Social-Emotional Learning Foundations for K-1 students at Risk for EBD: Findings from a 3-Year RCT

Ann P. Daunic, Nancy L. Corbett, Stephen W. Smith, James Algina,

Emily Crews, Delaney Boss, Jessica Vezzoli

University of Florida

DRAFT

NOT FOR DISTRIBUTION OR QUOTATION WITHOUT PERMISSION FROM THE AUTHORS

Abstract

Studies have shown that children's social-emotional growth is inextricably connected to academic learning. The Social-Emotional Learning Foundations (SELF) intervention, a grade K-1 curriculum merging social-emotional learning (SEL) and literacy instruction, was developed to promote language supported self-regulation, specifically for primary grade children at early risk for emotional or behavioral difficulties. We report findings from a pretest-posttest cluster randomized efficacy trial with one fixed between-subjects factor to test the effects of teacher-delivered SEL instruction against those of business as usual (BAU). We recruited 163 kindergarten and 141 first grade teachers from 52 schools across 11 distinct school districts within one southeastern state. Our student sample (n = 1,154) consisted of 627 kindergarteners and 527 first graders identified by teachers as at risk for internalizing or externalizing emotional and behavioral problems using the *Systematic Screening for Behavioral Disorders*; 613 of these students participated in the SELF condition and 541 in the BAU condition. We randomly assigned schools to SELF or BAU and used a multilevel model with three levels (i.e., children, classrooms, schools) to analyze data on subscales of six (four teacher-report and two direct) assessments related to self-regulation, social-emotional learning, social-emotional vocabulary, and general behavioral functioning. We found positive main effects of SELF compared to BAU on all but one measure, with effect sizes (calculated using Hedges' g) ranging from 0.20 to 0.65. Findings provide evidence for guiding future SEL intervention research and informing practice to improve student outcomes, particularly for children at risk for behavior problems.

Key words: Social-emotional learning, self-regulation, Tier 2 instruction, emotional and behavioral problems, academic and social-emotional learning instruction

Social-Emotional Learning Foundations for K-1 students at Risk for EBD: Findings from a 3-Year RCT

As many as 20% of children enter school exhibiting aggressive, non-compliant, and disruptive behaviors that impact their ability to fully benefit from early educational experiences (Cooper, Masi, & Vick, 2009; World Health Organization, 2004) and increase their risk for long-term academic and behavioral problems (O'Connor, Dearing, & Collins, 2011). Teachers cite behavioral issues as one of their most pressing concerns (Meister & Melnick, 2003; Pavri, 2004), and children as young as pre-kindergarten have been removed from school because of their behavior (Gilliam & Shahar, 2006; U.S. Department of Education Office for Civil Rights, 2014). Additionally, students who have social-emotional issues are more likely to have a higher number of disciplinary infractions and/or be referred for special education because of problematic behaviors (Elias & Haynes, 2008; Stoiber, 2011). Most students at risk for behavioral issues or identified with emotional behavioral disorders (EBD) receive the majority of their instruction in the general education classroom. Even though teachers acknowledge the importance of promoting the social and emotional development of their students (Pasi, 2001), most say they received minimal preparation on how to implement effective and efficient practices (Pavri, 2004; Smith & Smith, 2000).

In the current context of high stakes evaluations and initiatives (Every Student Succeeds Act, 2015), school professionals are compelled to focus on developing academic skills to the neglect of social-emotional learning (SEL), even though children's early school success depends heavily on successful social-emotional development (see e.g., Blair & Diamond, 2008; Downer & Pianta, 2006). Particularly in early childhood, emotion regulation is positively associated with

academic success (Graziano, Reavis, Keane, & Calkins, 2007). An increasing number of researchers (Hemmeter, Ostrosky, & Fox, 2006; Riggs, Greenberg, Kusche´, & Pentz, 2006; Zins, Bloodworth, Weissberg, & Walberg, 2004) emphasize the role that motivation, self-esteem, and self-regulation play in a child's adjustment and connection to school, particularly at the critical transition from pre-school settings to kindergarten and the primary grades. SEL skills have also been shown to be a protective factor for at-risk children, including those from low socioeconomic status backgrounds (Elias & Haynes, 2008). A proactive, evidence-based approach that provides teachers access to SEL instruction is critical for promoting the development of skills such as communication, empathy, problem-solving, and self-management in young children (Eklund, Kilpatrick, Kilgus, & Haider, 2018).

Since social-emotional growth and academic learning are intricately connected (Blair & Diamond, 2008), it is both conceptually sound and practically relevant to develop interventions that integrate behavioral and academic goals (Backes, 2009), such as literacy and language arts instruction with an intentional focus on SEL. Within a multi-tiered instructional framework, SEL programs are often described as universal (Tier 1) approaches because they are designed to promote positive academic and behavioral outcomes for all students (Zins & Elias, 2007). However, students at early risk for EBD may not benefit sufficiently from universally delivered programs (Greenberg, 2010). SEL researchers have addressed this issue by providing more intensive instruction (Tier 2) designed specifically for students with or at risk for EBD (August, Hektner, Egan, Realmuto, & Bloomquist, 2002; August, Lee, Bloomquist, Realmuto & Hektner, 2003; Walker et al., 2009; Webster, Stratton, Reid, & Hammond, 2004), and this approach has shown promise. We developed the Social-Emotional Learning Foundations (SELF) curriculum to extend this work by providing intensive instruction that integrates SEL and literacy and

strengthens (a) language related to SEL, (b) self-regulation, and (c) social-emotional competence for students at early risk for social and emotional problems.

Underlying Theory

Successful social-emotional growth requires the development of self-regulatory skills that underlie healthy social, emotional, and behavioral functioning (Blair & Diamond, 2008; Riggs et al., 2006). When these processes are under-developed, children may exhibit a variety of maladaptive behavior, particularly related to skills needed for successful adjustment to school. For example, children with aggressive tendencies are often impulsive, tend to misinterpret others' intentions, lack appropriate social decision-making skills, and are often rejected by peers (Dodge et al., 2002). Behavioral difficulties tend to remain stable and are often predictive of problems in adolescence and adulthood (Bradley, Doolittle, & Bartolotta, 2008; Lochman, Dunn, & Klimes-Dougan, 1993), including increased discipline referrals, poor attendance, and higher dropout rates (Greenberg et al., 2003).

Self-regulatory skills have been closely linked to neurocognitive mechanisms known as executive function (EF: Riggs & Greenberg, 2004; Blair & Razza, 2007). These mechanisms are thought to contribute significantly to a child's social-cognitive, behavioral, and academic functioning (Blakemore & Choudhury, 2006; Greenberg et al., 2004; Zelazo & Cunningham, 2007; Tangney, Baumeister, & Boone, 2004) and are closely related to the development of self-regulation (McClelland & Cameron, 2012). Emerging research indicates that students classified as having social, emotional, or behavior disorders in schools exhibit EF deficits (Feifer & Rattan, 2007; Mattison, Hooper, & Carlson 2006). Most important, teaching students to verbally identify and label their feelings can have a powerful effect on the ability to manage emotions and regulate behavior, and encouraging children to talk about emotional experiences further strengthens the

neural integration that contributes to self-regulation (Greenberg, Kusche, & Riggs, 2004). Thus, SEL is dependent on integrated neurocognitive development that can be targeted by interventions that foster the development of self-regulatory skills foundational to social-emotional adjustment.

Research Purpose

The SELF curriculum was designed specifically to promote social-emotional development for primary grade children at early risk for EBD. Using carefully coordinated materials and pedagogy, SELF lessons focus on developing language-supported self-regulation and social-emotional competence. To evaluate the efficacy of SELF compared to a business as usual (BAU) control condition for children at risk for EBD in general education kindergarten (K) and first grade classrooms under routine conditions, we conducted a three-year pretest-posttest cluster randomized efficacy trial. As such, we examined whether children in the SELF condition evidenced more positive performance on outcomes related to social-emotional competence, social-emotional vocabulary development, self-regulation, and successful school adjustment, compared to children in the BAU condition. We also investigated whether intervention outcomes differed for children in K versus first grade classrooms and for children identified as at risk for internalizing versus externalizing behaviors.

