Efficacy of the social-emotional learning foundations curriculum for kindergarten and first grade students at risk for emotional and behavioral disorders

Ann P. Daunic [a,*], Nancy L. Corbett [a], Stephen W. Smith [a], James Algina [a], Daniel Poling [b], Megan Worth [a], Delaney Boss [a], Emily Crews [a], Jessica Vezzoli [a]

[a] *University of Florida, United States of America*
[b] *Appalachian State University, United States of America*

ARTICLE INFO

ABSTRACT

Researchers have shown that children's social-emotional growth is inextricably connected to academic learning. We developed the Social-Emotional Learning Foundations (SELF) intervention, a Grade K-1 curriculum merging social-emotional learning (SEL) and literacy instruction, 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 (K) and 141 first grade teachers from 52 schools across 11 school districts within one southeastern state. Our student sample ($n = 1154$) 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 participated 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.

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 et al., 2009; World Health Organization, 2004) and that also increase their risk for long-term academic and behavioral problems (O'Connor et al., 2011). Teachers cite behavioral issues as one of their most pressing concerns (Bottiani et al., 2019; Herman et al., 2018; Martin et al., 2012), 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). 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 these problematic behaviors (Elias & Haynes, 2008; Stoiber, 2011).

Moreover, although internalizing disorders may be less apparent to school professionals than non-compliant or disruptive behaviors, they have been shown to be a greater burden than externalizing problems on children's mental health (Kimber et al., 2008). As children first encounter the demands of kindergarten and the primary grades (Zins et al., 2004), the associated challenges may create anxiety related to their academic performance and social interactions. Internalizing behaviors such as anxiety and depression are the most frequently occurring mental health concern in children and adolescents (see Copeland et al., 2014), with a lifetime prevalence reaching approximately 30% (Kessler et al., 2005). Internalizing behaviors negatively impact children's academic performance, social interactions with adults and peers, and physical health (Flook et al., 2005; Merrell & Walker, 2004). These issues, however, tend to receive less attention than externalizing problems in school settings, as children who are relatively withdrawn and quiet create fewer classroom disruptions and management problems for teachers (Weist et al., 2018).

Schools are a prime location for providing preventive services for children at behavioral risk, whether for internalizing or externalizing behaviors (Neil & Christensen, 2009; Weissberg & Cascarino, 2013). 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 teachers indicate that they have received minimal preparation on how to implement effective and efficient practices (Pavri, 2004; Peterson-Ahmad et al., 2018; Smith & Smith, 2000).

1. Social-emotional learning

Social-emotional learning (SEL) is defined as the process by which children acquire and use knowledge and skills to identify and manage emotions, work toward goals, develop positive relationships, and make prosocial choices (Jones & Kahn, 2017), and children's early school success has been shown to depend heavily on successful social-emotional development (e.g., Blair & Diamond, 2008; Downer & Pianta, 2006; Graziano et al., 2007). The Collaborative for Social and Emotional Learning (CASEL) was established to promote the integration of evidence-based SEL as an essential part of P-12 education (Weissberg & Cascarino, 2013). CASEL identified skills shown to provide protective factors for at-risk children, including those from low socioeconomic status backgrounds (Elias & Haynes, 2008). An increasing number of researchers (e.g., Hemmeter et al., 2006; Riggs et al., 2006; Zins et al., 2004) have emphasized 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.

Given that social-emotional growth and academic learning are inextricably connected (Blair & Diamond, 2008; Downer & Pianta, 2006), to neglect either in favor of the other would be counterproductive. Thus, it is conceptually and practically sound to integrate a SEL curriculum to reduce risk for EBD within academic subjects typically taught in the elementary grades, such as literacy and language arts. A proactive, evidence-based approach that provides teachers access to SEL instruction is critical, therefore, for promoting the development of skills such as communication, empathy, problem-solving, and self-management in young children (Eklund et al., 2018). Despite this assertion, in the current context of high stakes evaluations and initiatives (Every Student Succeeds Act [ESSA], 2015), many school professionals continue to focus heavily on developing academic skills to the detriment of SEL.

2. SEL and multi-tiered systems of support

A Multi-Tiered System of Support (MTSS) is a problem-solving framework designed to improve outcomes for all students using a continuum of evidence-based practices matched to student needs. Schools implement MTSS to align academic, behavioral, social, and emotional supports to improve education for all students (Sugai & Horner, 2009). 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). Students at early risk for EBD, however, may not benefit sufficiently from universally delivered programs (Greenberg, 2010). SEL researchers have addressed this issue by providing more targeted instruction (Tier 2) designed specifically for students with or at risk for EBD, and this approach has shown promise, as demonstrated by interventions such as Early Risers (August et al., n.d.), Incredible Years (Webster-Stratton, 1992), and First Steps to Success (Walker et al., 1997).

Early Risers is a 6-week summer program that targets elementary students who are at risk for EBD and who exhibit aggressive and disruptive behaviors and includes a parent component (August et al., 2002). In two studies of K–2 students (August et al., 2002, 2003) Early Risers evidenced positive effects on social outcomes and academic performance, but no discernible effects on externalizing or internalizing behavior. The Incredible Years, delivered in 60, 20–30 min lessons for 18–22 weeks, focuses on developing anger management, social problem solving, getting along with friends, and emotional awareness among pre-school and elementary aged students. Studies (Walker et al., 2009; Webster-Stratton et al., 2004) including children ages 4–8 years with oppositional defiant disorder showed positive effects on externalizing behavior and social outcomes. Finally, First Steps to Success, designed to include 50–60 h of instruction over a 3-month period, helped increase prosocial skills and reduce problem behavior for aggressive children at risk for EBD. Three studies (Nelson et al., 2009; Walker et al., 1997, 2009) with Grade 1–3 students indicated its effectiveness in reducing externalizing behavior and social-withdrawal.

3. The social-emotional learning foundations curriculum

Developed as a Tier 2 intervention with a universal component to promote social-emotional development for primary grade children at early risk for EBD, the Social-Emotional Learning Foundations (SELF) curriculum differs from other targeted interventions

by integrating SEL with literacy instruction to strengthen (a) language related to SEL and (b) self-regulation through small-group lessons taught within the general education setting. Using carefully coordinated materials and pedagogy, SELF lessons focus on developing language-supported self-regulation and social-emotional competence within the framework of literacy related instruction using children's storybooks. The curriculum is organized around SEL topics within five critical competencies identified by CASEL: self-awareness, self-management, social awareness, relationship management, and responsible decision-making (Zins et al., 2004). Thus, SELF simultaneously promotes SEL and literacy, making it feasible for teachers to incorporate social-emotional competencies during academic instruction. (See the Method section for a more detailed description of the curriculum and its implementation.)

4. Underlying theory

The ability to self-regulate both emotions and behavior has been closely linked to neurocognitive mechanisms known as executive function (EF; Blair & Razza, 2007; Riggs & Greenberg, 2004). These mechanisms are thought to contribute significantly to a child's social-cognitive, behavioral, and academic functioning (Blakemore & Choudhury, 2006; Greenberg et al., 2004; Tangney et al., 2004; Zelazo & Cunningham, 2007) and are closely related to the successful development of self-regulation (McClelland & Cameron, 2012). Emerging research indicates that students classified as having social, emotional, or behavioral disorders in schools exhibit EF deficits (Feifer & Rattan, 2007; Mattison et al., 2006). Most importantly, 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 et al., 2004). In sum, 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. We illustrate the connections between intervention components and proximal and distal outcomes in Fig. 1, hypothesizing that SELF will result in positive social-emotional and behavioral outcomes in support of learning.

5. Study purpose

To evaluate the efficacy of the SELF curriculum as compared to a business as usual (BAU) control condition for children at risk for EBD in general education kindergarten and first grade classrooms under routine conditions, we conducted a three-year pretest-posttest cluster randomized efficacy trial. Specifically, we hypothesized that children in the SELF condition would evidence more positive teacher reports of (a) knowledge related to SEL, (b) social-emotional competence, (c) self-regulatory skills, and (d) school adjustment as compared to children in the BAU condition. We also hypothesized that direct assessments of social-emotional vocabulary development and self-regulation would indicate better outcomes for students taught the SELF curriculum as compared to children in the BAU condition. In addition, we explored whether intervention outcomes differed for children in K vs. first grade classrooms and for children identified as at risk for internalizing vs. externalizing behaviors.

