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 min of observation followed by 10 min 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 (Hamre, Mashburn, Pianta, & LoCasale-Crouch, 2008) and provides support for its intended purpose. For the current study, each classroom was observed once (2 cycles) during the first data collection window to determine if 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), and interrater correlations based on paired observations ($N = 18$) were moderate to high (.56–.76) for the CLASS domain scores.

Procedure

Recruitment. Data were collected as part of a multiyear project including two separate efficacy trials of the SSIS-CIP in primary classrooms. The present study was the initial trial from this project and 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. Reminder letters were sent to all parents approximately 4 days after the initial invitation letter was sent home. By the conclusion of the recruitment period, approximately 52% of students received parental consent to participate in the data collection associated with the efficacy trial.

Data collection. 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 SSIS-RST for all participating children from their classroom. All participating teachers were paid for the time required to complete questionnaires. In addition, research data collectors completed CLOCK observations for a randomly selected subsample of participating students (three boys and three girls) from each classroom. As noted previously, 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 ($N = 27$) had at least a bachelor's degree in psychology, education, or a related discipline. In addition, data collectors completed formal training (approximately 12 hr of didactic instruction, practice observations, and individualized feedback) regarding the application and use of the 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 (see Table 2). Observations were distributed approximately evenly across observers.

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. Specifically, the skills that comprise the 10 SSIS-CIP units include (a) listening to others, (b) following directions, (c) following classroom rules, (d) ignoring peer distractions, (e) asking for help, (f) taking turns in conversations, (g) cooperating with others, (h) controlling temper in conflict situations, (i)

acting responsibly with others, and (j) showing kindness to others. Each unit focuses on a single social skill and includes three scripted lessons. Each lesson, in turn, relies on six instructional strategies (describe, model, role-play, do, practice, monitor progress, and generalize) to help children learn the target skill for that unit. In addition, brief (30–90-s) video vignettes are shown at specific junctures within each lesson, and students complete an instructional booklet with practice exercises/activities that correspond with the aforementioned instructional strategies. Each lesson requires approximately 20–25 min to complete in the primary grades. Additional information regarding the SSIS-CIP is available in the instructor's handbook (Elliott & Gresham, 2008).

Teachers whose classrooms were 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. (Because the participating districts were in different locations, separate workshops were completed for the intervention teachers in each district. This approach also allowed for a smaller number of participants in each workshop.) Each training workshop followed the same structured protocol. Specifically, 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. As a result of school holidays, special activities (e.g., field trips), and occasional closings due to weather, some teachers were not able to complete one unit each week. All participating teachers, however, completed all units (30 lessons) within a 12-week period.

Fidelity of SSIS-CIP implementation. Two complimentary methods were used to evaluate and ensure fidelity of implementation of the SSIS-CIP lessons. First, implementing teachers were asked to provide weekly self-reports regarding their implementation of the SSIS-CIP unit for that week. These reports required completion of a standardized checklist where the teacher indicated the level of implementation (using a 4-point scale ranging from *Not Implemented* to *Full Implementation*) for the five core components (introduce, define, discuss, identify and practice steps, and model/role-play) of each lesson within the unit. Teachers also provided feedback regarding the total time (min) required to teach each lesson ($M = 27.15$, $SD = 5.25$) and prepare for teaching all of the lessons per unit ($M = 38.73$, $SD = 23.18$).

In addition to teachers self-report of fidelity, independent observers completed direct observations for approximately 20% of the SSIS-CIP lessons taught by each teacher. For these fidelity observations, staff observed the entire lesson that was being taught that day. They then completed a structured report form that corresponded to the specific instructional objectives/sequence for the observed lesson. The report form included 20 specific instructional actions/activities, and observers recorded if each was completed (or not) during the observed lesson. Similar to the teacher reports, observers also 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, the lead researchers monitored fidelity (both self-report and independent observations) to ensure that teachers (a) demonstrated at least 90% fidelity in their implementation of the lessons within an instructional unit, and (b) stayed on schedule relative to the target implementation calendar. 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 if 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 ef-

forts. 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 Analyses

This study used a multisite CRT 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 take into account the nested data structure of students being nested within classes within schools. As a result of the modeling complexity attributed to three-level structures, we initially tested the degree to which the schools differed with respect to each of the outcomes of this investigation. These unconditional models yielded intraclass correlation (ICC) coefficients that indicated the degree to which the assumption of independence was violated due to the clustering of students in classes in schools (Raudenbush, 1997). If the test of the outcome variance at the school level demonstrated nonsignificance, then we planned to simplify the models to two levels. Otherwise, we planned to test three-level models.

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' sex (1 = male, 0 = female), race ethnicity (1 = White, 0 = other), and receipt of supplementary services (1 = yes, 0 = no). The dummy variable predictors were grand-mean centered. 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 (sex, race, and receipt of supplementary services) were tested by adding product terms between SSIS-CIP and each of the variables to the model. If a 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 teacher ratings of social skills. We used the Glimmix procedure for teacher ratings of problem behavior and all classroom observation data. Because problem behaviors were observed infrequently and classroom observations consisted of frequency data that were highly skewed, we used Poisson distribution and log link for the Glimmix procedure.

In addition, we estimated effect sizes of SSIS-CIP as 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) for "ES computation based on results from HLM analyses in studies with cluster-level assignment" (WWC, n.d., p. 45). Pooled within-group standard deviation of pretest scores was used because pretest scores were not affected by treatment.

Results

Table 2 presents ICCs at both class- and school-levels for all outcome measures.