Materials and Method

Recruitment and Sample Description

We recruited elementary schools across three years from multiple school districts in a southeastern state. We targeted schools eligible for Title I funding, as they were more likely to include children at early risk for emotional and behavioral difficulties (Kupersmidt, et al., 1995). After obtaining permission from district level administration, we solicited and confirmed participation with school principals and K-1 teachers, informing them prior to their consent that

they had an equal chance of being assigned to either the SELF intervention condition or a BAU condition. We recruited a distinct group of schools each year, with an average of 5.85 teachers per school across K and 1st grade to ensure an adequate sample size, allowing for typical rates of attrition. Across all participating schools, an average of 75% of the student population was eligible for free or reduced-price lunch. All recruitment and study procedures met university Institutional Review Board standards.

Child Screening Measures, Inclusion, and Exclusion Criteria

Participating K-1 teachers screened children for early risk for developing EBD prior to random assignment, ruling out those with school identified developmental delays. Teachers identified students at possible risk for EBD using the *Systematic Screening for Behavior Disorders* (SSBD; Walker & Severson, 1992), a cost-effective, validated, multiple-gating procedure to identify students with behavior disorders on externalizing and internalizing dimensions. In Stage 1, teachers place students in rank order from "most like" to "least like" on items describing each dimension. The three highest-ranked students on each dimension move through Gate 1 into Stage 2, where teachers complete the Critical Events Index and the Combined Frequency Index. The SSBD distinguishes among students with externalizing behaviors, internalizing behaviors, emotional disturbances, and those with typical development (e.g., Walker et al. 1994), with coefficient alphas above .90 for the standardization sample (Walker et al., 1990). In this study, teachers used Gates 1 and 2 to prioritize the top three to four students in the internalizing category and the top three to four students in the externalizing category for possible participation in either SELF or BAU. Teachers rank ordered students to identify the top two to three students in each category.

Random Assignment and Consent Procedures

Following recruitment and screening, we solicited active consent from the parents/guardians of selected students in all participating classrooms. We required that each participating teacher obtain consent from at least two and no more than five students among those designated as most at risk for internalizing or externalizing problems through the SSBD procedure. All teachers and research project staff involved in recruitment were trained on the informed consent process, and we followed all our institution's participant consent protocols. After student participants were consented, we randomly assigned schools within district to SELF or BAU so that some schools in each district would be provided SELF during their participation year. We randomized at the school level to address potential contamination between classrooms within schools, as most elementary schools operate with grade level teams whose teachers interact regularly. Project staff contacted school principals to inform them of their school's assignment to the treatment or the control group. (Note: School personnel assigned to BAU were informed they would be offered SELF, including professional development [PD] and related materials, following study participation.)

Sample Description

The sample aggregated across three years is described in the following table.

	Districts	Schools	K Teachers /Students	1st Teachers /Students	SELF Teachers /Students	BAU Teachers /Students	Total Teachers /Students
Year 1	5	19	51/203	42/162	46/184	47/181	93/365
Year 2	5	21	76/292	60/221	66/254	70/259	136/513
Year 3	6	12	35/132	38/144	46/175	27/101	73/276
Total	11*	52	162/627	140/527	158/613	144/541	302/1154

*Number of distinct participating districts.

Across demographic categories, we were able to obtain data on an average of 91.88% of the sample. Of these students, 61.75% were male and 38.25% female, 83.14% met criteria for free or reduced lunch status, 3.31% were classified as English Language Learners, and 20.31% received services through an Individualized Education Plan or accommodations via a Section 504 plan. Additionally, 64.67% were White (non-Hispanic), 19.48% African American, 9.57% Hispanic, and 6.36% of another race.

Description of the SELF Curriculum

We developed SELF (and its accompanying PD) to support students identified as at risk for EBD. While SELF includes whole-group lessons taught to all students in the class, SELF's emphasis is on more intensive intervention for the students identified as at-risk. SELF's small-group lessons provide repeated opportunities to practice SEL skills within the framework of literacy instruction.

SELF lessons, approximately 50 at each grade level, are organized using 16 children's storybooks that address SEL topics within 5 critical competencies: self-awareness, self-management, social awareness, relationship management, and responsible decision-making (Zins et al., 2004). Lessons comprise 5 units corresponding to the 5 competencies, and each unit includes 2-4 related and developmentally appropriate topics, e.g., "expressing my feelings." Lessons incorporate instructional strategies that promote children's use of SEL related vocabulary, self-talk, critical thinking, and application of learned concepts. To promote SEL for all students in the class, the first lesson in each topic directs the teacher to read the corresponding storybook, selected specifically for related social-emotional concepts and vocabulary appropriate at each respective grade level, to the whole class. Subsequently, the teacher teaches the second lesson in each SEL topic to targeted students only in a small-group setting using dialogic reading

(DR) to promote discussion about the concept addressed. In the third lesson per topic, teachers support generalization by having target students apply social-emotional concepts and skills in problem-solving situations and practice using the selected vocabulary, again within the small-group setting.

Used in the second lesson per topic, DR is a read-aloud method that provides a socially interactive context within which children can learn and apply verbal and conceptual skills (Neuman, 1996). DR typically involves using scripted prompts and questions (Flynn, 2011; Whitehurst et al., 1988) and is widely used with early childhood populations. It has been shown to increase vocabulary (Coogle, Floyd, & Rahn, 2018; Opel, Ameer, & Aboud, 2009) and expressive and receptive language skills (Simsek & Erdogan, 2015) and has been incorporated in individual, small-group, and whole-group instruction. The structure of DR enables a teacher to promote active learning and provide feedback that models sophisticated language (DeTemple & Snow, 2003) and fosters dialog (Al Otaiba, 2004; Van Kleeck & Vander Woude, 2003). When followed by a 10- to 20-minute role-play or application activity, DR was found to increase the use of social-emotional skills such as turn-taking, problem solving, and praising peers during play **(**Fettig, Cook, Morizio, Gould, & Brodsky, 2018).

Professional Development

Prior to implementing the SELF curriculum, all treatment teachers participated in two six-hour days of PD during the fall of their participation year to familiarize them with the conceptual foundations of SEL and associated competencies and to introduce essential components of SELF lessons. We also focused on curriculum implementation and deepening teachers' pedagogical knowledge to help them integrate DR and targeted vocabulary instruction with lesson objectives.

Data Collection Procedures and Sources

Assessment of Treatment Efficacy

To assess the effects of SELF compared to BAU on outcome measures, teachers completed the *Behavior Rating Inventory of Executive Function2* (BRIEF2), *Devereux Student Strengths Assessment* (DESSA), *Clinical Assessment of Behavior Teacher Rating Form* (CAB-T), and *Student Knowledge Questionnaire* (SKQ) online using the Qualtrics platform. Each participating teacher completed these measures prior to the onset of intervention (late November-early December) and near the end of the school year to coincide with completion of SELF implementation in the treatment group (late April-mid May). We enlisted the help of assessors blind to condition to administer all direct, individualized assessments, i.e., *Head-Toes-Knees-Shoulders* and the SELF Vocabulary Measure, of child participants within the same time period. The assessors, the majority of whom were former teachers, completed approximately 2-3 hours of training on direct assessments prior to administration each year.