6. Method

6.1. 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
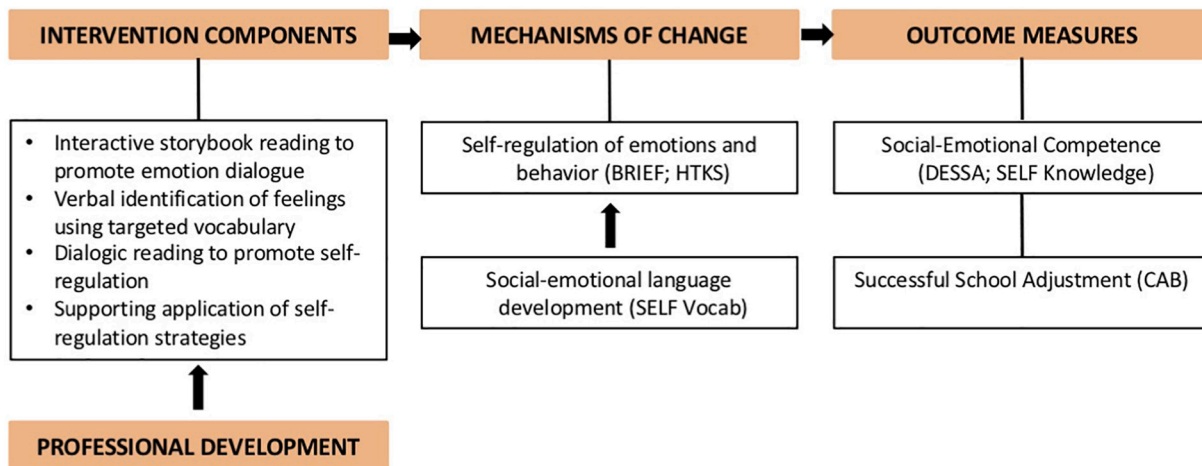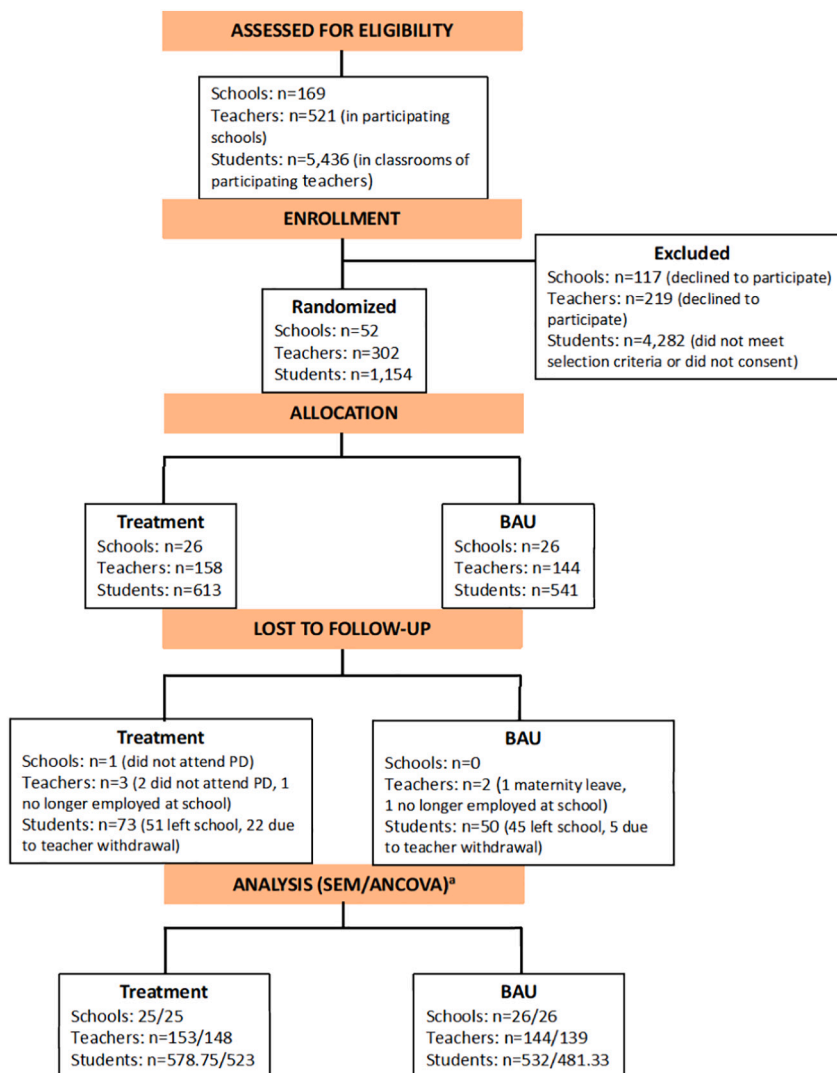

Fig. 1. SELF intervention theory of change.

SELF intervention condition or a BAU condition. All participating teachers, treatment and BAU, were compensated for their involvement in the study. We recruited a distinct group of schools each year, with an average of 5.85 teachers per school across Grades K–1 to ensure an adequate sample size, allowing for typical rates of attrition. Across all participating schools ($N = 52$ schools across 11 school districts), 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.

6.2. 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. We informed parents of our intent to have all children in the classroom screened by their teacher to determine who might benefit most from the SELF intervention if the school were randomly assigned to the treatment condition. Parents were given the opportunity to opt out of the screening process, if desired, by signing and returning the information letter, a procedure approved by the university IRB. A total of 246 students from 302 classrooms across three years of data collection opted out of the screening process. With K–1 class size limited to 18 students (Education Commission of the States, 2009), 246 students constitute 4.5% of total possible participants.

Teachers identified students at possible risk for EBD using the Systematic Screening for Behavior Disorders, 2nd Edition (SSBD; Walker & Severson, 1992), a cost-effective, validated, multiple-stage 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

ASSESSED FOR ELIGIBILITY
Schools: n=169
Teachers: n=521 (in participating schools)
Students: n=5,436 (in classrooms of participating teachers)

ENROLLMENT

Excluded
Schools: n=117 (declined to participate)
Teachers: n=219 (declined to participate)
Students: n=4,282 (did not meet selection criteria or did not consent)

Randomized
Schools: n=52
Teachers: n=302
Students: n=1,154

ALLOCATION

Treatment
Schools: n=26
Teachers: n=158
Students: n=613

BAU
Schools: n=26
Teachers: n=144
Students: n=541

LOST TO FOLLOW-UP

Treatment
Schools: n=1 (did not attend PD)
Teachers: n=3 (2 did not attend PD, 1 no longer employed at school)
Students: n=73 (51 left school, 22 due to teacher withdrawal)

BAU
Schools: n=0
Teachers: n=2 (1 maternity leave, 1 no longer employed at school)
Students: n=50 (45 left school, 5 due to teacher withdrawal)

ANALYSIS (SEM/ANCOVA)[a]

Treatment
Schools: 25/25
Teachers: n=153/148
Students: n=578.75/523

BAU
Schools: n=26/26
Teachers: n=144/139
Students: n=532/481.33

[a] See appendix table A.7 for sample sizes specific to each measure

Fig. 2. CONSORT flow diagram of the cluster-randomized trial.

describing each dimension. The three highest-ranked students on each dimension move through Stage 1 into Stage 2, where teachers typically 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 0.90 for the standardization sample (Walker et al., 1990).

In the current study, we used Stage 1 of the SSBD as a guide for selecting students who might benefit most from instruction because of being at risk for emotional or behavioral difficulties, if the school were randomized to the treatment condition. Stage 1 of the SSBD provides teachers with research-validated descriptions of internalizing and externalizing behavior problems. Teachers evaluated all students in their class (except for students whose parents opted the child out) based on these behavioral dimensions and identified up to five students each for internalizing and externalizing categories. Using the Stage 2 questionnaires, teachers rank-ordered students based on their evaluations of the degree to which identified students' behavior or characteristics aligned with the behavior profiles. The top 3–4 students in the internalizing and the externalizing categories were selected for possible participation in either SELF or BAU. Although instructed that internalizing and externalizing categories were not mutually exclusive, teachers only categorized nine students as at risk for both types of behavior. We excluded these students from subsequent statistical analyses involving internalizing vs. externalizing as a moderator.

6.3. Random assignment and consent procedures

Following recruitment and screening, we solicited active consent from the parents/guardians of the selected students in all participating classrooms. We required that each participating teacher obtain consent for at least two and no more than five students among those designated most at risk for internalizing or externalizing problems through the SSBD procedure. Four teachers (1.3% of total teachers across three years) were unable to obtain parental consent for at least two students and were therefore dropped from the study. To assure compliance with human subject protection, 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. School personnel assigned to BAU were informed they would be offered SELF, including professional development (PD) and related materials, following their year of participation in the study. Two schools responded to this offer, both of which participated in the BAU condition during Year 3. We provided PD and curriculum materials in the summer following their participation. A flow chart that illustrates study participant follow through is presented in Fig. 2.

6.4. Sample description

In Table 1, we describe the sample aggregated across three years. Across demographic categories, we were able to obtain data on an average of 91.88% of the students included; we present these characteristics in Table 2. Chi-square tests comparing BAU and SELF intervention groups were non-significant for all demographic variables, with or without correction for clustering. We present comparisons between BAU and SELF groups on all outcome variables at pretest in Table 3.

6.5. Description of the SELF curriculum

We developed SELF (and accompanying PD) to support students identified as at risk for EBD, whether because of externalizing or internalizing behavioral tendencies. Although SELF includes whole-group lessons taught to all students in the class, SELF's emphasis is on targeted intervention for the students identified as at risk. As such, SELF's small-group lessons provide repeated opportunities to practice SEL skills within the framework of literacy instruction.

We designed SELF small-group lessons to be taught by the classroom teacher during the literacy block or intervention period when at-risk students typically receive tiered instruction. The 52 K and 54 first grade lessons are organized using 16 children's storybooks that address SEL topics within the five critical competencies identified by CASEL. Lessons comprise five units corresponding to the five competencies, and each unit includes 2–4 related and developmentally appropriate topics (e.g., "expressing my feelings.").

Table 1
Sample size for grade and condition across three years.

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

Note. K = kindergarten; BAU = business as usual.
[a] Number of distinct participating districts.

Table 2
Student demographic characteristics for BAU and SELF conditions.