Class-level ICCs for posttest outcome measures ranged from medium (.09 for classroom observations of interference) to large (.32 for positive social). 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 2-level model was used for each outcome to provide proper standard error estimates.

School-level variances of all posttest outcome measures were small and statistically nonsignificant based on *z* tests (2-tailed *p*s $> .05$). However, school-level ICC for posttest teacher ratings of empathy (.09) and classroom observations of interference (.14) were considered

medium-sized[1] (Raudenbush, Spybrook, Liu, & Congdon, 2005). We examined both 2- and 3-level models for these two 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 empathy (Δ deviance = 2.2, 2-tailed $p = .07$) and interference (Δ deviance = 0) and the fixed effect estimates and test results were similar between the 2- and 3-level counterparts, we decided to report the 2-level model results for parsimony.

Figure 1 depicts the flow of classroom and student participants throughout the study. Given the low percentage of missing data (1.3%–1.5%) and that these data were missing completely at random (MCAR; Little's chi-square = 147.26, $df = 155$, $p > .05$), cases were deleted listwise for analysis. 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 2-level models) on any of the pretest measures[2] (Tables 3, 4, 5, and 6).

Parameter estimates for the final multilevel model for each of the social skills and problem behavior outcome variables are presented in Tables 7 and 8, respectively. There was a statistically significant interaction between treatment condition and class-level pretest on teacher ratings of most social skills measures including social skills composite, communication, cooperation, responsibility, and empathy. As shown in Figures 2, 3, 4, 5, and 6, the adjusted differences between treatment and control classrooms were larger for classes that had lower average pretest scores. For classes that had high average pretest scores (e.g., .33 standard deviation units above the mean of the social skills composite pretest scores in Figure 2), SSIS-CIP participation did not improve their average scores when holding other variables constant. For classes that had low average pretest scores (e.g., < .33 standard deviation units above the mean for social skills composite), however, SSIS-CIP participation resulted in higher average adjusted posttest scores with larger differences for classes that had lower pretest scores. None of the other interaction effects examined on the social skills measures were statistically significant at the .05 level. Similarly, none of the interaction effects for problem behavior variables were statistically significant at the .05 level.

As expected, both student-level and class-level pretest scores were statistically significant predictors for most of the posttest outcome scores (except for classroom observations of positive social and interference). Student's race-ethnicity (i.e., White or racial minority) did not make a statistically significant difference on any of the adjusted posttest outcome scores. Student's sex (male or female) generally did not make a statistically significant difference on the adjusted posttest outcome scores except for teacher ratings of assertion and empathy. After adjustment for pretest differences in teacher ratings, male students received lower average teacher ratings on assertion (approximately .12 points) and empathy (approximately .11 points) than female students. Moreover, receipt of supplementary service did not make a statistically significant difference on any of problem behavior measures but did make a difference in teacher ratings on most of the social skills measures (except empathy and self-control). 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 these measures than students who did not receive any such service.

SSIS-CIP implementation yielded a statistically significant difference on posttest teacher ratings of engagement (adjusted standardized difference = .34), and internalizing behavior (adjusted standardized difference = −.24) after controlling for their respective pretest ratings and students' sex, race-ethnicity, and receipt of supplementary service. The magnitude of these effect sizes are small-medium according to Cohen's (1988) criterion but "substantively impor-

[1] We also conducted deviance difference test between 2- and 3-level unconditional models for each outcome measure. All but empathy and interference were statistically nonsignificant at the .05 level.

[2] As noted in Figure 1, some cases were excluded from the final analyses due to missing demographic data. There were no statistically significant differences on any of the pretest measures between these cases and those retained in the analyses.