The BRIEF2 (Gioia, Isquith, Guy, & Kenworthy, 2000) evaluates emotional and behavioral self-regulation and contains 86 items comprising eight clinical scales (Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor) that form the Behavioral Regulation Index (BRI), Cognitive Regulation Index (CRI), and the Emotion Regulation Index (ERI). The BRIEF2 assesses behavioral aspects of children's EF from respondents with unique knowledge directly relevant to self-regulation in the school environment. Sample derived Cronbach's alphas for the BRI, CRI, and ERI were .957, .972, and .937, respectively, at pretest and .962, .976, and .947, respectively, at posttest.

The DESSA (LeBuffe, Shapiro, & Naglieri, 2008) is a 72-item, standardized, norm-referenced behavior rating scale measuring SEL competencies that serve as protective factors for

children in grades K-8. Items rated on a five-point scale indicating how often the child engaged in a specified behavior over the previous four weeks are organized into eight conceptually derived scales corresponding to key social-emotional competencies. Students were assessed on the five competencies that correspond to those covered in SELF: Self-Awareness, Social-Awareness, Self-Management, Relationship Skills, and Decision Making. Sample derived Cronbach's alphas for the five competencies at pretest were .886, .899, .897, .930, and .910, respectively, and at posttest were .914, .938, .941, .956, and .945, respectively.

The CAB-T (Bracken & Keith, 2004) consists of 70 questions that comprise three clinical scales (Internalizing, Critical, Externalizing Behavior), three adaptive scales (Social Skills, Competence, Adaptive Behavior), and four educationally related clinical clusters, including EF. CAB subscales have demonstrated adequate internal reliability and are sensitive to item gradients. We analyzed scores from four subscales: two clinical (Internalizing, Externalizing) and two adaptive (Social Skills, Competence). Sample derived Cronbach's alphas at pretest were .877, .970, .933, and .934 for the four subscales, respectively. Sample Cronbach's alphas at posttest were .900, .974, .942, and .947, respectively.

The SKQ, developed by the researchers, assesses how much students know about concepts taught directly in SELF and related to SEL competencies espoused by the Collaborative for Academic, Social, and Emotional Learning (CASEL; see Zins et al., 2004). They include self-awareness, self-management, social awareness, relationship management, and responsible decision-making. The SKQ asks teachers to rate students on 11 items using a 3-point Likert scale (1 = rarely, 2 = sometimes, 3 = often). The sample derived Cronbach's alphas for the total SKQ score were .887 at pretest and .927 at posttest.

Head-Toes-Knees-Shoulders (HTKS; Cameron, Ponitz, McClelland, Matthews, & Morrison, 2009) was developed as a direct assessment of self-regulation for children ages 4-6 and is designed to assess how well children apply cognitive skills to overt behavior. It integrates multiple aspects of EF in a game involving four paired behavioral rules: "touch your head" and "touch your toes"; "touch your shoulders" and "touch your knees." After instructions that require children to respond "naturally," they are instructed to switch and respond in the "opposite" way, e.g., to touch their head when asked to touch their toes. The game thus requires children to pay attention, use working memory to remember the rule, inhibit the initial response, and initiate the correct, "unnatural" response. Particularly relevant to this project, the HTKS assesses how well children apply cognitive skills to overt behavior, a process required in classrooms settings. The HTKS has demonstrated good inter-rater reliability and scoring agreement, test-retest reliability of .93 over a three-month period, and predictive validity with school achievement (McClelland & Cameron, 2012). Sample derived Cronbach's alphas for total HTKS scores were .949 at pretest and .927 at posttest.

The *SELF vocabulary measure* is a researcher developed, curriculum-based measure designed to assess knowledge of key SEL related receptive and expressive vocabulary (van der Wissel, 1988). Each item includes three tasks: (A) provide a definition, (B) use the vocabulary word in an example, and (C) apply the word by answering a multiple-choice question. The definition task, adapted from the Oral Vocabulary subtest of the Test of Language Development (Hammill & Newcomer, 1988), yields scores of 0, 1, or 2 for incorrect, partially correct, or correct responses, respectively. Responses to the child-provided example and the multiple-choice task are scored 0 for incorrect or 1 for correct. These two tasks are adapted from vocabulary measures designed to evaluate the effect of listening to stories on primary grade children's

vocabulary growth (Justice, Meier, & Walpole, 2005; Penno, Wilkinson, & Moore, 2002). Cronbach's alphas derived from sample data at pretest were .802, .842, and .734 for parts a, b, and c, respectively. At posttest, sample derived Cronbach's alphas were .824, .863, and .770 for parts A, B, and C, respectively. The Cronbach's alpha for the total score at pretest was .918; Cronbach's alpha for the total score at posttest was .928.

Assessment of Treatment Fidelity

To monitor whether teachers implemented SELF as intended, we used two forms of the Direct Observation of Practice Protocol (DOPP), developed previously by the research team. The "DOPP-lo" is a low inference instrument that enables observers to record simple adherence to each lesson component in the manualized curriculum, such as covering all content, introducing vocabulary and prompting students as directed, and completing all activities. The "DOPP-hi" is a high inference instrument that requires the observer to rate criteria within three domains related to quality of instruction, student responsiveness, and use of key curricular strategies to promote the development of social-emotional language and self-regulation. The DOPP-hi includes three domains: Preparing the Classroom for Instruction, Development of Language to Support Self-Regulation, and Quality of Instructional Delivery. Raters assign a score of 1, 2, or 3 to criteria within each domain. Research assistants completed 4-6 hours of training and multiple practice ratings on the DOPP-hi measure before they completed any observations. The research team, either through in situ observations or by video, completed a DOPP-lo on 14.72% (1185/8048) and completed a DOPP-hi on 13.88% (1117/8048) of all lessons taught across teachers.

Research assistant and teacher schedules determined specific teachers and lessons observed, but observations were spread across the school year. Observers did not regularly provide lesson feedback or instructional support. Pairs of observers conducted 9.23% of these

observations to obtain inter-observer agreement on the DOPP-lo and 9.57% to obtain inter-observer agreement on the DOPP-hi.

Assessment of the BAU Condition

To evaluate the level of SEL instruction in the control condition, BAU teachers (*n* = 135; 93.75% response rate) completed a survey at the end of their year of study participation. Specifically, BAU teachers reported on (1) the level of storybook interactions (i.e., whether the book was available, read in whole group, read in small group, read one-on-one) for the 16 storybooks used at each grade level in the SELF curriculum, (2) whether they used the books to teach SEL concepts, and (3) whether or not they used another SEL curriculum in their classroom.

Social Validity

To assess whether teachers thought the SELF curriculum was socially valid, we conducted semi-structured group exit interviews with participating K-1 teachers at each treatment school at the end of the year in which they participated in the study. Interviews included questions about the effectiveness of PD, ease of use of curricular materials, scheduling feasibility, treatment efficacy, lesson appeal, and sustainability.

Design and Analysis

In this three-year cluster randomized trial, kindergarten and first-grade children (n = 1,154) were nested in teachers/classrooms (n = 302), and teachers were nested in schools (n = 52). We randomly assigned schools to the SELF or a BAU comparison condition, and used a multilevel model with three levels (i.e., children, classrooms, schools) to analyze the data on each of the subscales of BRIEF2, DESSA, CAB, SKQ, HTKS, and SELF Vocabulary. Different methods were used to analyze the data collected from the HTKS and SELF Vocabulary measures versus data collected using BRIEF2, DESSA, CAB, and SKQ, as described below.

We used a three-level ANCOVA model to analyze data from HTKS and SELF Vocabulary. Fixed effects included Treatment, Grade, Treatment x Grade (TxG) interaction, coefficients for the child-level, teacher-level, and school-level covariates and Covariate x Treatment (CxT) interactions at the child, teacher, and school levels. The child-level covariate was the class-mean centered pretest score. Following Brincks et al. (2017), the teacher-level covariate was the class mean score centered around the school mean pretest score, and the school-level covariate was the school mean pretest score centered around the grand mean pretest score. The treatment factor was coded -.5 for BAU and .5 for SELF, and the grade factor was coded -.5 for kindergarten and .5 for first grade. Each model included random effects for the child, teacher, and school.