Characteristic	BAU			SELF	
Gender	n	%		n	%
Male	319	62.9		350	61.5
Female	219	37.1		188	38.5
Receiving Free or Reduced Lunch	371	81.0		416	82.4
English Language Learner	14	2.7		24	4.2
IEP or Section 504 Plan	107	21.1		110	19.3
Race					
White (non-Hispanic)	306	60.4		379	66.6
African American	114	22.5		103	18.1
Hispanic	48	9.5		51	9.9
Other Race	39	7.7		36	6.3

Note. IEP = Individualized Education Program; BAU = business as usual.

Table 3
Summary of inferential tests of pretest group differences by outcome measures.

Scale	Subscale	Estimate	SE	z	p	g
BRIEF	BRI	0.70	0.61	1.14	.253	0.08
	CRI	1.07	1.38	0.78	.437	0.06
	ERI	1.07	0.72	1.49	.137	0.12
CAB	COM	−1.58	1.16	−1.36	.174	−0.11
	EXT	−1.00	1.63	−0.62	.538	−0.05
	INT	−1.48	0.94	−1.58	.115	−0.13
	SOC	−2.14	1.20	−1.79	.074	−0.15
DESSA	DM	−1.18	0.51	−2.29	0.022[a]	−0.19
	RS	−2.13	0.67	−3.20	0.001[a]	−0.28
	SA	−1.49	0.44	−3.37	0.001[a]	−0.28
	SM	−1.21	0.65	−1.85	0.065	−0.16
	Soc-A	−1.28	0.54	−2.35	0.190	−0.20
SKQ		−2.94	0.43	−6.80	<.001	−0.58
HTKS		1.66	1.25	1.33	.183	0.14
SELF	A	0.24	0.63	0.38	0.703	0.04
SELF	B	0.46	0.68	0.67	0.502	0.05
SELF	C	0.23	0.33	0.71	0.479	0.06
SELF	Total	0.97	1.43	0.68	0.497	0.06

Note. BRIEF2 = Behavior Rating Inventory of Executive Function, Second Edition; BRI = Behavior Regulation Index; CRI = Cognition Regulation Index; ERI = Emotion Regulation Index; CAB = Clinical Assessment of Behavior; DESSA = Devereux Student Strengths Assessment, DM = Decision Making, RS = Relationship Skills, SA = Self Awareness, SM = Self-Management, Soc-A = Social Awareness; SKQ = SELF Knowledge Questionnaire; HTKS = Head-Toes-Knees-Shoulders; SELF A = definition, SELF B = contextual use, SELF C = application; *n* = 1112 for BRIEF, CAB, and SKQ, 1106 for DESSA, 1127 for HTKS, & 1130 for SELF.
[a] For DESSA, tests with *p* < .05 are significant by BHFDR.

Lessons incorporate instructional strategies that promote children's use of SEL related vocabulary, self-talk, critical thinking, and application of learned concepts. With two exceptions, each topic consists of three lessons designed so that the teacher can introduce a SEL topic to all students in a class and then teach 2–3 additional lessons to students identified as at risk for EBD. 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. Activities in the third lesson are also designed to help students understand story structure, compare and contrast key events in a story, and summarize social-emotional concepts. As such, SELF provides evidence-based literacy instruction integrated within social-emotional learning topics, originating with the storybook read-aloud and culminating with activities that highlight story elements (Baumann & Bergeron, 1993; Lynch & van den Broek, 2007) and foster social-emotional competence.

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 et al., 2018; Opel et al., 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-min 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 et al., 2018).

6.6. Professional development

Prior to implementing the SELF curriculum, all treatment teachers participated in two 6 h days of PD during the fall of their participation year. The first day of PD included (a) a description of the conceptual foundations of SEL, (b) an introduction to SEL competencies, and (c) discussion of teacher opportunities to make connections among the conceptual underpinnings of SELF, their students' social-emotional development, and the SEL instruction they were currently providing. We also introduced the instructional features of SELF lessons; described how SELF objectives and activities align with the national Common Core Standards in reading, speaking, and listening; and clarified the importance of vocabulary acquisition and use specifically for Grades K–1. The second day of PD focused on curriculum implementation and deepening teachers' pedagogical knowledge to help them integrate DR and targeted vocabulary instruction with SEL lesson objectives. Finally, using video examples and discussion activities, teachers practiced each of the three lesson types (i.e., storybook reading with prompts, DR and vocabulary instruction, application activities) with a focus on associated SEL objectives, and reflected on curriculum implementation. During SELF curriculum implementation, graduate research assistants visited treatment teachers' classrooms an average of once per week. Although primarily scheduled to monitor treatment fidelity, these visits also allowed research assistants to answer questions about procedural issues and suggest strategies for addressing general classroom issues (e.g., recommend using positive reinforcement to help manage problem behavior, help solve scheduling concerns).

6.7. Data collection procedures and sources

6.7.1. 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; Gioia et al., 2000), Devereux Student Strengths Assessment (DESSA; LeBuffe et al., 2008), Clinical Assessment of Behavior Teacher Rating Form (CAB-T; Bracken & Keith, 2004), 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 who were blind to condition to administer all direct, individualized assessments (i.e., Head-Toes-Knees-Shoulders, SELF Vocabulary Measure) of child participants within the same time period. The assessors, the majority of whom were former teachers, completed approximately 2–3 h of training on direct assessments prior to administration each year.

6.7.1.1. Teacher-completed assessments. The BRIEF2 evaluates emotional and behavioral self-regulation and contains 60 items comprising nine clinical scales (Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, Task-Monitor, Self-Monitor) that form the Behavior Regulation Index (BRI), Cognitive Regulation Index (CRI), and the Emotion Regulation Index (ERI). Using a 3-point Likert-type scale to indicate *never*, *sometimes*, or *often* for each item, 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 0.957, 0.972, and 0.937, respectively, at pretest and 0.962, 0.976, and 0.947, respectively, at posttest.

The DESSA is a 72-item, standardized, norm-referenced behavior rating scale measuring SEL competencies that serve as protective factors for children in grades K-8. The rater is asked to indicate how often the student engaged in each designated behavior over the previous four weeks using a 5-point scale ranging from *never* (0) to *very frequently* (4). Items are organized into eight conceptually derived scales corresponding to key social-emotional competencies. Students were assessed on the five competencies that correspond to those encompassed 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 0.886, 0.899, 0.897, 0.930, and 0.910, respectively, and at posttest were 0.914, 0.938, 0.941, 0.956, and 0.945, respectively.

The CAB-T 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. Raters use a 5-point Likert-type scale to describe how often a student has recently engaged in a particular behavior, with a rating of 1 indicating *always or very frequently* and a rating of 5 indicating *never*. 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 0.877, 0.970, 0.933, and 0.934 for the four subscales, respectively. Sample Cronbach's alphas at posttest were 0.900, 0.974, 0.942, and 0.947, respectively.

The SKQ, developed by the researchers for the current study, assesses how much students know about concepts taught directly in SELF and related to SEL competencies espoused by CASEL and identified earlier in this paper. 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 0.887 at pretest and 0.927 at posttest.

6.7.1.2. Individually administered assessments. Head-Toes-Knees-Shoulders (HTKS; Ponitz et al., 2009) is a direct measure of self-regulation for children ages 4–6 years and assesses 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 that is required in classroom settings. The HTKS includes 2 training, 4 practice, and 20 testing items across two parts. For each item, participants earn one of three possible scores based on their behavioral response (i.e., 0 = *incorrect*, 1 = *self-corrected without assessor intervention*, 2 = *correct*). Assessors administered each item of Part 1 (i.e., head/toes) and Part 2 (i.e., head/toes & knees/shoulders) for a total possible score of 50. (Summed scores did not include training items.) The HTKS has demonstrated good inter-rater reliability and scoring agreement, test-retest reliability of 0.93 over a three-month period, and predictive validity with school achievement (McClelland & Cameron, 2012). Sample derived Cronbach's alphas for total HTKS scores were 0.949 at pretest and 0.927 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 et al., 2005; Penno et al., 2002). Cronbach's alphas derived from sample data at pretest were 0.802, 0.842, and 0.734 for parts a, b, and c, respectively. At posttest, sample derived Cronbach's alphas were 0.824, 0.863, and 0.770 for parts a, b, and c, respectively. The Cronbach's alpha for the total score at pretest was 0.918 and at posttest was 0.928.

6.7.2. 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 the manualized curriculum, such as covering all lesson content, introducing targeted vocabulary, prompting students as directed, and completing all activities. It consists of lesson-specific checklists of lesson components, and observers record whether the components were included in the lesson or not. The research team, either through in situ observations or by video, completed a DOPP-lo on 14.72% (1185/8048) of all lessons taught (including whole & small group lessons) across all teachers and schools. Pairs of observers conducted 9.23% of the 1185 observations to obtain inter-observer agreement on the DOPP-lo.

The "DOPP-hi" is a high inference instrument that requires the observer to rate criteria 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 (2 items), Development of Language to Support Self-Regulation (6 items), and Quality of Instructional Delivery (6 items). Raters assign a score of 1 (*teacher demonstrated multiple examples of the behavior specified in the item*), 2 (*teacher demonstrated at least one or two clear examples of the behavior specified*), or 3 (*there was little or no evidence that the teacher performed the specified behavior*) to each criterion within each domain.

