Reducing child problem behaviors and improving teacher-child interactions and relationships: A randomized controlled trial of BEST in CLASS

Kevin S. Sutherland[a],[*], Maureen A. Conroy[b], James Algina[b], Crystal Ladwig[b], Gabriel Jessee[a], Maria Gyure[a]

[a] Virginia Commonwealth University, United States
[b] University of Florida, United States

ARTICLE INFO

ABSTRACT

Research has consistently linked early problem behavior with later adjustment problems, including antisocial behavior, learning problems and risk for the development of emotional/behavioral disorders (EBDs). Researchers have focused upon developing effective intervention programs for young children who arrive in preschool exhibiting chronic problem behaviors; however, Tier-2 interventions that can be delivered by teachers with fidelity in authentic settings are lacking. This study examined the effect of BEST in CLASS, a Tier-2 intervention delivered by teachers, on child problem behavior, teacher-child interactions and teacher-child relationships using a cluster randomized controlled trial design. Participants were 465 children (3–5 year olds) identified at risk for the development of EBDs and their 185 teachers from early childhood programs located in two southeastern states. Significant effects were found across both teacher reported (*ES* ranging from 0.23 to 0.42) and observed child outcomes (*ES* ranging form 0.44–0.46), as well as teacher-child relationships (*ES* ranging from 0.26 to 0.29) and observed teacher-children interactions (*ES* ranging from 0.26 to 0.45), favoring the BEST in CLASS condition. Results suggest the promise of BEST in CLASS as a Tier-2 intervention for use in authentic early childhood classroom contexts and provide implications for future research on transactional models of teacher and child behavior.

1. Introduction

Children displaying problem behavior early in life are more likely to face developmental challenges later in school and into adulthood, including antisocial behavior (Campbell, Spieker, Burchinal, & Poe, 2006), learning problems (Hetzner, Johnson, & Brooks-Gunn, 2010), and risk for later identification of emotional/behavioral disorders (EBDs) (Fanti & Henrich, 2010). During the preschool years, EBDs develop progressively, beginning with early exposure to a number of child (e.g., developmental delays, temperament, conduct problems) (APA, 2000; Benson & Aman, 1999), family (e.g., poverty, substance abuse, violence, poor parenting) (Kaiser, Cai, Hancock, & Foster, 2002; Qi & Kaiser, 2003), and school (e.g., teacher-child relationships, classroom climate) risk factors (Curby, Rimm-Kaufman, & Ponitz, 2009; Pianta et al., 2005). While no single risk factor *causes* EBDs, research suggests that exposure to a higher number and combination of risk factors are likely to result in chronic behavior problems, eventually leading to social, emotional, and behavioral disorders (e.g., Conduct Disorder [CD] or Oppositional Defiant Disorder [ODD]) (Qi & Kaiser, 2003; Webster-Stratton, 1997).

These findings are troubling on many levels. First, a national survey found that preschool age children are expelled at three times the rate of K-12 students (Gilliam, 2005), which is likely linked to their display of chronic problem behaviors. Additionally, the early onset of behavior problems in young children predicts problems in school as well as serious problems in adolescence including drug abuse, juvenile delinquency, violence, and school dropout (Beyer, Postert, Muller, & Furniss, 2012; Campbell et al., 2006; McClelland et al., 2007; Phillips & Shonkoff, 2000). Finally, data from the National Longitudinal Transition Study-2 (NLTS2) indicates that young children who lack social, emotional, and behavioral competence are at a noticeable disadvantage in classroom settings (Markowitz et al., 2006). When children with significant problem behaviors are not identified or treated at an early age, the severity and intensity of their problems often increase, ultimately requiring more intensive services and resources and increasing the likelihood of poor academic outcomes, peer rejection, adult mental health concerns, and adverse effects on their families (Beyer et al., 2012; Dunlap et al., 2006; Marchant, Young, & West, 2004).

* Corresponding author.
  E-mail address: kssuther@vcu.edu (K.S. Sutherland).