We analyzed the data for each subscale from the BRIEF2, DESSA, CAB, and SKQ using a bivariate multilevel model, with the pretest and the posttest as dependent variables, instead of the multilevel ANCOVA model. At pretest, teachers completed these instruments after being informed of the study purpose during the consent process and, subsequently, of their assignment to condition, information that could have affected teachers' initial assessment of their students. In general, compared to students of teachers in the BAU condition, students of teachers in the SELF condition had less positive average pretest scores on BRIEF2, DESSA, CAB, and SKQ. With this pattern of pretest differences, ANCOVA would have adjusted treatment effects to be more positive than would an analysis of effects from a comparison of posttest means. For this reason, we used the bivariate multilevel model, with pretest and posttest scores as dependent variables. Fixed effects included Treatment, Grade, and TxG. As in the ANCOVA model, the treatment factor was coded -.5 for BAU and .5 for SELF, and the grade factor was coded -.5 for kindergarten and .5 for first grade. Each model included random effects for the child, teacher,

and school for pretest and posttest. At each level, random effects were allowed to correlate over time.

For BRIEF2, DESSA, CAB, and SELF Vocabulary, hypothesis tests were conducted using Benjamin-Hochberg false discovery rate (BHFDR) procedure, with a false discovery rate of .05. We applied this procedure to the set of scales within the BRIEF2, DESSA, CAB, and SELF Vocabulary, respectively. Analyses were conducted using the robust maximum likelihood procedure in M*plus* 8.3. This procedure provides maximum likelihood parameter estimates and standard errors that account for non-normality of the data. Effect sizes were calculated based on Hedges effect size and are subsequently referred to as Hedge's *g*. The adjusted posttest means for SELF Vocabulary and HTKS or the model-estimated posttest means for the other variables were divided by the pooled standard deviation (over the treatment group) for the posttest data in the analytic sample and multiplied by $1 - 3/(4N - 9)$, where *N* is number of posttest scores in the analytic sample.

Results

Descriptive statistics are reported in Table 1 for SELF Vocabulary and HTKS. Tables 2 and 3 provide results of estimation and hypothesis testing for SELF Vocabulary and HTKS, respectively. For the SELF Vocabulary and HTKS measures, CxT interactions were not statistically significant and were deleted from the model; we report effects based on the simplified model in Tables 2 and 3. Wald test statistics are reported in M*plus*; these are reported as *z* in Tables 2 and 3 and subsequent tables containing inferential statistics. Estimated parameters are reported in the *b* columns and standard errors are reported in the *SE* columns. Unadjusted probability (*p*) values are reported in Table 2 and a footnote to the table indicates statistical tests that are significant by the BHFDR procedure. (On the HTKS, only a total score is

obtained; thus, the BHFDR procedure was not required for that measure.) Results in Table 2 indicate that the TxG interaction was not significant for any of the SELF Vocabulary subtests or for the total score. Grade effects were significant for all SELF Vocabulary variables, indicating better performance for children in first grade. Treatment effects were also significant for the four SELF vocabulary variables. Adjusted means for the SELF and BAU groups and Hedge's *g,* respectively, were 15.6, 13.7, and 0.22 for the A subtest; 21.2, 18.2, and 0.16 for the B subtest; 13.3, 12.5, and 0.19 for the C subtest; and 50.1, 44.3, and 0.31for the total score. Results for HTKS in Table 3 indicate a non-significant TxG interaction, as well as non-significant treatment and grade effects.

Descriptive statistics are reported in Table 4 for BRIEF2, CAB, DESSA, and SKQ. On BRIEF2, lower scores indicate better self-regulation. Results in Table 5 indicate that there were no significant treatment, grade, or TxG effects for BRIEF2 pretests. At posttest, the treatment effect was significant for all three subtests and indicated better self-regulation.

Results of estimation and hypothesis testing for CAB are presented in Table 6. For the pretest data, only the grade effect on Internalizing was significant, with smaller Internalizing means for children in first grade. Significant treatment effects were found for all four CAB subscales at posttest.

Table 7 contains inferential results for DESSA subscales. Significant treatment effects were found at pretest for Decision Making, Relationship Skills, Self-Awareness, and Social Awareness, with effect sizes ranging from -0.18 to -0.27. At posttest, significant grade effects were found for Relationship Skills and Social Awareness, with effect sizes of -0.18 and -0.19, respectively. Treatment effects at posttest were significant for all DESSA subtests.

Inferential Results for SKQ are reported in Table 8. The treatment effect was significant at pretest with an effect size of -0.57, indicating that compared to teachers assigned to BAU, teachers assigned to SELF viewed their children as having less pretest knowledge of the concepts taught in SELF. Grade was significant at posttest with an effect size of -0.15. Treatment was significant at posttest with an effect size of 0.65.

In summary, findings indicate that SELF had a positive effect at posttest on self-regulation (BRIEF2 indices), SEL (SKQ, DESSA subscales), general behavioral functioning (CAB subscales), and SEL vocabulary (SELF Vocabulary). There were no significant treatment effects on the HTKS. All posttest comparisons between SELF and BAU indicate that the scores of selected (at-risk) students in the SELF condition were more positive than those of selected students in the BAU condition. Effect sizes ranged from 0.65 (SKQ) to 0.20 (BRI & ERI indices of the BRIEF). Although we did not report pretest adjusted results for BRIEF2, CAB, DESSA, and SKQ, we conducted such analyses and found significant treatment effects for all variables. In general, effect sizes were larger after adjusting for pretest than in the analyses without adjustment because teachers assigned to SELF had less positive views of their students at pretest than did teachers assigned to BAU. We found no moderation of main effects by grade or teacher identification as internalizing versus externalizing.

Treatment Fidelity, BAU, and Social Validity

As measured by the DOPP-lo (adherence), the average percentage of instructional components taught across all lessons observed was 92.09. The mean inter-observer agreement based on 110 lessons scored by two independent observers was 90.61 with a SD of 11.26. As measured by the DOPP-hi (quality of implementation), the mean observational rating for SELF condition teachers was 37.73 out of 42 possible points (90%). Based on 100 observations rated

independently by two members of the research team, the mean IOA was 77.71 with a standard deviation of 18.2.

In surveys of BAU teachers at each grade level conducted following their year of participation in the study, most teachers indicated that few of the 16 books used in the SELF curriculum were available in their classrooms. Across all three years of the study, only one to three books were accessed (i.e., available or read in large groups) in half or more of BAU classrooms. Of the teachers who reported having read a SELF storybook in their classrooms (n = 118), just over half (55.09%) indicated use of the book to teach SEL concepts. Finally, only 15.50% of BAU teachers reported using an explicit SEL curriculum.

During the semi-structured group exit interviews, SELF teachers expressed overall satisfaction with the PD, materials, and the intervention's applicability to their students. Although a few teachers thought additional video examples and practice time would have been helpful, participating teachers generally felt the two-day PD prepared them appropriately to implement SELF in their classrooms. The majority of participants said the curriculum was easy to use, with well-organized materials, interesting books, and engaging lessons relevant to students' lives, e.g., interrupting, dealing with frustration and anger, showing empathy. Teachers also reported observing students applying SELF strategies across school settings, using SEL vocabulary, and discussing how peers should have applied emotion control techniques, e.g., Breathe and Think (BAT). Most teachers across all three years said it was likely to highly likely that they would use the SELF curriculum the following year, although some teachers expressed a preference for teaching lessons solely in a whole-group setting.