Research assistants completed 4–6 h 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-hi on 13.88% (1117/8048) of all small-group lessons taught across all teachers in all schools. Pairs of observers conducted 9.57% of the 1117 lesson observations to obtain inter-observer agreement on the DOPP-hi.

DOPP observations accommodated research assistants' and teachers' schedules and were spread across the school year. Observers did not provide feedback or instructional support related to SELF lesson delivery. If teachers were not following the curriculum manual for a particular lesson, the observer would suggest that they adhere to the lesson format and manualized prompts as much as possible, providing specific examples if necessary.

Although teachers differed in how much time they spent for a given lesson in discussions with the targeted students, there were few occasions observed in which teachers did not include all lesson components. To assess intervention dosage, we asked SELF teachers to record the students who were present for each small-group lesson taught.

6.7.3. BAU description

To describe SEL instruction in BAU classrooms, we surveyed teachers at the end of their year of participation in the study. We asked whether they used a specific SEL curriculum in their classrooms, and because the SELF curriculum incorporates commercially-available children's storybooks, we asked about the extent to which the teachers used any of the same books as part of their typical instruction (academic or SEL). For each of the 16 storybooks per grade level, BAU teachers reported whether the books were available in their classroom and whether they were used to teach SEL concepts, specifically in small group settings.

6.7.4. Social validity

To ascertain teacher perceptions of the SELF curriculum's social validity, we gathered informal feedback from teachers at each treatment school ($n = 26$) at the end of the year in which they participated in the study. SELF research assistants met with groups of 1–8 teachers at each school in a 30–45 min feedback session. The researchers asked seven questions (see Appendix A) developed by the investigators, summarized teacher responses using laptop computers, and discussed summaries with project staff. Questions concerned the effectiveness of PD, ease of use of curricular materials, scheduling feasibility, perceived treatment efficacy, lesson appeal, and potential sustainability. All were designed to gather teachers' impressions of the curriculum and potential issues (Stewart et al., 2009)

related to the feasibility of implementing SELF in the general education classroom.

6.8. Design and analysis

In this three-year cluster randomized efficacy trial, K and first-grade children ($n = 1154$) 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. As described below, we used different models to analyze the data collected using the HTKS and SELF Vocabulary measures vs. data collected using BRIEF2, DESSA, CAB, and SKQ. For all measures, we used raw scores in analyses.

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 −0.5 for BAU and 0.5 for SELF, and the grade factor was coded −0.5 for K and 0.5 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 the 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 −0.5 for BAU and 0.5 for SELF, and the grade factor was coded −0.5 for K and 0.5 for first grade. Each model included random effects for the child, teacher, and school for pretest and posttest. At each level, we allowed random effects to correlate over time.

We conducted analyses using the robust maximum likelihood procedure in M*plus* 8.3 (Muthén & Muthén, 1998, 2017). This procedure provides maximum likelihood parameter estimates and standard errors that account for non-normality of the data. The pretest plays different roles in the multilevel ANCOVA model vs. the bivariate multilevel model. The pretest is an exogenous variable in the ANCOVA model and an endogenous variable in the bivariate model. Due to these different roles, the treatment of missing data varies across the models. For the ANCOVA model, the analytic sample comprises cases that include both the pretest and posttest. For the bivariate model, the analytic sample comprises cases that include either the pretest, posttest, or both variables.

For the BRIEF2, DESSA, CAB, and SELF Vocabulary, we conducted hypothesis tests using the Benjamin-Hochberg false discovery rate (BHFDR) procedure, with a false discovery rate of 0.05. We applied this procedure to the set of scales within the BRIEF2, DESSA, CAB, and SELF Vocabulary, respectively. We calculated effect sizes based on Hedges' effect size, subsequently referred to as Hedges' g, where g indicates group difference in standard deviations units. To calculate g, the difference in adjusted posttest means for SELF Vocabulary and HTKS, or the model-estimated posttest means for the other variables, was divided by the pooled standard deviation (over treatment groups) 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.

7. Results

Descriptive statistics for the SELF Vocabulary and HTKS are presented in Table 4. We include missing data counts (percentages) for these measures in Appendix Table B.1. The results of estimation and hypothesis testing for SELF Vocabulary and HTKS, respectively,

Table 4
Descriptive statistics by grade and condition for SELF vocabulary and HTKS.

| | | | 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; BAU = business as usual; K = kindergarten.

are provided in Tables 5 and 6. 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 5 and 6. Wald test statistics are reported in M*plus*; these are reported as *z* in these tables and subsequent tables containing inferential statistics. Estimated parameters are reported in the *b* columns and standard errors are reported in the *SE* columns. We report unadjusted probability (*p*) values in Table 5, 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.) Results in Table 5 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 Hedges' *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 0.31 for the total score. Results for HTKS in Table 6 indicate a non-significant TxG interaction, as well as non-significant treatment and grade effects.

Descriptive statistics are reported in Table 7 for the BRIEF2, CAB, DESSA, and SKQ. On the BRIEF2 measure, lower scores indicate better self-regulation (less risk). We present missing data counts (percentages) for BRIEF2, CAB, DESSA, and SKQ variables in Appendix Table B.2Table B.2.

Results in Table 8 indicate no significant treatment, grade, or TxG effects for BRIEF2 pretests. At posttest the treatment effect was significant for all three subtests, with better self-regulation evidenced for SELF vs. BAU. We computed maximum likelihood estimates (MLEs) from the structural equation model results. Estimates of the SELF mean, BAU mean, and Hedges' *g,* respectively, at posttest were 23.7, 25.4, and − 0.20 for BRI; 54.0, 58.4, and − 0.25 for CRI; and 26.1, 27.8, and − 0.20 for ERI.

Results of estimation and hypothesis testing for CAB are presented in Table 9. 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. MLEs of SELF and BAU means and Hedges' *g,* respectively, were 61.0, 56.5, and 0.29 for Competence; 69.1, 64.6, and 0.24 for Externalizing; 59.9, 56.0, and 0.33 for Internalizing; and 63.9, 59.4, and 0.30 for Social Skills.

Table 10 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. MLEs of SELF and BAU group means and Hedges' *g,* respectively, were 20.1, 17.8, and 0.33 for Decision Making; 25.8, 23.0, and 0.33 for Relationship Skills; 16.9, 14.4, and 0.44 for Self-Awareness; 26.7, 23.7, and 0.33 for Self-Management; and 22.3, 19.8, and 0.34 for Social Awareness.

Inferential results for SKQ are reported in Table 11. The treatment effect was significant at pretest with an effect size of −0.57, indicating that compared to teachers assigned to BAU, SELF teachers rated their children as having less pretest knowledge. Grade was significant at posttest with an effect size of −0.15. Treatment was significant at posttest with MLEs for SELF and BAU equal to 26.8 and 23.1, respectively, and Hedges' *g* equal to 0.65.

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. These analyses are included in Tables B.3-B.6 in the Appendix. In general, effect sizes were larger after adjusting for pretest than in the analyses without adjustment because SELF teacher ratings were less positive at pretest than were those of BAU teachers. Pretest, posttest, and analytic sample sizes across outcome variables are presented in Table B.7

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 indicated 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). We found no moderation of main effects by grade or teacher identification as internalizing vs.

Table 5
Summary of inferential tests by subscale and total score for SELF vocabulary measure.

Effect	SELF A					SELF B				
	b	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	2.09	0.60	3.49	<.001[a]	0.29	3.02	0.79	3.83	<.001[a]	0.33
Grade (G)	1.57	0.55	2.87	.004[a]	0.22	1.47	0.69	2.12	.034[a]	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[a]		0.63	0.04	17.22	<.001[a]	
Teacher Covariate	0.69	0.05	13.48	<.001[a]		0.58	0.06	9.75	<.001[a]	
School Covariate	0.81	0.11	7.66	<.001[a]		0.75	0.11	6.66	<.001[a]	
	SELF C					SELF Total				
Effect	*b*	*SE*	*z*	*p*	*g*	*b*	*SE*	*z*	*p*	*g*
Treatment (T)	0.75	0.24	3.20	.001[a]	0.19	5.89	1.20	4.89	<.001[a]	0.31
Grade (G)	0.90	0.22	4.01	<.001[a]	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[a]		0.75	0.03	23.69	<.001[a]	
Teacher Covariate	0.67	0.06	11.95	<.001[a]		0.75	0.06	13.69	<.001[a]	
School Covariate	0.71	0.08	8.37	<.001[a]		0.90	0.08	11.25	<.001[a]	

Note. SELF A = definition; SELF B = contextual use; SELF C = application.
[a] Significant effect by BHFDR.

Table 6
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	

externalizing.

To interpret the magnitude of effect sizes, the rule of thumb suggested for Cohen's *d* (i.e., 0.20 = small, 0.50 = medium, 0.80 = large) is suitable, because the difference between Hedges' *g* and Cohen's *d* is likely to be very small in sample sizes above 20 (Lakens, 2013). To gain a sense of their practical significance, however, it is preferable to relate effect sizes to those obtained in relevant literature. A meta-analysis (Wilson & Lipsey, 2007) assessing the effects of school-based psychosocial prevention and intervention programs on student maladaptive behavior (e.g., fighting, bullying, classroom disruption, acting out) indicated that the effect size for selected and indicated interventions was 0.29. The study's authors concluded that these effect sizes were both statistically and practically significant, projecting that such programs would lead to a 25%–33% reduction in the level of inappropriate behavior in a typical school (see also Greenberg, 2010). In addition, studies of the universally delivered Strong Start program for children in Grades K–2 with internalizing problems found effect sizes ranging from 0.19–0.38 (Caldarella et al., 2009; Kramer et al., 2010; Whitcomb & Merrell, 2012), and a study of the individually administered Coping Cat intervention for students ages 7–14 years found an effect size of 0.35 for reductions in teacher-reported internalizing problems (Kendall et al., 2008). Although these studies vary in intervention approach and target population, the effect sizes reported provide a context for evaluating the magnitude of effects obtained in the current study.