Over the past several decades researchers have approached interventions designed to address children's behavior using multi-tiered systems of supports (MTSS), that offer children the necessary supports needed to prevent problem behaviors from occurring and learn the appropriate behaviors and skills needed to be successful (e.g., see Fox, Carta, Strain, Dunlap, & Hemmeter, 2010 for a discussion on tiered systems of support). Within the early childhood field, several evidence-based Tier 2 programs have been developed and implemented within early childhood programs to help address young children's chronic problem behaviors and improve their social-emotional learning skills (e.g., Banking Time; Driscoll & Pianta, 2010; Williford, Wolcott, Whittaker, & LoCasale-Crouch, 2015; Chicago School Readiness Project; Raver et al., 2008; Playing-2-gether; Vancraeyveldt et al., 2015). Although EBPs do exist, there continues to be a need for programs that can be implemented by teachers within authentic early childhood programs to ameliorate young children's problem behaviors. This paper presents findings from a study examining the effect of BEST in CLASS, a Tier-2 intervention program designed to decrease young children's chronic problem behaviors.

1.1. Efforts to prevent and reduce problem behavior

A recent large-scale study (Head Start CARES; Morris et al., 2014) evaluated the effects of enhancements to Head Start for children with social-emotional learning needs, comparing the effects of Preschool PATHS (Domitrovich, Cortes, & Greenberg, 2007), Incredible Years (Webster-Stratton, Reid, & Hammond, 2001), Tools of the Mind (Diamond, Barnett, Thomas, & Munro, 2007) and Head Start (as usual) on a variety of child outcomes, including problem behavior. Although initial positive effects on child social-emotional learning outcomes were found when enhancements to Head Start were included (*ES* range = 0.04–0.29), significant effects on children's problem behavior assessed in the spring of the school year were not found (*ES* range = −0.06–0.02). Another recent study (Schindler et al., 2015) used meta-analytic techniques to examine the effects of three levels of programming on child problem behavior. These researchers reported effect sizes across typical early childhood education programs, broadly focused social-emotional programming (e.g., High Scope curriculum; High Scope Educational Research Foundation, 2014), and intensive focused programming on social-emotional learning (e.g., Preschool PATHS; Incredible Years). Increasing effect sizes on child problem behavior were found across increasing levels of intensity of programming, with intensive focused programming (i.e., Preschool PATHS; Incredible Years) having the greatest effect (*ES* = −0.26). In addition, data suggested that intensive child focused programming, such as Preschool PATHS, had a greater impact on reducing child problem behavior (*ES* = −0.46) than programs that focused solely on changing caregiver (i.e., teacher and/or parents) behavior.

Findings from these two studies highlight the importance of providing targeted interventions for young children who face adversities as well as the need for interventions that can be delivered in classrooms by teachers that effectively reduce young children's problem behaviors. Universal programs such as Preschool PATHS, Incredible Years and Tools of the Mind have demonstrated positive effects on children's social-emotional learning outcomes that may help prepare them for Kindergarten and beyond. However, this level of programming may not provide enough supports to impact the children who continue to exhibit chronic problem behavior despite universal (i.e., Tier 1) supports (Qi & Kaiser, 2004). Tier 2 supports (e.g., Banking Time; Driscoll & Pianta, 2010) that systematically identify children with problem behavior may be necessary to provide the intensity of support needed by these children. In addition to increasing supports to children, many teachers lack the knowledge and skills to ameliorate chronic problem behaviors (Hemmeter, Fox, & Snyder, 2009) and require additional knowledge, skills, and supports to address the needs of these children in their classrooms. As highlighted by Schindler et al. (2015)

this aspect of intervention presents additional challenges related to changing adult behavior to the degree necessary to effectively address and ameliorate young children's severe problem behaviors.