Discussion

There is a substantial focus on academic achievement in schools today, often at the expense of students' SEL, despite extensive evidence that school success, especially in the early years, is deeply dependent on successful social-emotional development (Blair & Diamond, 2008; Downer & Pianta, 2006). Further, learning skills to engage self-regulatory processing can contribute significantly to a child's social-cognitive and behavioral functioning (Riggs et al., 2006). As such, we designed a randomized controlled field trial to investigate the efficacy of SELF, a K-1 curriculum to promote student social-emotional development through lessons focusing on language-supported self-regulation and social-emotional competence. We hypothesized that the intervention would improve direct assessments of SEL related vocabulary and teacher-reports of behaviors consistent with EF and social–emotional adjustment, outcomes associated with the prevention of developmental risk for significant emotional and behavioral difficulties.

Outcomes Directly Related to Social-Emotional Learning

We found significant differences in teacher reported SEL-related knowledge between students who were taught the SELF curriculum and their peers in the BAU condition. Findings show that students in the SELF condition learned broad concepts related to self-awareness, self-management, social awareness, relationship management, and responsible decision-making addressed in the SELF curriculum. Since conceptual knowledge is foundational to making effective choices across multiple contexts, it is of consequence that SELF teachers perceived their students as knowledgeable about important SEL concepts.

Our findings also indicate that students taught the SELF curriculum on average showed more positive social-emotional vocabulary development at posttest than did students in the BAU

condition. Being able to identify and label feelings can have a profound effect on the capacity to manage emotions and regulate behavior. Children who have the vocabulary to express their feelings accurately can develop emotional literacy, which is a key component of emotional competence (Joseph & Strain, 2003). Moreover, engaging in dialog about emotional experiences, as encouraged in SELF lessons, strengthens neural integration that can contribute to self-regulation (Greenberg, Kusche, & Riggs, 2004). The findings related to vocabulary development are consistent with significantly better social-emotional competence ratings on the DESSA for students in the SELF group. Specifically, students who received SELF instruction were rated higher by their teachers than were students in the BAU condition on relationship skills, ability to engage in self-management, demonstrating self and social awareness, and being able to solve social problems.

Outcomes Related to Self-Regulation and Behavioral Functioning

In addition to SELF effects on knowledge of SEL concepts, competencies, and related vocabulary, there were also treatment effects on teacher reports of children's self-regulation of behavior, cognition, and emotions as measured by BRIEF2 indices. SELF teachers had more positive perceptions of students' ability to develop multiple solutions to problems and evaluate the consequences of a given response in various social environments, recognize the physiological signs of emotions, and use calming down strategies, each representing higher order processes that govern the use of self-control (see Duckworth, Gendler, & Gross, 2014). It appears, therefore, that the modeling and practice opportunities provided by small-group SELF lessons positively affected foundational cognitive processes that facilitate behavioral self-regulation.

Study outcomes also indicate that SELF affected more general behavioral functioning as measured by the CAB. Teacher reports in the SELF condition evidence a) lower risk associated

with internalizing and externalizing behaviors and b) more positive adaptive behavior associated with social skills and competent functioning within the school setting. These findings are promising, in that subscales of the CAB, like those of the BRIEF2, contain items more distally related to SELF lesson content. As such, the fact that students in the SELF condition were rated more positively by teachers in these categories than students who were not taught SELF indicates the potential of this intervention to affect psycho-educational outcomes important to future school success.

Treatment Fidelity and Social Validity

Noteworthy is the high level of adherence on the part of SELF teachers to delivering SELF lessons as intended (92%) as measured by the DOPP-lo. This fidelity outcome was based on observations of a significant number of teachers across diverse schools over three years of the study. Moreover, DOPP-hi scores reveal more than acceptable instructional quality, heightening confidence about the feasibility of SELF implementation. Teacher feedback during semi-structured group exit interviews was generally positive regarding the PD, curriculum materials, and, importantly, SELF's appeal, ease of use, and value to students. In sum, these findings indicate that SELF is a feasible, useful, and acceptable intervention.

Limitations and Future Directions

Similar to many school- and classroom-based studies, we acknowledge that bias and expectancy effects may exist when classroom teachers who complete outcome measures also implement the intervention. Teachers are uniquely positioned, however, to provide information about their students because of frequent and prolonged contact and the opportunity to observe them in diverse contexts. Although in the current study, findings from the direct assessment of SEL vocabulary by assessors blind to condition serve to strengthen results from teacher report

measures, we acknowledge the need for additional informant sources, such as school counselors and parents, when possible, as well as additional direct assessments, in future research efforts.

In addition to the limitation related to informant sources, we note that participating teachers were informed of their assignment to either the treatment or the control condition prior to completing teacher-report pretest assessments. This could have affected baseline assessments measured by the SKQ and DESSA, measures most closely related to competencies covered in SELF mentioned during the recruitment process. As presented in the **Design and Analysis** section, we did, indeed, find significant baseline differences on multiple subscales and consequently, chose our statistical analyses accordingly.

Of interest for future research is an evaluation of PD implementation to ensure that teachers have a clear understanding of basic theoretical elements (e.g., EF, SR), structural components (e.g., whole-group and small-group activities), and instructional strategies (e.g., dialogic reading, prompting student responses, teaching social-emotional vocabulary) important for enhancing effective treatment delivery (Gerber & Solari, 2005; Guskey, 2003). Relatedly, future researchers could investigate the relation between quality of lesson implementation and teacher characteristics such as classroom organization and discipline style, levels of emotional and instructional support, and attitudes toward students who exhibit at-risk behavior.

In sum, our findings suggest that teachers can increase student behavioral knowledge, requisite vocabulary, and skill development by integrating SEL with academic instruction using curricula such as SELF. We are, therefore, hopeful that SELF and other interventions in which explicit SEL instruction is embedded in academics (e.g., literacy) will continue to be the focus of rigorous investigations involving students who need support beyond what is typically provided. Our findings can help guide future studies and inform practice for students with social-emotional

needs with the goal of improved educational outcomes for all students, particularly students at risk for EBD.

Acknowledgement: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A160136 to The University of Florida. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Al Otaiba, S. (2007). Weaving moral elements and research-based reading practices in inclusive classrooms using shared book reading techniques. *Early Child Development and Care, 174*(6), 575-589.

August, G. J., Hektner, J. M., Egan, E. A., Realmuto, G. M., & Bloomquist, M. L. (2002). The early risers longitudinal prevention trial: Examination of 3-year outcomes in aggressive children with intent-to-treat and as-intended analyses. *Psychology of Addictive Behaviors*, *16*(4S), S27.

August, G. J., Lee, S. S., Bloomquist, M. L., Realmuto, G. M., & Hektner, J. M. (2003). Dissemination of an evidence-based prevention innovation for aggressive children living in culturally diverse, urban neighborhoods: The Early Risers effectiveness study. *Prevention Science*, *4*(4), 271-286.

Backes, B. (2009). NIJ seeks to strengthen the practitioner-researcher bond. *Corrections Today* 71 (4). 78-80.

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*, 899-911.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child development*, *78*(2), 647-663.

Blakemore, S., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3), 296-312.

Bracken, B. A., & Keith, L. K. (2004). *Clinical Assessment of Behavior*. Lutz, FL: Psychological Assessment Resources.

Bradley, R., Doolittle, J., & Bartolotta, R. (2008). Building on the data and adding to the discussion: The experiences and outcomes of students with emotional disturbance. *Journal of Behavioral Education, 17*(1), 4-23. doi:10.1007/s10864-007-9058-6

Brincks, A. M, Enders, C. K., Llabre, M. M., Bulotsky-Shearer, R. J., Prado, G. & Feaster, D. J. (2017) Centering predictor variables in three-level contextual models, *Multivariate Behavioral Research*, 52(2), 149-163. doi:10.1080/00273171.2016.1256753

Cameron, Ponitz, C., McClelland, M. M., Matthews, J. M., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*, 605–619.