7.1. Treatment fidelity and dosage

As measured by the DOPP-lo (adherence), the average percentage of instructional components taught across all lessons observed was 92.09. The inter-observer agreement is expressed as a percentage and was calculated using $100(A/[A + B])$ where A is the number of items on which two observers agreed and B the number of items on which the two observers disagreed during a paired observation. The mean inter-observer agreement for the DOPP-lo 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 for the DOPP-hi was 77.71 with a standard deviation of 18.2.

We calculated SELF lesson dosage as a percentage of possible lessons (54 for K; 52 for first grade) received by a target child. The mean percentage of SELF lessons received across both grade levels was 88.4 with a standard deviation of 21.5 (M = 89.1, SD = 19.6 for K; M = 87.7, SD = 23.1 for first grade).

7.2. Description of the business as usual (BAU) comparison condition

Based on grade-specific end-of-year surveys completed by 93.75% of BAU teachers, only 15.50% reported using a formal SEL curriculum. In addition, BAU teachers, on average, read 2–3 of the storybooks used in SELF during their whole group instruction. Few teachers used any of the books in small group settings. Specifically, across all three years of the study, four K teachers (5.40%) reported using a SELF storybook during small group instruction; among the four, the number of storybooks ranged from 1 to 3. Nine first grade BAU teachers (14.06%) reported reading a SELF storybook in a small group setting, with the number of books ranging from one to four. Particularly at the Tier 2 level where targeted instruction for students at risk for EBD was provided through SELF, the BAU condition was sufficiently distinct from the SELF condition in terms of its intentional focus on SEL and targeted instruction using storybooks that were integral to the SELF curriculum.

7.3. Social validity

At each treatment school, 1–8 teachers provided feedback during group sessions about the effectiveness of PD, ease of use of the SELF curriculum, feasibility of implementation and sustainability, appeal to students, and perceived efficacy. For example, one teacher stated that "Kids like the books about anger; they made the connections to their own emotions." Another teacher reported that "I heard a student at recess (say) 'You made me frustrated,' and I heard a student saying the blow-the-anger-out chant to himself." Other examples included the use of vocabulary and use of problem-solving strategies, as follows: "Students are using the (emotion) vocabulary. One student would use the target vocabulary correctly; he would remember from whole group and use it in small group. This wasn't a typically verbal student"; "When I remind about consequences, like 'Do you want to sit on the bench or do you want to sit with your friends?', now they really do stop, and they're like 'Oh yeah, I have to make choices and there are consequences'." Teachers

Table 7

Descriptive statistics by grade and condition for teacher-rated assessments.

Instrument	Scale	Grade	SELF						BAU					
			Pretest			Posttest			Pretest			Posttest		
			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; 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; K = kindergarten; BAU = business as usual. Higher scores on the BRIEF2 subscales indicate more risk; lower scores indicate less risk.

Table 8

Summary of inferential tests for the behavior rating inventory of executive function (BRIEF 2) assessment.

Pretest Results

Effect	Behavior Regulation Index					Cognitive Regulation Index					Emotional Regulation Index				
	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	1.16	1.37	0.85	.395	0.07	1.09	0.73	1.49	.137	0.12
Grade (G)	0.54	0.63	0.86	.389	0.06	2.43	1.53	1.58	.113	0.14	1.07	0.66	1.62	.105	0.12
TxG	−1.27	1.28	−1.00	.320	−0.15	−5.34	3.20	−1.67	.095	−0.30	−1.35	1.31	−1.03	.304	−0.15

Posttest Results

Effect	Behavior Regulation Index					Cognitive Regulation Index					Emotional Regulation Index				
	b	SE	z	p	g	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	−1.68	0.71	−2.35	.019[a]	−0.20	−4.40	1.70	−2.60	.009[a]	−0.25	−1.71	0.73	−2.35	.019	[a] -0.20
Grade (G)	0.21	0.61	0.35	.727	0.03	1.15	1.50	0.77	.442	0.06	0.70	0.66	1.07	.283	0.08
TxG	−1.46	1.21	−1.20	.230	−0.17	−4.42	3.05	−1.45	.147	−0.25	−1.36	1.31	−1.03	.302	−0.16

Note. Higher scores on the BRIEF2 subscales indicate more risk; lower scores indicate less risk.

[a] Significant by BHFDR.

Table 9
Summary of inferential statistics for the clinical assessment of behavior subscales.

Pretest Results

Effect	Competence					Externalizing				
	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

Effect	Internalizing					Social Skills				
	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[a]	−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

Effect	Competence					Externalizing				
	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	4.49	1.47	3.04	.002[a]	0.29	4.54	1.68	2.70	.007[a]	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

[a] Significant by BHFDR.

interviewed indicated they would most likely use the SELF curriculum the year following their participation in the study, but some expressed a preference for teaching lessons solely in a whole-group setting: "I'd do more whole group or all whole group lessons; it's hard to keep up and fit in SELF."

8. Discussion

There is a substantial focus on academic achievement in schools today, often at the expense of students' SEL, despite evidence that school success, especially in the early years, is deeply dependent on successful social-emotional development (Blair & Diamond, 2008; Downer & Pianta, 2006). Furthermore, learning skills that engage self-regulatory processes can contribute significantly to a child's social-cognitive and behavioral functioning (Riggs et al., 2006). As such, we conducted a cluster randomized efficacy trial to investigate SELF. Unique in its design as a Tier 2 K–1 intervention with a universal component, SELF is intended to promote student social-emotional development through small-group and whole-class lessons focused on language-supported self-regulation and the development of social-emotional competence using children's storybooks readily available to classroom teachers. We hypothesized that the intervention would improve (a) direct assessments of social-emotional learning related vocabulary and self-regulation, (b) teacher-reports of behaviors consistent with executive function skill development and social-emotional adjustment, and (c) outcomes associated with the prevention of developmental risk for significant emotional and behavioral difficulties.

8.1. Social-emotional learning and vocabulary development

We found significant differences in teacher reported social-emotional vocabulary development between students who were taught the SELF curriculum and their peers in the BAU condition. This is particularly salient given that having the tools to identify and label feelings can profoundly affect a child's neural capacity to manage emotions and regulate behavior (Blair & Raver, 2015; Greenberg et al., 2004). For example, children entering school who are uncomfortable about meeting the challenges of a new situation can feel overwhelmed. Teaching them vocabulary for labeling their emotions provides the foundation for using self-talk, and self-talk is critical to managing the interplay between emotions and cognitive functioning (Greenberg, 2010). Through its small-group structure, SELF lessons provided self-talk models and multiple opportunities to engage in dialog with teacher and peers about emotional experiences, thereby strengthening neural connections that contribute to self-regulatory processes (Petersen et al., 2015).

8.2. Social-emotional learning and self-regulation

We also found significant differences between students taught SELF and their peers in the BAU condition in teacher reported knowledge of important social-emotional concepts, including self-awareness, self-management, and responsible decision-making, all of which relate to knowledge that is foundational to making effective choices in varying social situations. Furthermore, SELF teachers perceived their students as more capable of cognitive processes, such as recognizing the physiological signs of emotions, evaluating the consequences of a given response, and developing multiple solutions to social problems (see Duckworth et al., 2014). These outcomes

Table 10
Summary of inferential statistics for the Devereux student strengths assessment (DESSA).

Pretest Results

Effect	Decision Making					Relationship Skills					Self-Awareness				
	b	SE	z	p	g	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	−1.09	0.52	−2.10	.036[a]	−0.18	−2.02	0.68	−2.99	.003[a]	0–0.27	−1.41	0.44	−3.19	.001[a]	−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

Effect	Self-Management					Social Awareness				
	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[a]	−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

Posttest Results

Effect	Decision Making					Relationship Skills					Self-Awareness				
	b	SE	z	p	g	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	2.26	0.70	3.21	.001[a]	0.33	2.85	0.88	3.23	.001[a]	0.33	2.54	0.64	3.97	<.001[a]	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

Effect	Self-Management					Social Awareness				
	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	2.99	0.90	3.33	.001[a]	0.33	2.50	0.71	3.52	<.001[a]	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

[a] Significant by BHFDR.

Table 11
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

support the hypothesis that the modeling and practice opportunities provided by small-group SELF lessons enhanced cognitive development foundational to behavioral self-regulation and more competent functioning within the school setting. Importantly, the fact that the internalizing/externalizing behavioral dimension did not moderate treatment effects indicates that SELF lessons were beneficial for all targeted children, regardless of type of behavioral risk. This is noteworthy, as there is a need for effective school-based programming for children with internalizing problems, as well as for children with disruptive behavior that typically draws more attention from teachers (Neil & Christensen, 2009; Weist et al., 2018).