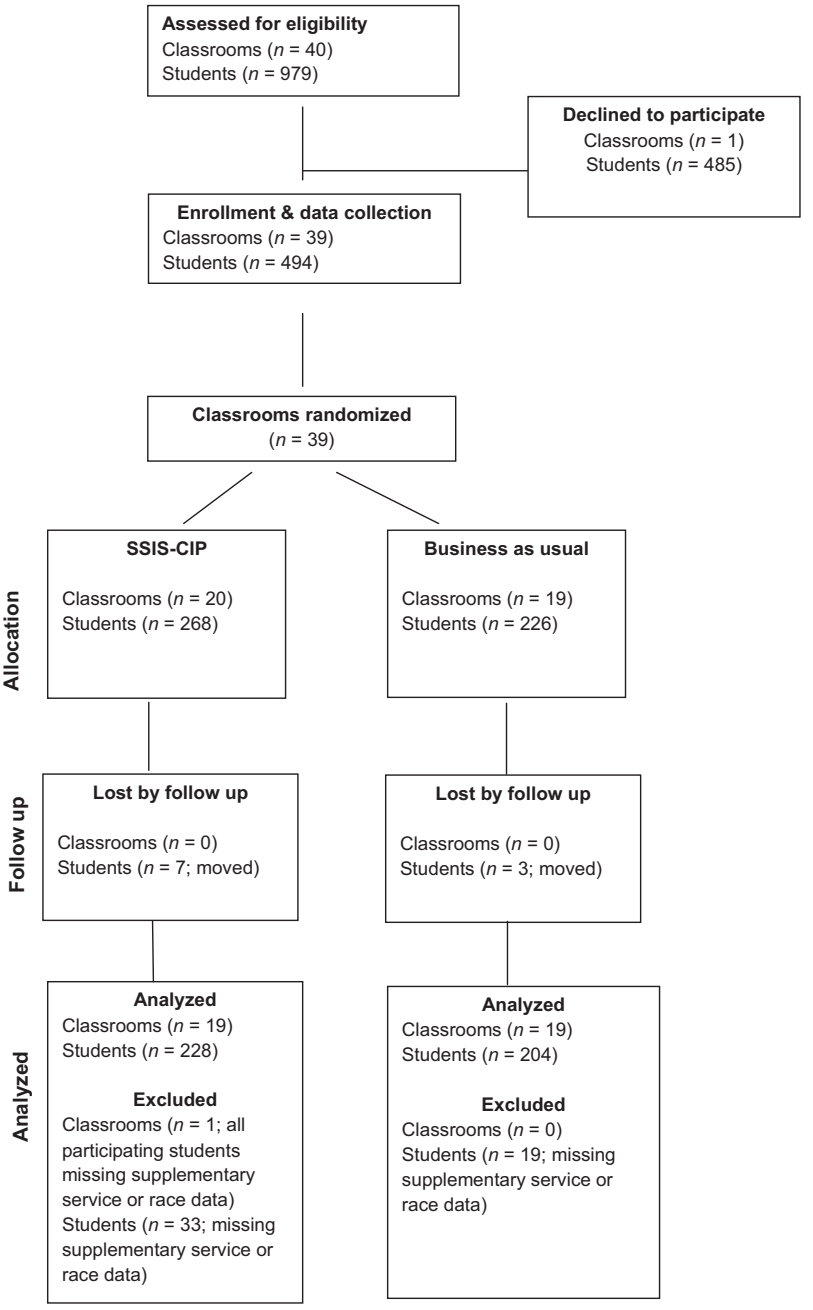

Figure 1. Social Skills Improvement System Classwide Intervention Program (SSIS-CIP) randomized trial participant flow chart.

tant" based on WWC's .25 or greater criterion (WWC, n.d., p .60).

Although the main effects associated with SSIS-CIP participation were significant for teacher ratings of overall social skills, communication, cooperation, responsibility, and empathy, these effect sizes (calculated at the mean of classes' pretest levels) should be interpreted

Table 3
Student-Level Mean Item Scores (Standard Deviations) on Social Skills Measures by Time and Treatment Conditions

	Pretest		Posttest	
	SSIS-CIP	Control	SSIS-CIP	Control
Teacher rating[a]				
Social skills (composite)	2.20 (.49)	2.14 (.56)	2.39 (.47)	2.14 (.57)
Communication	2.33 (.54)	2.26 (.62)	2.50 (.49)	2.26 (.61)
Cooperation	2.10 (.65)	2.03 (.66)	2.28 (.63)	2.02 (.70)
Assertion	1.97 (.55)	1.97 (.62)	2.23 (.55)	2.04 (.60)
Responsibility	2.29 (.58)	2.22 (.64)	2.45 (.54)	2.20 (.64)
Empathy	2.20 (.56)	2.21 (.66)	2.38 (.54)	2.16 (.67)
Engagement	2.30 (.56)	2.17 (.62)	2.49 (.52)	2.17 (.65)
Self-control	2.23 (.58)	2.10 (.69)	2.38 (.59)	2.11 (.67)
Observation[b]				
Positive social	.54 (.66)	.37 (.46)	.39 (.60)	.25 (.44)

Note. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program.
[a] SSIS-CIP: $N = 228$; Control: $N = 204$. [b] SSIS-CIP: $N = 106$; Control: $N = 102$.

with caution because the effect of the SSIS-CIP appeared to be dependent on class pretest levels of the respective outcomes for these measures. As noted previously, SSIS-CIP demonstrated a stronger positive effect for classes with lower pretest scores for these measures. The effect of SSIS-CIP participation on the other teacher-rated variables (assertion, self-control, positive social, problem behavior composite, externalizing, bullying, hyperactive-inattentive) was in the expected direction, although the adjusted differences were nonsignificant and effect sizes were small. Similarly, direct observation data (positive social, interference) did not demonstrate any significant differences between treatment and control conditions.

Discussion

The purpose of this study was to test the efficacy of the SSIS-CIP Primary Version on students' classroom behavior in the second grade. Participants' prosocial and problem behaviors were the primary outcomes of interest, and multiple measures were used to assess these variables. Students in SSIS-CIP

Table 4
Class-Level Mean Item Scores (Standard Deviations) on Social Skills Measures by Time and Treatment Conditions

Measures	Pretest		Posttest		Adjusted standardized difference[a]
	SSIS-CIP	Control	SSIS-CIP	Control	
Teacher rating					
Social skills (composite)	2.20 (.26)	2.19 (.38)	2.39 (.23)	2.18 (.36)	.36
Communication	2.35 (.30)	2.31 (.43)	2.51 (.26)	2.31 (.42)	.30
Cooperation	2.08 (.27)	2.08 (.37)	2.28 (.26)	2.07 (.40)	.30
Assertion	1.98 (.41)	2.03 (.43)	2.26 (.37)	2.08 (.35)	.26
Responsibility	2.28 (.31)	2.27 (.36)	2.44 (.25)	2.25 (.36)	.29
Empathy	2.17 (.24)	2.27 (.36)	2.35 (.22)	2.20 (.39)	.35
Engagement	2.32 (.30)	2.22 (.42)	2.50 (.25)	2.20 (.37)	.34
Self-control	2.22 (.29)	2.15 (.42)	2.36 (.32)	2.16 (.41)	.22
Observation					
Positive social	.54 (.45)	.34 (.28)	.37 (.42)	.25 (.27)	.02

Note. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program. SSIS-CIP: $N = 19$, Control: $N = 19$.
[a] Adjust for pretest scores and other student- and class-level covariates.