Coogle, C. G., Floyd, K. K., & Rahn, N. L. (2018). Dialogic reading and adapted dialogic reading with preschoolers with autism spectrum disorder. *Journal of Early Intervention, 40*(4), 363-379.

Cooper, J., Masi, R., & Vick, J. (2009). Social and emotional development in early childhood: what every policymaker should know. New York: National Center for Children in Poverty (NCCP).

DeTemple, J., & Snow, C. E. (2003). Learning words from books. In S. Stahl, A. van Kleeck & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 16-36). Mahwah, NJ: Erlbaum.

Dodge, K.A., Laird, R., Lochman, J.E., Zelli, A., & the Conduct Problems Prevention Research Group (2002). Multi-dimensional latent construct analysis of children's

social information processing patterns: Correlations with aggressive behavior problems. *Psychological Assessment, 14,* 60-73.

Downer, J. T., Jamil, F., Maier, M. F., & Pianta, R. C. (2012). Implications of information processing theory for professional development of early educators. *Effective early childhood professional development: Improving teacher practice and child outcomes*, 131-155.

Downer, J. T., & Pianta, R. C. (2006). Academic and cognitive functioning in first grade: Associations with earlier home and child care predictors and with concurrent home and classroom experiences. *School Psychology Review, 35*(1), 11-30.

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2014). Self-control in school-age children. *Educational Psychologist*, *49*(3), 199-217.

Eklund, K., Kilpatrick, K. D., Kilgus, S. P., & Haider, A. (2018). A systematic review of state-level social-emotional learning standards: implications for practice and research. *School Psychology Review, 47*(3), 316-326. doi: 10.1710s/spr-2017.0116.v47-3

Elias M., Haynes N. (2008). Social competence, social support, and academic achievement in minority, low-income urban elementary school children. *School Psychology Quarterly*, *23*(4), 474-495.

ESSA (2015). Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).

Feifer, S. G., & Rattan, G. (2007). Executive functioning skills in male students with social-emotional disorders. *International Journal of Neuroscience, 117,* 1565- 1577. doi:10.1080/00207450701239350

Fettig, A., Cook, A. L., Morizio, L., Gould, K., & Brodsky, L. (2018). Using dialogic reading strategies to promote social-emotional skills for young students: An exploratory case study in an after-school program. *Journal of Early Childhood Research, 16*(4), 436-448.

Flynn, K. S. (2011) Developing children's oral language skills through dialogic reading: Guidelines for implementation. *Teaching Exceptional Children, 44,* 8–16.

Gerber, M. M., & Solari, E. J. (2005). Teaching effort and the future of cognitive-behavioral interventions. *Behavioral Disorders*, *30*(3), 289-299.

Gilliam, W. S. & Shahar, G. (2006). Preschool and child care expulsion and suspension: Rates and predictors in one state. *Infants & Young Children, 19*, 228-245.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function Professional Manual.* Lutz, FL: Psychological Assessment Resources, Inc.

Graziano, P., Reavis, R., Keane, S., & Calkins, S. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology, 45,* 3-19. http://dx.doi.org/10.1016/j.jsp.2006.09.002

Greenberg, M. T. (2010). School-Based Prevention: Current Status and Future Challenges. *Effective Education. 2*, 27-52.

Greenberg, M. T., Kusche, C. A., & Riggs, N. (2004). The PATHS curriculum: Theory and research on neurocognitive development and school success. In J. E. Zins, R. P. Weissberg, M. C. Wang. & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (170-188). New York: Teachers College Press.

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*(6-7), 446–474. doi:10.1037/0003-066X.58.6-7.466

Guskey, T. R. (2003). What makes professional development effective?. *Phi delta kappan*, *84*(10), 748-750.

Hemmeter, M. L., Ostrosky, M. M., & Fox, L. (2006). Social and emotional foundations for early learning: A conceptual model for intervention. *School Psychology Review, 35(*4), 583-601.

Joseph, G.E. & Strain, P.S. (2003). Enhancing emotion vocabulary in young children. *Young Exceptional Children*, 6:18-26

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: Findings from an intervention with at-risk kindergarteners. *Language, Speech, and Hearing Services in Schools*, *36*, 17-32.

Kupersmidt, J. B, Griesler, P. C., DeRosier, M. E., Patterson, C. J., & Davis, P. W. (1995). Childhood aggression and peer relations in the context of family and neighborhood factors. *Child Development, 66*, 360-375.

LeBuffe, P. A., Shapiro, V. B., & Naglieri, J.A. (2008). *The Devereux Student Strengths Assessment (DESSA)*. Lewisville, NC: Kaplan Company.

Lochman, J. E., Dunn, S. E., & Klimes-Dougan, B. (1993). An intervention and consultation model from a social cognitive perspective: A description of the Anger Coping Program. *School Psychology Review, 22*(3), 458-471.

Mattison, R. E., Hooper, S. R., & Carlson, G. A. (2006). Neuropsychological characteristics of special education students with serious Emotional/Behavioral disorders. *Behavioral Disorders, 31*(2), 176-188.

McClelland, M. M., & Cameron, C. E. (2012). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Development Perspectives, 6*(2), 136–142.

Meister, D. G., & Melnick, S. A. (2003). National new teacher study: Beginning teachers' concerns. *Action in Teacher Education*, *24*(4), 87-94.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Neuman, S.B. (1996). Children engaging in storybook reading: The influence of access to print resources, opportunity and parental interaction. *Early Childhood Research Quarterly, 11*, 495-514.

Hammill, D. D., & Newcomer, P. L. (1988). *Test of Language Development-Intermediate (TOLD-I: 2)*. Pro-ed.

O'Connor, E., Dearing, E., & Collins, B.A. (2011). Teacher–child relationship trajectories: Predictors of behavior problem trajectories and mediators of child and family factors. *American Educational Research Journal, 48*, 120–162. doi: 10.3102/0002831210365008

Opel, A., Ameer, S. S., & Aboud, F. E. (2009). The effect of preschool dialogic reading on vocabulary in Bangladesi children. *International Journal of Educational Research, 48*(1), 12-20.

Pasi, R. J. (2001). *Higher expectations: Promoting social emotional learning and academic achievement in your school*. New York: Teachers College Press.

Pavri, S. (2004). General and special education teachers' preparation needs in providing social support: A needs assessment. *Teacher Education and Special Education 27*(4), 433–443.

Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of educational psychology*, *94*(1), 23.

Riggs, N. R., & Greenberg, M. T. (2004). The role of neuro-cognitive models in prevention research. In D. Fishbein (Ed.), *The science, treatment, and prevention of anti-social behaviors* (Vol. 2). Kingston, NJ: Civic Research Institute.

Riggs, N. R., Greenberg, M. T., Kusche´, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: *Effects of the PATHS curriculum. Prevention Science, 7*(1), 91-102.

Simsek, Z. C. & Erdogan, N. I. (2015). Effects of the dialogic and traditional reading techniques on children's language development. *Procedia-Social and Behavioral Sciences, 197,* 754-758.

Smith, M. & Smith, K. (2000). "I believe in inclusion, but…": Regular education early childhood teachers' perceptions of successful inclusion. *Journal of Research in Childhood Education, 14,* 161-180. doi:10.1080/02568540009594761.

Stoiber, K. C. (2011). Translating knowledge of social-emotional learning and evidence-based practice into responsive school innovation. *Journal of Educational and Psychological Consultation, 21*(1), 46-55. doi:10.1080/10474412.2011.549039

Sutherland, K. S., & Oswald, D. P. (2005). The relationship between teacher and student behavior in classrooms for students with emotional and behavioral disorders: Transactional processes. *Journal of Child and Family Studies*, *14*(1), 1.

Tangney, J.P., Baumeister, R.F., & Boone, A.L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271–324.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse.