8.3. Study limitations

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. Despite possible bias, teachers provide valuable, if not essential, perspectives about student behavior. Although the current study findings from the direct assessment of social-emotional vocabulary by assessors blind to condition strengthen the results from teacher-report measures, we acknowledge the possibility of teacher bias and the need for additional informant sources, such as school counselors, school psychologists, and parents, and additional direct assessments, in future research efforts. The fact that SELF and BAU students did not differ at posttest on the Head-Toes-Knees-Shoulders assessment in this study underscores the need for further exploration of how findings from teacher-report measures compare to those from direct assessments.

In addition to the limitation related to informant sources, we noted that we informed participating teachers of their assignment to either the treatment or the control condition prior to completing teacher-report pretest assessments. During the recruitment process, we explained to all prospective teachers that the SELF curriculum was designed to provide students with concepts, vocabulary, and skills related to social-emotional learning. Once assigned to condition, therefore, general knowledge about the intervention could have affected how they viewed their students' needs in key social-emotional learning areas based on whether they would subsequently be teaching SELF lessons. This could have affected baseline assessments measured by the SKQ and the DESSA, which are measures most closely related to competencies covered in SELF that were mentioned during the recruitment process. As stated in the Design and Analysis section, we did, indeed, find significant baseline differences on multiple subscales and chose our statistical analyses accordingly. We recommend conducting teacher-reported pretest assessments prior to teacher knowledge of group assignment in future studies.

Finally, our theory of change suggests that proximal outcomes related to the development of social-emotional language and self-regulation may be instrumental to change in more distal outcomes that include lessening risk for EBD and enhancing positive school adjustment. We deemed an exploration of mediation effects as beyond the scope of this paper. Such theoretical questions deserve investigation, and we hope to see them addressed in future studies of the SELF intervention.

8.4. Implications and future directions

As a Tier 2 intervention with a universal component, we developed SELF with a focus on students at risk for EBD. We designed small-group activities to increase opportunities to engage in interactive conversations and role-plays about identifying emotions and feelings, considering the feelings of others, and how to respond in emotionally challenging situations. Theoretically, the opportunities provided in the SELF small-group lessons constitute the mechanism that intensifies social-emotional learning instruction for students at risk. The high level of adherence and instructional quality demonstrated by SELF teachers in the current study supports SELF's feasibility, and the positive feedback from teachers provided during group feedback sessions supports SELF's social validity (e.g., appeal, ease of use, value to students). However, during post intervention feedback sessions, some SELF teachers expressed a preference for teaching the curriculum entirely in a whole-class setting. This could perhaps indicate a failure to grasp the fundamental role of the small-group lessons in promoting social-emotional language and providing students at risk for emotional and behavioral difficulties with multiple opportunities to practice social-emotional interactions. It could also suggest a concern with the added

responsibility of situating Tier 2 social-emotional learning instruction within academic skill development given federal mandates that prioritize academic progress (ESSA, 2015). We recommend, therefore, that researchers consider designing studies that include qualitative inquiry (e.g., verbatim transcriptions of interviews with individual teachers and other school stakeholders, systematically coded for in-depth analyses) to better understand the issues related to implementing small-group, Tier 2 social-emotional learning interventions within the general education classroom.

Relatedly, effective and efficient implementation of social-emotional curricula such as the SELF is particularly complex in schools where a significant number of students exhibit academic and behavioral problems, linguistic differences, and racial/ethnic and economic diversity (Downer et al., 2011). As such, researchers should continue to evaluate PD to ensure that it provides a clear understanding of a social-emotional learning intervention's basic theoretical elements (e.g., executive function, self-regulation), structural components (e.g., whole-group and small-group activities), and pedagogical content knowledge (e.g., dialogic reading vis a vis social-emotional learning, prompting related student responses, teaching social-emotional vocabulary) that enhance instruction (Gerber & Solari, 2005; Guskey, 2003; Shulman, 1987). Finally, follow-up studies to determine whether treatment effects are sustained over time would illuminate whether continued or more intensive intervention is needed to improve the social-emotional functioning of students who are behaviorally at risk.

8.5. The role of school psychologists

In recent decades, the field of school psychology has moved away from the "medical model" of identify-diagnose-treat in favor of more comprehensive, public health-based prevention models of care (Strein et al., 2003). In fact, the National Association of School Psychologists (2020) now specifies that prevention and responsive services are a key tenant of school psychology training and practice under the Services to Promote Safe and Supportive Schools domain. Our study findings, therefore, have relevance for school psychologists who are trained in the developmental, social, emotional, and educational needs and behaviors of students, as well as to methods for addressing them (Ross et al., 2002). Given their background and the field's shift toward preventative services, school psychologists could serve in a variety of capacities to further positive social-emotional development, such as providing support to teachers, particularly for students at risk, by reinforcing social-emotional learning concepts such as those included in SELF when meeting with students.

In sum, our findings suggest that teachers can increase student behavioral knowledge, requisite vocabulary, and skill development by integrating social-emotional learning with academic instruction using curricula such as SELF. We hope that SELF and other interventions in which explicit social-emotional learning 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.

Declaration of Competing Interest

None.

Appendix A. SELF Teacher Feedback Questions

1. Did the two-day professional development sufficiently prepare you to teach SELF? What do you recommend we add or emphasize in the professional development to better prepare teachers to implement the curriculum?
2. To implement SELF, you received a copy of the curriculum, books, Batty the puppet, hard copies of posters, vocabulary cards, and writing prompts. What suggestions do you have that would make the SELF materials more user friendly?
3. Describe the challenges of scheduling three SELF lessons per week. If you teach SELF in the future, how will you schedule the lessons?
4. From your perspective, did you observe any changes in student behavior and/or academics as a result of teaching SELF? If so describe those changes and what makes you think what you observed was influenced by SELF instruction.
5. Talk about a topic/lesson that was particularly relevant or engaging to you and your class. Why was it relevant or engaging?
6. Talk about a topic/lesson that was challenging to teach or challenging for your students to understand. Why do you think it was challenging for you and/or your students?
7. On a scale of 1–10 (1 = not at all likely; 10 = highly likely), how likely are you to use the SELF curriculum next year? Why or why not? What might influence your decision to use SELF in the future?

Appendix B. Missing data

Table B.1
Amount (percent) of missing data across SELF and BAU Groups for HTKS and SELF vocabulary measures.

	Total		SELF		BAU	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
HTKS	27 (2.3%)	147 (12.7%)	17 (2.8%)	87 (14.2%)	10 (1.9%)	60 (11.1%)
SELF A, B, & Total	24 (2.1%)	140 (12.1%)	11 (1.84%)	85 (13.9%)	13 (2.4%)	55 (10.2%)
SELF C	24 (2.1%)	139 (12.0%)	11 (1.8%)	84 (13.7%)	13 (2.4%)	55 (10.2%)

Note. HTKS = Head-Toes-Knees-Shoulders, SELF A = definition; SELF B = contextual use; SELF C = application; BAU = business as usual.

Table B.2.
Amount (percent) of missing data across SELF and BAU Groups for BRIEF, CAB, DESSA subscales and SKQ.

	Total		SELF		BAU	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
BRIEF2 & CAB	42 (3.6%)	141 (12.7%)	33 (5.4%)	87 (14.2%)	9 (1.7%)	60 (11.0%)
DESSA DM, S-A, SA	48 (4.2%)	144 (12.5%)	39 (6.4%)	82 (13.4%)	9 (1.7%)	62 (11.5%)
DESSA RS, SM	48 (4.2%)	144 (12.5%)	39 (6.2%)	82 (13.4%)	9 (1.7%)	62 (11.5%)
SKQ	42 (3.6%)	142 (12.3%)	33 (5.4%)	80 (13.1%)	9 (1.7%)	62 (11.5%)

Note. BRIEF2 = Behavioral Rating Inventory of Executive Function; CAB = Clinical Assessment of Behavior; DESSA = Devereux Student Strengths Assessment; DM = Decision Making; S-A = Self-Awareness; SA = Social Awareness; RS = Relationship Skills; SM = Self-Management; SKQ = Student Knowledge Questionnaire; BAU = business as usual.

Results of ANCOVA analyses

Table B.3
Summary of ANCOVA tests for the behavior rating inventory of Executive function (BRIEF2) assessment.

	Behavior Regulation Index					Cognitive Regulation Index					Emotion Regulation Index				
Effect	b	SE	z	p	g	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	−2.26	0.45	−5.02	<.000[a]	−0.27	−5.27	1.17	−4.49	<.001[a]	−0.29	−2.27	0.64	−3.58	<.001[a]	−0.26
Grade (G)	−0.25	0.36	−0.69	0.489	−0.03	−0.81	0.78	−1.04	0.300	−0.05	−0.03	0.46	−0.08	0.940	−0.01
G x T	−0.38	0.72	−0.53	0.595	−0.05	−0.09	1.60	−0.05	0.957	−0.00	−0.37	0.92	−0.40	0.689	−0.04
Child Covariate	0.76	0.02	34.70	<.000[a]		0.72	0.03	21.34	<.001[a]		0.68	0.03	21.72	<.001[a]	
Teacher Covariate	0.83	0.04	19.69	<.000[a]		0.79	0.05	17.45	<.001[a]		0.72	0.05	13.87	<.001[a]	
School Covariate	0.84	0.10	8.65	<.000[a]		0.77	0.13	5.79	<.001[a]		0.57	0.13	4.27	<.001[a]	

Note. Higher scores on the BRIEF2 subscales indicate more risk; lower scores indicate less risk.
[a] Significant by BHFDR.