Table 5
Student-Level Mean Item Scores (Standard Deviations) on Problem Behavior Measures by Time and Treatment Conditions

	Pretest		Posttest	
	SSIS-CIP	Control	SSIS-CIP	Control
Teacher rating[a]				
Problem behavior (composite)	.43 (.43)	.44 (.44)	.39 (.44)	.50 (.49)
Externalizing	.40 (.47)	.42 (.50)	.39 (.48)	.48 (.54)
Bullying	.20 (.43)	.24 (.43)	.21 (.41)	.27 (.48)
Hyperactive-inattentive	.68 (.63)	.63 (.56)	.59 (.62)	.67 (.60)
Internalizing	.40 (.48)	.43 (.47)	.34 (.49)	.50 (.52)
Observation[b]				
Interference	.38 (.55)	.26 (.31)	.27 (.39)	.27 (.51)

[a] SSIS-CIP: $N = 228$; Control: $N = 204$. [b] SSIS-CIP: $N = 106$; Control: $N = 102$.

classrooms were expected to demonstrate increases in social skills and decreases in problem behaviors relative to their peers in comparison classrooms. Based on teacher ratings via the SSIS-RST, students' overall (composite) social skills scores demonstrated significant positive increases (small-medium effects) relative to their peers in control classrooms. Similarly, significant positive increases (small-medium effects) were observed in teacher ratings of communication, cooperation, responsibility, empathy, and social engagement. These main effects are similar to those reported in a recent meta-analysis regarding the effects of classroom social skills training programs on elementary students' social behavior (January et al., 2011).

In addition to main effect differences, class pretest by treatment condition interactions were significant for teacher ratings of the social skills composite, communication, cooperation, responsibility, and empathy domains. These interactions indicated greater gains were observed (relative to peers in the control condition) for students in classrooms with lower mean teacher-rated scores prior to implementation of the SSIS-CIP. As such, it appears that students in classrooms with the greatest need for social skills intervention are likely to benefit the most from participating in the SSIS-CIP lessons. This finding is consistent with prior research indicating that at-risk children benefit the most from early intervention services (e.g., Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Elias & Allen, 1991).

Although findings were as expected in most of the social skill domains, observed differ-

Table 6
Class-Level Mean Item Scores (Standard Deviations) on Problem Behavior Measures by Time and Treatment Conditions

Measures	Pretest		Posttest		Adjusted standardized difference[a]
	SSIS-CIP	Control	SSIS-CIP	Control	
Teacher rating					
Problem behavior (composite)	.44 (.26)	.43 (.23)	.42 (.26)	.47 (.29)	−.18
Externalizing	.42 (.25)	.40 (.21)	.42 (.25)	.46 (.28)	−.13
Bullying	.22 (.19)	.22 (.17)	.25 (.23)	.25 (.26)	−.08
Hyperactive-inattentive	.68 (.30)	.61 (.28)	.62 (.30)	.64 (.33)	−.19
Internalizing	.41 (.31)	.42 (.30)	.36 (.31)	.48 (.32)	−.24
Classroom observation					
Interference	.39 (.34)	.27 (.19)	.28 (.27)	.28 (.26)	−.23

Note. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program. SSIS-CIP: $N = 19$, Control: $N = 19$.
[a] Adjust for pretest scores and other student- and class-level covariates.

Table 7
Mixed Model Estimates (Standard Errors) for Social Skills

	Teacher rating								Observation
	Soc. skill (comp.)	Comm.	Cooper.	Assertion	Respons.	Empathy	Engage.	Self-control	Positive social[a]
Fixed effects									
Intercept	2.18** (.04)	2.31** (.05)	2.07** (.05)	2.06** (.06)	2.24** (.05)	2.16** (.05)	2.22** (.05)	2.17** (.05)	−1.37** (.21)
Student-level pretest score	.76** (.03)	.65** (.04)	.73** (.03)	.65** (.04)	.73** (.03)	.64** (.04)	.74** (.04)	.74** (.03)	.10 (.21)
Class-level pretest score	.89** (.13)	.87** (.12)	.98** (.16)	.66** (.11)	.91** (.15)	.88** (.12)	.66** (.11)	.76** (.11)	1.23** (.30)
Condition	.19** (.06)	.18* (.07)	.19* (.07)	.15† (.08)	.18* (.07)	.21** (.06)	.20** (.07)	.14† (.07)	.03 (.30)
Class-level pretest score × Condition	−.63** (.22)	−.62** (.22)	−.64* (.26)	NA	−.61** (.22)	−.73** (.23)	NA	NA	NA
Male	−.05 (.03)	−.05 (.03)	−.06 (.04)	−.12** (.04)	−.06 (.04)	−.11** (.04)	−.03 (.04)	−.04 (.04)	−.38 (.26)
White	−.02 (.04)	−.03 (.04)	.06 (.05)	−.06 (.05)	−.05 (.04)	−.01 (.05)	−.06 (.05)	−.01 (.05)	.06 (.33)
Supp. services	−.13** (.04)	−.15** (.05)	−.10* (.05)	−.13** (.05)	−.09 (.05)	−.12** (.05)	−.09 (.05)	−.05 (.35)	−.11** (.04)
Random effects									
Intercept	.03** (.01)	.03** (.01)	.03** (.01)	.05** (.02)	.03** (.01)	.02** (.01)	.04** (.01)	.04** (.01)	.07 (.13)
Variance									
Residual	.08** (.01)	.11** (.01)	.16** (.01)	.13** (.01)	.12** (.01)	.16** (.01)	.13** (.01)	.14** (.01)	—
Variance									

Note. NA = not included in the model because it was not statistically significant at the .05 level. Soc. skill (comp.) = Social skill composite; Comm. = Communication; Cooper.= Cooperation; Respons.= Responsibility, Engage. = Engagement; Supp. = supplementary.
[a] Outcome variable is on log scale.
† $p < .10$. * $p < .05$. ** $p < .01$.