U.S. Department of Education Office for Civil Rights. Civil Rights Data Collection Data Snapshot: School Discipline. Issue Brief Number 1 [CRDC Data] (2014). Retrieved from http://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf

Van der Wissel, A. (1988). Hampered production of words as a characteristic of school failure. *Journal of Learning Disabilities, 21*, 517-518.

Van Kleeck, A., & Vander Woude, J. (2003). Book Sharing with Preschoolers with Language Delays. In van Kleeck, A., Stahl, S. A., & Bauer, E. B. (Eds*.), On Reading Books to Children: Parents and Teachers* (pp. 58-92). Mahwah, NJ: Erlbaum.

Walker, H.M., & Severson, H.H. (1992). *Systematic Screening for Behavior Disorders (SSBD): User's Guide and Administration Manual.* Longmont, CO: Sopris West.

Walker, H. M., Severson, H., Nicholson, F., Kehle, T., Jenson, W R., & Clark, E. (1994). Replication of the Systematic Screening for Behavior Disorders (SSBD) procedure for the identification of at-risk children. *Journal of Emotional and Behavioral Disorders, 2*, 66-77.

Walker, H. M., Severson, H., Todis, B. J., Block-Pedego, A. E., Williams, G. J., Haring, N. G., et al. (1990). Systematic Screening for Behavior Disorders (SSBD): Further validation, replication, and normative data. RASE: *Remedial and Special Education, 11*, 32-46.

Walker, H. M., Seeley, J. R., Small, J., Severson, H. H., Graham, B. A., Feil, E. G., ... & Forness, S. R. (2009). A randomized controlled trial of the First Step to Success early intervention: Demonstration of program efficacy outcomes in a diverse, urban school district. *Journal of Emotional and Behavioral Disorders*, *17*(4), 197-212.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of clinical child and adolescent psychology*, *33*(1), 105-124.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C, & Caulfield, M. (1988). Accelerating language development through picture book reading. *Development Psychology, 24,* 552-559. doi: 10.1037/0012-1649.24.4.552

World Health Organization. (2004). *International statistical classification of diseases and related health problems* (Vol. 1). World Health Organization.

Zelazo, P. D., & Cunningham, W. A. (2007). Executive Function: Mechanisms underlying emotion regulation. In J. Gross (Ed.), *Handbook of emotion regulation.* New York: Guilford.

Zins, J. E., Bloodworth, M. R., Weissberg, M. C., & Walberg, H. J. (2004). The scientific base linking social and emotional learning to school success. In *Building Academic Success on Social and Emotional Learning* (pp. 3-22). New York: Teachers College Press.

Zins, J. E., & Elias, M. J. (2007). Social and emotional learning: Promoting the development of all students. *Journal of Educational and Psychological Consultation*, *17*(2-3), 233-255.

Table 1

Descriptive Statistics by Grade and Condition for Direct Assessments of Vocabulary and Self-Regulation

			SELF						BAU					
			Pretest			Posttest			Pretest			Posttest		
Instrument	Scale	Grade	n	M	SD	n	M	SD	n	M	SD	n	M	SD
SELF Vocab	A	K	325	8.43	5.50	279	13.24	7.00	289	9.04	5.85	273	10.82	5.96
		1	277	15.22	6.23	249	19.17	7.33	239	14.16	6.51	213	16.33	6.18
	B	K	325	13.15	7.90	279	19.00	9.30	289	13.13	7.87	273	15.56	7.96
		1	277	20.49	8.73	249	24.92	9.04	239	19.66	8.78	213	20.82	8.65
	C	K	325	9.25	3.35	279	11.71	3.65	289	8.98	3.20	273	10.73	3.91
		1	277	13.34	3.30	250	15.46	3.15	239	12.96	3.74	213	14.18	3.36
	Total	K	325	30.83	15.12	279	43.95	18.04	289	31.15	15.36	273	37.12	16.07
		1	277	49.05	16.54	249	59.55	18.01	239	46.78	16.84	213	51.33	16.24
HTKS	Total	K	321	39.89	12.57	275	44.13	10.03	291	37.33	15.31	271	42.93	11.78
		1	275	47.03	7.28	251	48.26	5.16	240	45.42	7.72	210	47.54	5.96

Note. A = definition; B = contextual use; C = application; HTKS = Head Toes Knees Shoulders.

Table 2

Summary of Inferential Tests by Subscale and Total Score for SELF Vocabulary Measure

	SELF A					SELF B				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	2.09	0.60	3.49*	<.001	0.29	3.02	0.79	3.83*	<.001	0.33
Grade (G)	1.57	0.55	2.87*	.004	0.22	1.47	0.69	2.12*	.034	0.16
TxG	-1.16	0.91	-1.27	.204	-0.16	-0.06	1.15	-0.06	.955	-0.01
Child Covariate	0.69	0.03	19.92*	<.001		0.63	0.04	17.22*	<.001	
Teacher Covariate	0.69	0.05	13.48*	<.001		0.58	0.06	9.75*	<.001	
School Covariate	0.81	0.11	7.66*	<.001		0.75	0.11	6.66*	<.001	

	SELF C					SELF Total				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	0.75	0.24	3.20*	.001	0.19	5.89	1.20	4.89*	<.001	0.31
Grade (G)	0.90	0.22	4.01*	<.001	0.23	1.99	1.22	1.63	.102	0.11
TxG	0.29	0.33	0.87	.382	0.07	-1.15	2.00	-0.57	.567	-0.06
Child Covariate	0.61	0.03	19.27*	<.001		0.75	0.03	23.69*	<.001	
Teacher Covariate	0.67	0.06	11.95*	<.001		0.75	0.06	13.69*	<.001	
School Covariate	0.71	0.08	8.37*	<.001		0.90	0.08	11.25*	<.001	

Note. **S*ignificant effect by BHFDR; A = definition ; B = contextual use; C = application.

Table 3

Summary of Inferential Tests for the Head Toes Knees Shoulders Direct Assessment

Effect	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-0.05	0.66	-0.07	.945	-0.00
Grade (G)	0.86	0.56	1.55	.122	0.09
TxG	0.42	1.07	0.39	.698	0.05
Child Covariate	0.47	0.05	9.23	<.001	
Teacher Covariate	0.39	0.08	4.58	<.001	
School Covariate	0.58	0.07	8.59	<.001	