Table B.4
Summary of ANCOVA tests for the clinical assessment of behavior subtests.

	Competence					Externalizing				
Effect	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	5.95	0.84	7.09	<.001[a]	0.38	5.34	0.85	6.27	<.001[a]	0.29
Grade (G)	−0.47	0.70	−0.67	0.504	−0.03	−0.85	0.69	−1.24	0.214	−0.05
G x T	0.94	1.37	0.68	0.496	0.06	0.99	1.36	0.73	0.466	0.05
Child Covariate	0.74	0.03	27.40	<.001[a]		0.75	0.02	37.07	<.001[a]	
Teacher Covariate	0.84	0.05	16.67	<.001[a]		0.82	0.05	17.64	<.001[a]	
School Covariate	0.96	0.09	10.44	<.001[a]		0.82	0.07	11.01	<.001[a]	
	Internalizing					Social Skills				
Effect	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	4.73	0.80	5.92	<.001[a]	0.40	6.18	0.85	7.30	<.001[a]	0.42
Grade (G)	−0.47	0.65	−0.72	0.471	−0.04	−1.48	0.60	−2.48	0.013	−0.10
G x T	0.82	1.30	0.63	0.530	0.07	0.32	1.17	0.28	0.782	0.02
Child Covariate	0.65	0.03	20.75	<.001[a]		0.76	0.02	32.29	<.001[a]	
Teacher Covariate	0.73	0.05	15.63	<.001[a]		0.77	0.06	13.86	<.001[a]	
School Covariate	0.58	0.13	4.59	<.001[a]		0.84	0.09	9.08	<.001[a]	

[a] Significant by BHFDR.

Table B.5
Summary of ANCOVA tests for the Devereux student strengths assessment (DESSA).

Effect	Relationship Skills					Self-Awareness					Social Awareness				
	b	SE	z	p	g	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	4.65	0.63	7.39	<.001[a]	0.54	3.85	0.47	8.19	<.001[a]	0.66	3.55	0.54	6.55	<.001[a]	0.48
Grade (G)	−0.96	0.39	−2.47	0.013	−0.11	−0.49	0.33	−1.50	0.134	−0.08	−0.66	0.36	−1.82	0.069	−0.09
G x T	−0.91	0.76	−1.20	0.232	−0.10	0.25	0.65	0.38	0.704	0.07	−0.16	0.72	−0.21	0.830	−0.02
Child Covariate	0.65	0.04	15.77	<.001[a]		0.57	0.04	15.94	<.001[a]		0.73	0.04	19.33	<.001[a]	
Teacher Covariate (L2)	0.67	0.05	12.47	<.001[a]		0.68	0.05	14.81	<.001[a]		0.71	0.05	13.87	<.001[a]	
School Covariate	0.89	0.12	7.30	<.001[a]		0.93	0.12	7.71	<.001[a]		0.84	0.12	6.93	<.001[a]	
T x L2	−0.37	0.11	−3.41	.001[a]		−0.23	0.09	−2.44	0.015		−0.24	0.1	−2.39	0.017	

Effect	Decision Making					Self-Management				
	b	SE	z	p	g	b	SE	z	p	g
Treatment (T)	3.29	0.49	6.72	<.001[a]	0.48	4.06	0.66	6.18	<.001[a]	0.45
Grade (G)	−0.26	0.33	−0.78	0.433	−0.04	−0.37	0.49	−0.76	0.448	−0.04
G x T	−0.23	0.62	−0.37	0.714	−0.03	0.64	0.98	0.65	0.514	0.07
Child Covariate	0.71	0.04	20.04	<.001[a]		0.74	0.03	22.87	<.001[a]	
Teacher Covariate	0.73	0.07	10.50	<.001[a]		0.83	0.07	12.47	<.001[a]	
School Covariate	0.92	0.09	10.47	<.001[a]		0.92	0.10	9.30	<.001[a]	

[a] Significant by BHFDR.

Table B.6
Summary of ANCOVA tests for the SELF knowledge questionnaire.

Effect	b	SE	z	p	g
Treatment (T)	5.68	0.83	6.88	<0.001	1.00
Grade (G)	−0.75	0.29	−2.62	0.009	−0.13
G x T	0.53	0.58	0.92	0.360	0.09
Child Covariate	0.58	0.04	15.30	<0.001	
Teacher Covariate (L2)	0.53	0.06	8.93	<0.001	
School Covariate	0.69	0.18	3.91	<0.001	
T x L2	−0.51	0.12	−4.31	<0.001	

Table B.7
Pretest, posttest, and analytic sample sizes across outcome variables.

Variable	Condition	Student			Teacher			School		
		Analytic Sample	Pretest	Posttest	Analytic Sample	Pretest	Posttest	Analytic Sample	Pretest	Posttest
BRIEF2, CAB	BAU	532	532	480	144	144	136	26	26	26
	SELF	580	580	533	153	153	148	25	25	25
DESSA										
DM, RS, SOCA	BAU	532	532	479	144	144	136	26	26	26
	SELF	578	574	531	153	152	148	25	25	25
SA, SM	BAU	532	532	479	144	144	136	26	26	26
	SELF	577	574	530	153	152	148	25	25	25
SKQ	BAU	532	532	479	144	144	136	26	26	26
	SELF	580	580	533	153	153	148	25	25	25
HTKS	BAU	480	480	480	139	139	139	26	26	26
	SELF	518	518	518	148	148	148	25	25	25
SELF A, SELF B	BAU	482	482	482	139	139	139	26	26	26
	SELF	525	525	525	148	148	148	25	25	25
SELF C	BAU	482	482	482	139	139	139	26	26	26
	SELF	526	526	526	148	148	148	25	25	25

Note. BAU = business as usual; BRIEF2 = Behavior Rating Inventory of Executive Function, Second Edition; CAB = Clinical Assessment of Behavior; DESSA = Devereux Student Strengths Assessment; DM = Decision Making, RS = Relationship Skills, SOCA = Social Awareness, SA = Self Awareness, SM = Self-Management; SKQ = SELF Knowledge Questionnaire; HTKS = Head-Toes-Knees-Shoulders; SELF A = definition, SELF B = contextual use, SELF C = application.

The analytic sample for an analysis comprises cases (students, teachers, & schools) with data in the analysis. For an SEM analysis, sample sizes are different for pretest and posttest because the analytic sample included all cases with a pretest, posttest or both. For an ANCOVA analysis, sample sizes are the same for pretest and posttest because the analytic sample included only cases with a pretest and a posttest.

References

Al Otaiba, S. (2004). 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. https://doi.org/10.1080/0300443042000187086.

August, G., Realmuto, G., & Bloomquist, M. (2020). *Early risers "skills for success": A skill building program for children and their families*. University of Minnesota. https://doi.org/10.1037/t1488-006.

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. https://doi.org/10.1037/0893-164X.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. https://doi.org/10.1023/A:1026072316380.

Baumann, J. F., & Bergeron, B. S. (1993). Story map instruction using children's literature: Effects on first graders' comprehension of central narrative elements. *Journal of Reading Behavior, 25*(4), 407–437. https://doi.org/10.1080/10862969309547828.

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. https://doi.org/10.1017/S0954579408000436.

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology, 66*, 711–731. https://doi.org/10.1146/annurev-psych-010814-015221.

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. https://doi.org/10.1111/j.1467-8624.2007.01019.x.

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. https://doi.org/10.1111/j.1469-7610.2006.01611.x.

Bottiani, J. H., Duran, C. A. K., Pas, E. T., & Bradshaw, C. P. (2019). Teacher stress and burnout in urban middle schools: Associations with job demands, resources, and effective classroom practices. *Journal of School Psychology, 77*, 36–51.

Bracken, B. A., & Keith, L. K. (2004). *Clinical assessment of behavior*. Psychological Assessment Resources.

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. https://doi.org/10.1080/00273171.2016.1256753.

Caldarella, P., Christensen, L., Kramer, T. J., & Kronmiller, K. (2009). Promoting social and emotional learning in second grade students: A study of the Strong Start Curriculum. *Early Childhood Education Journal, 37*, 51–56. https://doi.org/10.1007/s10643-009-0321-4.

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. https://doi.org/10.1177/1053815118797887.

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

Copeland, W. E., Angold, A., Shanahan, L., & Costello, E. J. (2014). Longitudinal patterns of anxiety from childhood to adulthood: The great smoky mountains study. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(1), 21–33. https://doi.org/10.1016/j.jaac.2013.09.017.

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). Erlbaum.

Downer, J. T., Maier, M., & Jamil, F. (2011). Implications of information processing theory for professional development of early educators. In R. Pianta, C. Howes, & B. Hamre (Eds.), *Effective professional development in early childhood education* (pp. 233–244). Brookes.

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. https://doi.org/10.1080/02796015.2006.12087999.

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2014). Self-control in school-age children. *Educational Psychologist, 49*(3), 199–217. https://doi.org/10.1080/00461520.2014.926225.

Education Commission of the States. (2009). Class size: Maximum P-12 class-size policies. Retrieved from http://www.ecs.org/clearinghouse/82/91/8291.pdf.

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. https://doi.org/10.17105/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. https://doi.org/10.1037/1045-3830.23.4.474.

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. https://doi.org/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. https://doi.org/10.1177/1476718X18804848.