Table 8
Mixed Model Estimates (Standard Errors) for Problem Behavior Measures

	Teacher rating					
	Problem behavior (composite)	Externalizing	Bullying	Hyperactive-inattentive	Internalizing	Observation Interference
Fixed effects						
Intercept	−.93** (.11)	−1.01** (.11)	−1.66** (.17)	−.56** (.09)	−.97** (.11)	−1.26** (.20)
Student-level pretest score	1.16** (.16)	1.09** (.13)	1.01** (.14)	.91** (.10)	1.11** (.15)	.09 (.29)
Class-level pretest score	1.80** (.31)	1.86** (.37)	3.09** (.65)	1.31** (.22)	1.60** (.22)	1.91** (.45)
Condition	−.21 (.15)	−.18 (.15)	−.19 (.23)	−.22† (.12)	−.36* (.15)	−.45 (.31)
Male	.07 (.15)	.13 (.15)	.0001 (.20)	.15 (.13)	.02 (.15)	.35 (.28)
White	.04 (.16)	−.02 (.16)	−.10 (.23)	.02 (.14)	.12 (.17)	−.44 (.29)
Supp. services	.07 (.16)	.09 (.16)	.24 (.22)	.10 (.16)	.08 (.16)	−.09 (.35)
Random effect						
Intercept variance	<.0001	<.0001	.07	<.0001	<.0001	.03

Note. Outcome variables are on log scale. No significant interaction is observed for any of the problem behavior measures; Supp. = supplementary.
† $p < .10$. * $p < .05$. ** $p < .01$.

ences between intervention and control classrooms on teacher ratings of students' assertion and self-control behaviors did not meet a priori criterion for statistical significance. Interestingly, the authors of the SSIS-CIP (Elliott & Gresham, 2007) indicated that half of the units are focused primarily on cooperation behaviors, with the remaining units addressing the areas of self-control, assertion, empathy, and responsibility. Many of the target skills of the SSIS-CIP, however, appear to encompass multiple social skill domains (e.g., *taking turns* requires cooperation, communication, and self-control), and the results of the current study suggest that certain skill areas (e.g., cooperation, communication) possibly are impacted by the SSIS-CIP more so than others (self-control, assertion).

In addition to some variability among teacher ratings of social behavior outcomes, it is important to note that direct observations of positive social behavior did not indicate any significant differences between intervention and control conditions. There are several considerations regarding these findings. First, elementary classroom teachers' judgments are based on the universe of a child's behavior within their classroom; whereas the independent observations in the current study were based on a limited number of observations (three 12-min observations) with a subsample of participants during a few weeks before and after SSIS-CIP implementation. Second, in an attempt to maximize the likelihood that the observations captured positive social behavior when it occurred, a partial interval recording method was used to record occurrence of this domain of behavior. In addition, positive social behavior was broadly defined within the CLOCK observation system. Despite employing these strategies, these behaviors were still recorded with relatively low frequency in both intervention and control conditions. Third, although systematic direct observation is often considered an objective method for measuring student behavior, researchers (e.g., Hintze & Matthews, 2004) have raised questions about the generalizability of data based on fewer than eight observations (or more depending on variability of the target behavior).

Fourth, in light of the observed differences across teacher ratings and observation data in

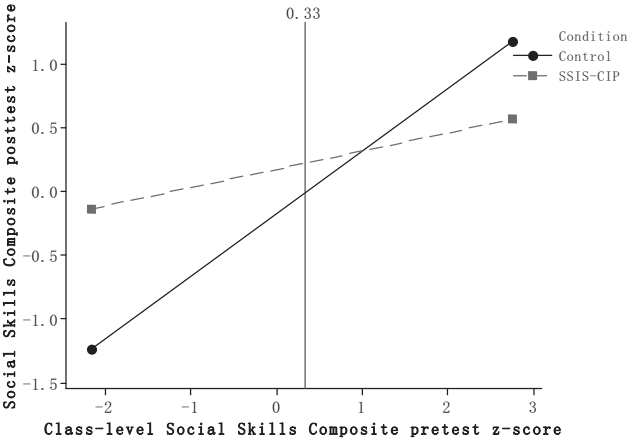

Figure 2. Interaction between treatment condition and class-level social skills composite pretest on social skills composite posttest score. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program. (Area to the left of the vertical line represents the region of statistically significant differences between conditions.)

the current study, it is possible that implementation of the SSIS-CIP changes teachers' *perceptions* of students' social behavior without changing the actual behavior itself. Given that teachers rely on their judgments of student behavior and achievement to guide their instructional practices and interactions with students (e.g., Abidin & Robinson, 2002; Dompnier, Pansu, & Bressoux, 2006; Hoge & Coladarci, 1989), changing their perceptions of student behavior has potential positive implications for students. Nonetheless, despite completing as many direct observations for as many students as resources allowed, methodological differences (between direct observation and teacher rating) and/or limitations (e.g., low frequencies and little variation in observation, potential teacher bias in teacher ratings) may have contributed to some of the differences noted across outcome measures and need to be explored further.