Table 4

Descriptive Statistics by Grade and Condition for Teacher-Rated Assessments

			SELF						BAU					
			Pretest			Posttest			Pretest			Posttest		
Instrument	Scale	Grade	n	M	SD	n	M	SD	n	M	SD	n	M	SD
BRIEF2	BRI	K	309	27.52	8.81	280	24.09	8.57	295	26.25	8.87	273	24.95	9.10
		1	271	27.38	7.84	253	23.48	7.77	237	27.36	8.27	207	25.87	8.34
	CRI	K	309	63.40	17.68	280	54.76	17.65	295	59.44	18.02	273	56.72	19.21
		1	271	63.27	16.91	253	53.63	17.38	237	64.45	17.11	207	59.27	16.84
	ERI	K	309	30.34	8.86	280	26.20	8.99	295	28.65	8.57	273	27.13	8.87
		1	271	30.71	8.40	253	26.11	8.04	237	30.43	8.94	207	28.14	8.47
CAB	Communication	K	309	51.62	15.00	280	60.81	15.52	295	54.21	15.46	273	58.15	16.57
		1	271	51.16	14.38	253	60.82	15.08	237	51.69	15.08	207	55.49	15.36
	Externalizing	K	309	62.64	20.17	280	69.64	18.30	295	63.23	19.67	273	65.67	19.66
		1	271	61.18	17.83	253	68.26	17.46	237	62.04	19.38	207	63.64	18.85
	Internalizing	K	309	53.41	11.85	280	60.52	11.57	295	55.32	11.92	273	57.56	12.11
		1	271	51.10	10.93	253	59.09	11.07	237	52.02	12.16	207	54.90	11.83
	Social Skills	K	309	55.49	14.66	280	64.41	14.64	295	57.57	14.64	273	60.72	15.15
		1	271	54.50	13.35	253	62.79	14.15	237	56.29	14.23	207	58.08	14.59
DESSA	Decision Making	K	304	15.41	5.83	278	20.44	6.78	295	16.55	6.71	272	18.29	7.29
		1	270	14.76	5.52	253	19.59	6.59	237	15.92	6.30	207	17.59	6.70
	Relat Skills	K	304	19.25	7.21	278	26.58	8.13	295	21.26	8.18	272	23.70	8.87
		1	270	18.30	7.11	253	24.83	8.40	237	20.54	7.92	207	22.61	8.69
	Self-Awareness	K	304	11.18	5.35	277	16.92	5.87	295	12.51	5.68	272	14.53	6.20
		1	270	11.24	4.81	253	16.63	5.53	237	13.04	5.52	207	14.50	5.55
	Self-Mgt	K	304	20.48	7.55	277	27.07	9.07	295	21.27	8.36	272	24.32	9.52
		1	270	19.54	7.05	253	26.29	8.63	237	21.16	7.98	207	23.58	8.65
	Soc Awareness	K	304	17.64	6.06	278	22.99	7.17	295	18.51	6.83	272	20.52	7.67
		1	270	16.16	5.61	253	21.49	7.07	237	17.81	6.79	207	19.43	7.34
SKQ	Total	K	309	18.05	4.63	280	26.85	5.16	295	21.58	5.57	273	23.71	6.26
		1	271	18.77	4.60	253	26.61	5.40	237	21.20	5.49	206	22.64	5.74

Note. BRIEF2 = Behavior Rating Inventory of Executive Function, Second Edition; Post PD = post professional development; BRI = Behavior Regulation Index; CRI = Cognition Regulation Index; ERI = Emotion Regulation Index; CAB = Clinical Assessment of Behavior; DESSA = Devereux Student Strengths Assessment; SKQ = SELF Knowledge Questionnaire

Table 5

Summary of Inferential Tests for the Behavior Rating Inventory of Executive Function (BRIEF 2) Assessment

	Pretest Results														
	Behavior Regulation Index					Cognitive Regulation Index					Emotional Regulation Index				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	0.68	0.62	1.10	.270	0.08	1.38	1.27	1.09	.277	0.08	1.09	0.73	1.49	.137	0.12
Grade (G)	0.54	0.63	0.86	.389	0.06	2.37	1.56	1.52	.129	0.13	1.07	0.66	1.62	.105	0.12
TxG	-1.27	1.28	-1.00	.320	-0.15	-5.19	3.12	-1.66	.097	-0.30	-1.35	1.31	-1.03	.304	-0.15

	Posttest Results														
	Behavior Regulation Index					Cognitive Regulation Index					Emotional Regulation Index				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-1.68	0.71	-2.35	.019*	-0.20	-4.16	1.61	-2.58	.010*	-0.23	-1.71	0.73	-2.35	.019*	-0.20
Grade (G)	0.21	0.61	0.35	.727	0.03	1.06	1.53	0.69	.489	0.06	0.70	0.66	1.07	.283	0.08
TxG	-1.46	1.21	-1.20	.230	-0.17	-4.20	3.04	-1.38	.167	-0.24	-1.36	1.31	-1.03	.302	-0.16

* Significant by BHFDR

Table 6

Summary of Inferential Statistics for the Clinical Assessment of Behavior Subscales

Pretest Results

	Competence					Externalizing				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-1.47	1.17	-1.26	.209	-0.10	-0.97	1.63	-0.60	.552	-0.05
Grade (G)	-1.59	1.32	-1.20	.230	-0.11	-1.22	1.49	-0.82	.411	-0.06
TxG	2.22	2.66	0.83	.406	0.15	-0.73	2.94	-0.25	.804	-0.04

	Internalizing					Social Skills				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-1.51	0.96	-1.58	.115	-0.13	-1.90	1.22	-1.56	.119	-0.13
Grade (G)	-2.59	1.00	-2.59	.009*	-0.22	-1.15	1.10	-1.04	.296	-0.08
TxG	0.84	1.99	0.42	.671	0.07	0.16	2.18	0.07	.941	0.01

Posttest Results

	Competence					Externalizing				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	4.49	1.47	3.04	.002*	0.29	4.54	1.68	2.70	.007*	0.24
Grade (G)	-1.83	1.31	-1.39	.165	-0.12	-1.80	1.37	-1.31	.189	-0.10
TxG	2.86	2.64	1.08	.279	0.18	0.31	2.77	0.11	.911	0.02

	Internalizing					Social Skills				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	3.89	0.99	3.94	<.001*	0.33	4.47	1.38	3.23	.001*	0.30
Grade (G)	-2.26	1.02	-2.21	.027	-0.19	-2.36	0.96	-2.45	.014	-0.16
TxG	1.42	2.05	0.69	.488	0.12	0.42	1.90	0.22	.826	0.03

*Significant by BHFDR

Table 7

Summary of Inferential Statistics for the Devereux Student Strengths Assessment (DESSA)

	Pretest Results														
	Decision Making					Relationship Skills					Self-Awareness				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-1.09	0.52	-2.10	.036*	-0.18	-2.02	0.68	-2.99	.003*	0-0.27	-1.41	0.44	-3.19	.001*	-0.26
Grade (G)	-0.69	0.55	-1.24	.214	-0.11	-0.82	0.69	-1.18	.237	-0.11	0.24	0.46	0.53	.596	0.05
TxG	0.16	1.09	0.14	.885	0.03	-0.15	1.35	-0.11	.909	-0.02	-0.27	0.94	-0.29	.771	-0.05

	Self-Management					Social Awareness				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-1.14	0.67	-1.70	.089	-0.15	-1.23	0.55	-2.24	.025*	-0.19
Grade (G)	-0.51	0.72	-0.71	.481	-0.07	-1.10	0.59	-1.86	.063	-0.17
TxG	-0.70	1.43	-0.49	.623	-0.09	-0.64	1.16	-0.55	.581	-0.10

*Significant by BHFDR

Table 7, continued

Posttest Results

	Decision Making					Relationship Skills					Self-Awareness				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	2.26	0.70	3.21	.001*	0.33	2.85	0.88	3.23	.001*	0.33	2.54	0.64	3.97	<.001*	0.44
Grade (G)	-0.78	0.51	-1.53	.126	-0.11	-1.53	0.65	-2.37	.018*	-0.18	-0.31	0.51	-0.60	.549	-0.05
TxG	-0.07	1.00	-0.07	.947	-0.01	-0.69	1.26	-0.55	.585	-0.08	0.06	1.01	0.06	.955	0.01

	Self-Management					Social Awareness				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	2.99	0.90	3.33	.001*	0.33	2.50	0.71	3.52	<.001*	0.34
Grade (G)	-0.79	0.72	-1.10	.272	-0.09	-1.43	0.56	-2.54	.011*	-0.19
TxG	0.12	1.43	0.08	.935	0.01	-0.32	1.09	-0.30	.766	-0.04

* Significant by BHFDR

Table 8

Summary of Inferential Statistics for the SELF Knowledge Questionnaire

Pretest Results

Effect	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	-2.90	0.41	-7.00	<.001	-0.57
Grade (G)	0.15	0.46	0.33	0.742	0.03
TxG	1.22	0.84	1.45	0.146	0.24

Posttest Results

Effect	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	3.68	0.63	5.81	<.001	0.65
Grade (G)	-0.82	0.40	-2.05	0.041	-0.15
TxG	1.11	0.79	1.41	0.160	0.20