Flook, L., Repetti, R. L., & Ullman, J. B. (2005). Classroom social experiences as predictors of academic performance. *Developmental Psychology, 41*(2), 319–327. https://doi.org/10.1037/0012-1649.41.2.319.

Flynn, K. S. (2011). Developing children's oral language skills through dialogic reading: Guidelines for implementation. *Teaching Exceptional Children, 44*, 8–16. https://doi.org/10.1177/004005991104400201.

Gerber, M. M., & Solari, E. J. (2005). Teaching effort and the future of cognitive-behavioral interventions. *Behavioral Disorders, 30*(3), 289–299. https://doi.org/10.1177/019874290503000301.

Gilliam, W. S., & Shahar, G. (2006). Preschool and childcare expulsion and suspension: Rates and predictors in one state. *Infants & Young Children, 19*, 228–245. https://doi.org/10.1097/00001163-200607000-00007.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function professional manual*. 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. https://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. https://doi.org/10.1080/19415531003616862.

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?* (pp. 170–188). Teachers College Press.

Guskey, T. R. (2003). What makes professional development effective? *Phi Delta Kappan, 84*(10), 748–750. https://doi.org/10.1177/003172170308401007.

Hammill, D. D., & Newcomer, P. L. (1988). *Test of language development-intermediate (TOLD-I: 2)*. Pro-ed.

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. https://doi.org/10.1177/1096250611427350.

Herman, K. C., Hickmon-Rosa, J., & Reinke, W. M. (2018). Empirically derived profiles of teacher stress, burnout, self-efficacy, and associated student outcomes. *Journal of Positive Behavior Interventions, 20*(2), 90–100. https://doi.org/10.1177%2F1098300717732066.

Jones, S. M., & Kahn, J. (2017). *The evidence base for how we learn: Supporting students' social, emotional, and academic development – Consensus statements of evidence from the Council of Distinguished Scientists*. Washington, DC: National Commission on Social, Emotional, and Academic Development, The Aspen Institute.

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. https://doi.org/10.1044/0161-1461(2005/003).

Kendall, P. C., Hudson, J. L., Gosch, E., & Flannery-Schroeder, E. (2008). Cognitive behavioral therapy for anxiety disordered youth: A randomized clinical trial evaluating child and family modalities. *Journal of Counseling and Clinical Psychology, 76*(2), 282–297. https://doi.org/10.1037/0022-006X.76.2.282.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593.

Kimber, B., Sandell, R., & Bremberg, S. (2008). Social and emotional training in Swedish schools for the promotion of mental health: An effectiveness study of five years of intervention. *Health Education Research, 23*(6), 931–940. https://doi.org/10.1093/heapro/dam046.

Kramer, T. J., Caldarella, P., Christensen, L., & Shatzer, R. H. (2010). Social and emotional learning in the kindergarten classroom: Evaluation of the Strong Start Curriculum. *Early Childhood Education Journal, 37*, 303–309. https://doi.org/10.1007/s10643-009-0354-8.

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. https://doi.org/10.1111/j.1467-8624.1995.tb00876.x.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863. https://doi.org/10.3389/fpsyg.2013.00863.

LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2008). *The Devereux student strengths assessment (DESSA)*. Kaplan Company.

Lynch, J. S., & van den Broek, P. (2007). Understanding the glue of narrative structure: Children's on-and off-line inferences about characters' goals. *Cognitive Development, 22*(3), 323–340. https://doi.org/10.1016/j.cogdev.2007.02.002.

Martin, N. K., Sass, D. A., & Schmitt, T. A. (2012). Teacher efficacy in student engagement, instructional management, student stressors, and burnout: A theoretical model using in-class variables to predict teachers' intent-to-leave. *Teaching and Teacher Education, 28*(4), 546–559. https://doi.org/10.1016/j.tate.2011.12.003.

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. https://doi.org/10.1177/019874290603100205.

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. https://doi.org/10.1111/j.1750-8606.2011.00191.x.

Merrell, K. W., & Walker, H. M. (2004). Deconstructing a definition: Social maladjustment versus emotional disturbance and moving the EBD field forward. *Psychology in the Schools, 41*(8), 899–910. https://doi.org/10.1002/pits.20046.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus: The comprehensive modeling program for applied researchers: User's guide*.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide: Statistical analysis with latent variables, user's guide*.

National Association of School Psychologists. (2020). *Model for comprehensive and integrated school psychological services*. Author.

Neil, A. L., & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention and early intervention programs for anxiety. *Clinical Psychology Review, 29*(3), 208–215. https://doi.org/10.1016/j.cpr.2009.01.002.

Nelson, J. R., Hurley, K. D., Synhorst, L., Epstein, M. H., Stage, S., & Buckley, J. (2009). The child outcomes of a behavior model. *Exceptional Children, 76*, 7–30.

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. https://doi.org/10.1016/S0885-2006(96)90019-8.

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. https://doi.org/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. https://doi.org/10.1016/j.ijer.2009.02.008.

Pasi, R. J. (2001). *Higher expectations: Promoting social emotional learning and academic achievement in your school*. 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. https://doi.org/10.1177/088840640402700410.

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), 23–33. https://doi.org/10.1037/0022-0663.94.1.23.

Petersen, I. T., Bates, J. E., & Staples, A. D. (2015). The role of language ability and self-regulation in the development of inattentive-hyperactive behavior problems. *Development and Psychopathology, 27*(1), 221–237. https://doi.org/10.1017/S0954579414000698.

Peterson-Ahmad, M. B., Hovey, K. A., & Peak, P. K. (2018). Pre-service teacher perceptions and knowledge regarding professional development: Implications for teacher preparation programs. *The Journal of Special Education Apprenticeship, 7*(2), 3.

Ponitz, C. 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. https://doi.org/10.1037/a0015365.

Riggs, N. R., & Greenberg, M. T. (2004). The role of neuro-cognitive models in prevention research. In D. Fishbein (Ed.), *Vol. 2. The science, treatment, and prevention of anti-social behaviors* (pp. 249–262). Civic Research Institute.

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

Ross, M. R., Powell, S. R., & Elias, M. J. (2002). New roles for school psychologists: Addressing the social and emotional learning needs of students. *School Psychology Review, 31*(1), 43–52. https://doi.org/10.1080/02796015.2002.12086141.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23. https://doi.org/10.17763/haer.57.1.j463w79r56455411.

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. https://doi.org/10.1016/j.sbspro.2015.07.172.

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. https://doi.org/10.1080/02568540009594761.

Stewart, D. W., Shamdasani, P. N., & Rook, D. W. (2009). Group depth interviews: Focus group research. In L. Bickman, & D. J. Rog (Eds.), *The SAGE handbook of applied social research methods* (pp. 589–616). SAGE Publications. https://doi.org/10.4135/9781483348858.

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. https://doi.org/10.1080/10474412.2011.549039.

Strein, W., Hoagwood, K., & Cohn, A. (2003). School psychology: A public health perspective: Prevention, populations, and systems change. *Journal of School Psychology, 41*(1), 23–38. https://doi.org/10.1016/S0022-4405(02)00142-5.

Sugai, G., & Horner, R. H. (2009). Responsiveness-to-intervention and school-wide positive behavior supports: Integration of multi-tiered system approaches. *Exceptionality, 17*(4), 223–237. https://doi.org/10.1080/09362830903235375.

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. https://doi.org/10.1111/j.0022-3506.2004.00263.x.

U.S. Department of Education Office for Civil Rights. (2014). Civil Rights Data Collection Data Snapshot: School Discipline. Issue Brief Number 1 [CRDC Data]. 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. https://doi.org/10.1177/002221948802100815.

Van Kleeck, A., & Vander Woude, J. (2003). Book sharing with preschoolers with language delays. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 58–92). Erlbaum.

Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Severson, H. H., & Feil, E. G. (1997). *First step to success: An early intervention program for antisocial kindergartners*. Sopris West. https://doi.org/10.1177/106342669800600201.

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. https://doi.org/10.1177/1063426609341645.

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. https://doi.org/10.1177/106342669400200201.

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. https://doi.org/10.1177/074193259001100206.

Walker, H. M., & Severson, H. H. (1992). *Systematic screening for behavior disorders (SSBD): User's guide and administration manual.* Sopris West.

Webster-Stratton, C. (1992). *The incredible years. A trouble-shooting guide for parents of children aged* (pp. 2–8).

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. https://doi.org/10.1207/S15374424JCCP3301_11.

Weissberg, R. P., & Cascarino, J. (2013). Academic learning+ social-emotional learning= national priority. *Phi Delta Kappan, 95*(2), 8–13. https://doi.org/10.1177/003172171309500203.

Weist, M. D., Eber, L., Horner, R., Splett, J., Putnam, R., Barrett, S., … Hoover, S. (2018). Improving multitiered systems of support for students with "internalizing" emotional/behavioral problems. *Journal of Positive Behavior Interventions, 20*(3), 172–184. https://doi.org/10.1177/1098300717753832.

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

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. *Developmental Psychology, 24*, 552–559. https://doi.org/10.1037/0012-1649.24.4.552.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*(2), S130–S143. https://doi.org/10.1016/j.amepre.2007.04.011.

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* (pp. 135–158). 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 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?* (pp. 3–22). Teachers College Press. https://doi.org/10.1080/10474410701413145.

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. https://doi.org/10.1080/10474410701413152.