With regard to problem behaviors, observed outcomes were mixed, and as a result, some-

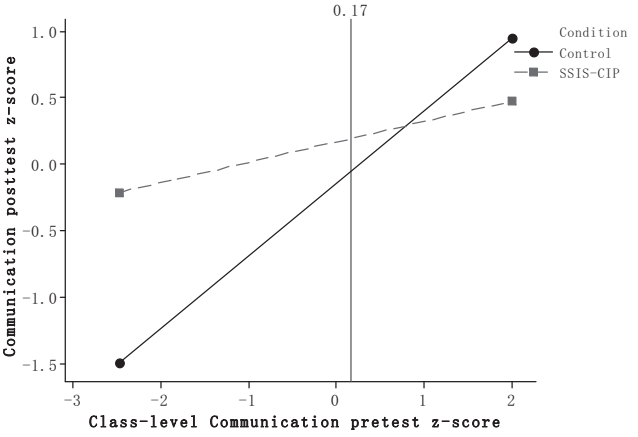

Figure 3. Interaction between treatment condition and class-level communication pretest score on communication posttest score. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program.(Area to the left of the vertical line represents the region of significant differences between treatment conditions.)

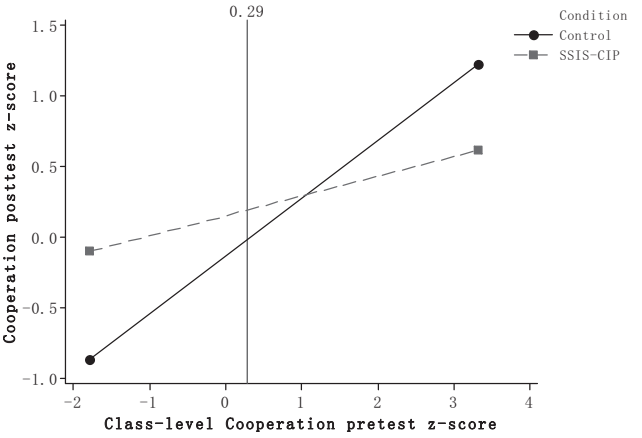

Figure 4. Interaction between treatment condition and class-level cooperation pretest score on cooperation posttest score. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program. (Area to the left of the vertical line represents the region of significant differences between treatment conditions.)

what inconsistent with our second hypothesis. Specifically, although greater reductions were observed in SSIS-CIP classrooms than control classrooms across all SSIS-RST problem behavior domains, the only domain where these differences achieved statistical significance was internalizing behavior. Similarly, other studies (e.g., Jones, Brown, Hoglund, & Aber, 2010; Weiss, Harris, Catron, & Han, 2003) have found that interventions focused on social behavior demonstrate greater impact on internalizing than externalizing behaviors. In addition, Schneider and Byrne (1985) hypothesized that internalizing problems may be more amenable to social skills training as these behaviors are more closely related to social skill deficits; whereas externalizing problems are related to the application of acquired skills and inability to use these skills in appropriate situations. Consistent with this hypothesis, many of the areas assessed

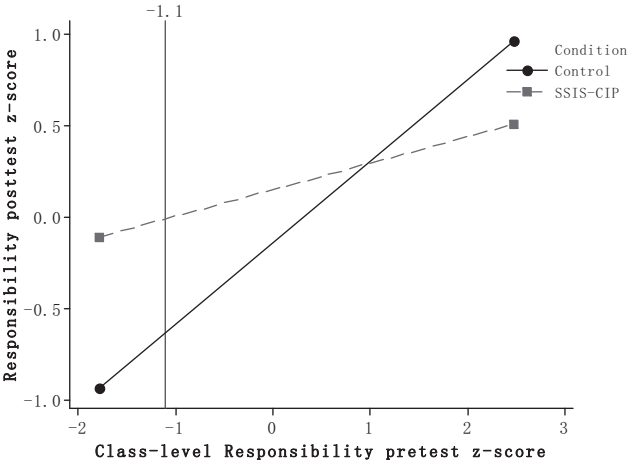

Figure 5. Interaction between treatment condition and class-level responsibility pretest score on responsibility posttest score. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program. (Area to the left of the vertical line represents the region of significant differences between treatment conditions.)

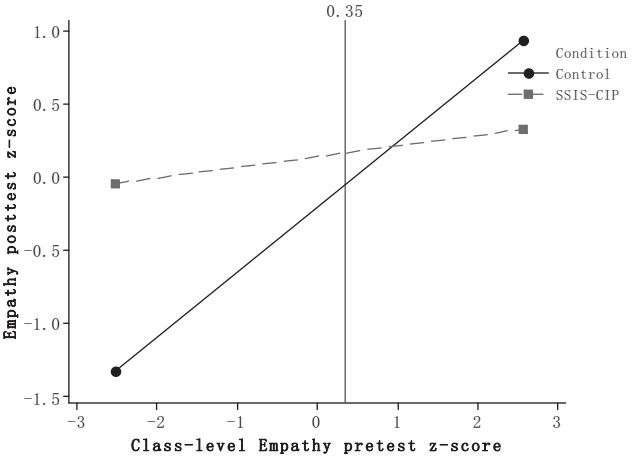

Figure 6. Interaction between treatment condition and class-level empathy pretest score on empathy posttest score. SSIS-CIP = Social Skills Improvement System Classwide Intervention Program. (Area to the left of the vertical line represents the region of significant differences between treatment conditions.)

on the SSIS-RST Internalizing subscale (e.g., depression) have been shown to be related to deficits in social skills (e.g., Segrin, 2000).

In light of the results and limitations of the current study, there are several important directions for future research regarding the efficacy of the SSIS-CIP. First, although there were a sufficient number of students, schools, and classrooms to test for the hypothesized effects, future studies should include a larger sample of schools to increase representativeness and test for school-level effects. In addition, the overall participation rate was lower than other recent studies of universal interventions (e.g., Bradshaw et al., 2009). Although participating students' pretest scores were in the average range relative to the SSIS-TRS standardization sample, it is unclear if the current results generalize to the subpopulation of students represented by those children who did not receive permission to participate in the current study. All data also were drawn from a single elementary grade level (second grade), and future studies are necessary to test the effects of the SSIS-CIP Primary Version for students in Grades K–1, as well as the effects of the SSIS-CIP Intermediate Version for students in Grades 3–8.

Beyond sample considerations, one of the outcome measures in the current study (SSIS-TRS) informed the development of the SSIS-CIP. As such, future studies could utilize additional measures of social skills to ensure that results are not unique to the method of outcome measurement. Similarly, the current measures of intervention fidelity focused primarily on coverage of required lesson content, and future studies could benefit from the inclusion of additional process-related implementation factors. Longitudinal studies also are necessary to determine if gains observed immediately following implementation of the SSIS-CIP are maintained through the end of the academic year and beyond. Some (e.g., Durlak & Weissberg, 2007) have suggested that there are benefits in other domains (e.g., improved academic performance) that result from implementing social-emotional programs in schools. As such, future studies should examine the possibility of positive outcomes as well as potential lost opportunity costs (e.g., decreased instructional time) that result from implementation of classwide social skills curricula such as the SSIS-CIP. Finally, future studies should examine if some units or lessons within the SSIS-CIP have a more significant impact on social skills than others.

In sum, results from the current study suggest that the SSIS-CIP is potentially efficacious in promoting the development of prosocial skills in the second grade—particularly in the areas of cooperation, communication, responsibility, social engagement, and empathy. In addition, par-

ticipation in the SSIS-CIP appears to reduce young students' problem behaviors in the internalizing domain (e.g., withdrawal, shyness). In contrast to our predictions, however, the SSIS-CIP did not yield significant behavior change in the externalizing domains (e.g., aggression, hyperactivity). Thus, if schools are interested in addressing this domain as well, the SSIS-CIP may need to be supplemented with additional lessons/units specifically targeting such behavior. Although results of the current study are promising, it represents the first attempt to test the efficacy of the SSIS-CIP. As such, additional research is necessary to ensure that the current findings are accurate (replicable), sustain over time, and generalize to students in other elementary grades.

References

Abidin, R. R., & Robinson, L. L. (2002). Stress, biases, or professionalism: What drives teachers' referral judgments of students with challenging behaviors? *Journal of Emotional and Behavioral Disorders, 10,* 204–212. doi:10.1177/10634266020100040201

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. doi:10.1901/jaba.1969.2-119

Beelmann, A., Pfingsten, U., & Losel, F. (1994). Effects of training social competence in children: A meta-analysis of recent evaluation studies. *Journal of Clinical Child Psychology, 23,* 260–271. doi:10.1207/s15374424jccp2303_4

Bierman, K. L., Coi, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (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. doi:10.1037/a0018607

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. doi:10.1037/a0016586

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37,* 231–242. doi:10.1037/0012-1649.37.2.231

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. doi:10.1037/a0027725

Childs, G. (1997). A concurrent validity study of teachers' ratings for nominated 'problem' children. *British Journal of Educational Psychology, 67,* 457–474. doi:10.1111/j.2044-8279.1997.tb01258.x

Cohen, J. (1988). *Statistical power for analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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,* 648–657. doi:10.1037/0022-006X.67.5.648

Daunic, A. P., Smith, S. W., Brank, E. M., & Penfield, R. D. (2006). Classroom-based cognitive-behavioral intervention to prevent aggression: Efficacy and social validity. *Journal of School Psychology, 44,* 123–139. doi:10.1016/j.jsp.2006.01.005

Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the Preschool PATHS curriculum. *The Journal of Primary Prevention, 28,* 67–91. doi:10.1007/s10935-007-0081-0

Dompnier, B., Pansu, P., & Bressoux, P. (2006). An integrative model of scholastic judgments: Pupils' characteristics, class context, halo effect and internal attributions. *European Journal of Psychology of Education, 21,* 119–133. doi:10.1007/BF03173572

DuPaul, G. J., & Eckert, T. L. (1994). The effects of social skills curricula: Now you see them, now you don't. *School Psychology Quarterly, 9,* 113–132. doi:10.1037/h0088847

Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social and Emotional Learning.

Elias, M. J., & Allen, G. J. (1991). A comparison of instructional methods for delivering a preventive social competence/social decision making program to at risk, average, and competent students. *School Psychology Quarterly, 6,* 251–272. doi:10.1037/h0088819

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.

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. doi:10.1023/A:1011591613728

Froiland, J. M. (2011). Response to intervention as a vehicle for powerful mental health interventions in the schools. *Contemporary School Psychology, 15,* 35–42. doi:10.1007/BF03340961

Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS Curriculum. *Development and Psychopathology, 7,* 117–136. doi:10.1017/S0954579400006374

Gresham, F. M. (2002). Teaching social skills to high risk children and youth: Preventive and remedial strategies. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial strategies* (pp. 403–432). Bethesda, MD: National Association of School Psychologists.

Gresham, F., & Elliott, S. N. (2008). *Social skills improvement system rating scales*. Minneapolis, MN: Pearson Assessments.

Hamre, K. B., Mashburn, A. J., Pianta, R. C., & LoCasale-Crouch, J. (2008). *Classroom assessment scoring system, Pre-K: Technical appendix*. Baltimore, MD: Brookes.

Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21,* 143–150. doi:10.1207/s15374424jccp2102_6

Hintze, J. M., & Matthews, W. J. (2004). The generalizability of systematic direct observations across time and settings: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review, 33,* 258–270.

Hoge, R., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research, 59,* 297–313. doi:10.3102/00346543059003297

January, A. M., Casey, R. J., & Paulson, D. (2011). A meta-analysis of classroom-wide interventions to build social skills: Do they work? *School Psychology Review, 40,* 242–256.

Jones, R. E., Yssel, N., & Grant, C. (2012). Reading instruction in tier 1: Bridging the gaps by nesting evidence-based interventions within differentiated instruction. *Psychology in the Schools, 49,* 210–218. doi:10.1002/pits.21591

Jones, S. M., Brown, J. L., Hoglund, W. L. G., & Aber, J. L. (2010). A school-randomized clinical trial of an integrated social-emotional learning and literacy intervention: Impacts after 1 school year. *Journal of Consulting and Clinical Psychology, 78,* 829–842. doi:10.1037/a0021383

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

Lannie, A. L., & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the Good Behavior Game on student and teacher behavior. *Education and Treatment of Children, 30,* 85–98. doi:10.1353/etc.2007.0002

Lembke, E. S., Hampton, D., & Beyers, S. J. (2012). Response to intervention in mathematics: Critical elements. *Psychology in the Schools, 49,* 257–272. doi:10.1002/pits.21596

Li, L., Washburn, I., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., . . . Flay, B. R. (2011). Effects of the Positive Action programme on problem behaviours in elementary school students: A matched-pair randomized control trial in Chicago. *Psychology and Health, 26,* 187–204. doi:10.1080/08870446.2011.531574

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

Pyle, N., & Vaughn, S. (2012). Remediating reading difficulties in a response to intervention model with secondary students. *Psychology in the Schools, 49,* 273–284. doi:10.1002/pits.21593

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2,* 173–185. doi:10.1037/1082-989X.2.2.173

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].

Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal of Positive Behavior Interventions, 15,* 39–51. doi:10.1177/1098300712459079

Reinke, W. M., Lewis-Palmer, T., & Martin, E. (2007). The effect of visual performance feedback on teacher use of behavior specific praise. *Behavior Modification, 31,* 247–263. doi:10.1177/0145445506288967

Scarpa, A., & Lorenzi, J. (2013). Cognitive-behavioral therapy with children and adolescents: History and principles. In A. Scarpa, S. Williams White, & T. Attwood (Eds.), *CBT for children and adolescents with high-functioning autism spectrum disorders* (pp. 3–26). New York, NY: Guilford Press.

Schneider, B. H., & Byrne, B. M. (1985). Children's social skills: A meta-analysis. In B. H. Schneider, K. H. Rubin, & J. E. Ledingham (Eds.), *Children's peer relations: Issues in assessment and interven-*

tion (pp. 175–192). New York, NY: Springer-Verlag. doi:10.1007/978-1-4684-6325-5_11

Segrin, C. (2000). Social skills deficits associated with depression. *Clinical Psychology Review, 20,* 379–403. doi:10.1016/S0272-7358(98)00104-4

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

Spence, S. H. (2003). Social skills training with children and young people: Theory, evidence and practice. *Child and Adolescent Mental Health, 8,* 84–96. doi:10.1111/1475-3588.00051

Sutherland, S. S., Wehby, J. H., & Copeland, S. R. (2000). Effect of varying rates of behavior-specific praise on the on-task behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 8,* 2–8. doi:10.1177/106342660000800101

Vaughn, S., Kim, A., Sloan, C. V. M., Hughes, M. T., Elbaum, B., & Sridhar, D. (2003). Social skills interventions for young children with disabilities: A synthesis of group design studies. *Remedial and special education, 24,* 2–15. doi:10.1177/074193250302400101

Volpe, R. J., & DiPerna, J. C. (2010). *Cooperative learning observation code for kids*. Unpublished observation code.

Volpe, R. J., & Missal, K. (2007). *Behavioral observation of young students*. Unpublished observation code.

Washburn, I. J., Acock, A., Vuchinich, S., Snyder, F., Li, K., Ji, P., . . . Flay, B. R. (2011). Effects of a social-emotional and character development program on the trajectory of behaviors associated with social-emotional and character development: Findings from three randomized trials. *Prevention Science, 12,* 314–323. doi:10.1007/s1112101102309

Weiss, B., Harris, V., Catron, T., & Han, S. S. (2003). Efficacy of the RECAP intervention program for children with concurrent internalizing and externalizing problems. *Journal of Consulting and Clinical Psychology, 71,* 364–374. doi:10.1037/0022-006X.71.2.364

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

Whitcomb, S. A., & Merrell, K. W. (2012). Understanding implementation and effectiveness of Strong Start K-2 on social-emotional behavior. *Early Childhood Education, 40,* 63–71. doi:10.1007/s10643-011-0490-9

Received November 18, 2013
Revision received June 3, 2014
Accepted June 6, 2014 ■