1.2. Teacher-child interactions and problem behavior

Unfortunately, the combination of chronic problem behavior and teachers lack of skills to effectively work with these young children may result in negative, and often coercive, teacher-child interactions which contribute to poor teacher-child relationships. The salience of teacher-child interactions is highlighted by research that suggests that the proximal nature of these interactions predicts greater growth in children's academic achievement than other more distal processes such as the education and training of teachers (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Buyse, Verschueren, Verachtert, Van Damme, & Leuven, 2009; Early et al., 2006; Graves & Howes, 2011; Hamre, 2014). Researchers have found that early childhood teachers often react negatively (e.g., increased restrictions, punitive care) to children who demonstrate problem behaviors (Barnett & Boocock, 1998; Scott-Little & Holloway, 1992). As a result, teachers' interactions with children with problem behavior tend to be less positive than their interactions with children who do not demonstrate problematic behavior, leading to fewer learning opportunities for these children, less time engaged in classroom activities, and missed opportunities for learning critical school readiness skills (Howes & Smith, 1995; Raver & Knitzer, 2002).

The application of transactional theory (Sameroff, 1983) to teacher-child interactions may be useful in helping explain how a child's behavior impacts the teacher and how the teacher's expectations and behavior, in turn, affect the behavior of the child. Sameroff and Mackenzie (2003), discussing the transactional process in relation to parenting, pose the question: "Is inept caregiving an expression of parent inadequacy or a reaction to prior experiences with the child?" (p. 634). This question could be rephrased for teachers of children who display significant problem behavior: Is a lack of effective instructional practices for young children with problem behavior an expression of teacher inadequacy or a reaction to prior experiences with the child? For children with a history of problem behavior, pre-existing negative and coercive patterns of interaction with caregivers may be carried over into the classroom setting (Patterson, Reid, & Eddy, 2002; Wahler & Dumas, 1986), and teacher characteristics such as insufficient classroom management skills, developmentally inappropriate expectations, or a lack of experience handling problem behaviors may contribute to the development or maintenance of coercive interaction patterns (Stormont, Beckner, Mitchell, & Richter, 2005). Over time, children experiencing such patterns of interactions tend to receive fewer instructional opportunities and positive social interactions than their peers, which may contribute to the documented long-term adverse effects for children who exhibit chronic problem behavior at early ages (Doumen et al., 2008).

1.3. Teacher-child relationships and problem behavior

Teacher-child relationships are typically characterized by "closeness", and children and teachers with established trusting relationships may be more likely to put forth more effort in classroom situations (Driscoll & Pianta, 2010). Henricsson and Rydell (2004) report that poor teacher-child relationships tend to be stable over time, as well as have a negative effect on school adjustment. This is particularly relevant for children at-risk for EBDs, as research suggests children who exhibit aggressive and disruptive behavior are more likely to develop negative relationships with their teachers (Ladd & Burgess, 1999). Developing negative relationships with teachers early in school can be particularly harmful as problematic relationships between teachers and students with behavior problems in Kindergarten are associated with academic and behavioral problems through eighth grade

(Hamre & Pianta, 2001). These problematic relationships with teachers may contribute to the documented low rates of positive teacher attention, such as academic interactions, teacher praise, and opportunities to respond for children with or at-risk for EBDs (e.g., Van Acker, Grant, & Henry, 1996; Wehby, Symons, & Shores, 1995).

Teacher-child interactions may affect teacher-child relationships through the moment-to-moment sequences of behavioral exchanges between teachers and children, and transactional theory (Sameroff, 1983) would suggest that over time these sequences of behavior and teacher-child interactions may have an impact on subsequent teacher-child relationships as well. For example, it is likely that challenges presented by children with chronic problem behavior result in their receiving differential rates of desired teacher instructional practices over time (Sutherland & Oswald, 2005). O'Connor and McCartney (2007) found that, even when controlling for child and family factors, both the change in quality of teacher-student relationships across time (from preschool to 3rd grade) as well as the quality of teacher-student relationships in 3rd grade had an effect on students' academic achievement. In fact, these authors found that teacher-student relationships in 3rd grade were stronger predictors of academic achievement than either insecure maternal attachment or peer relationships. Therefore, focusing intervention efforts upon the ongoing interactions between teachers and young children with chronic problem behavior, and their subsequent relationships, appears to be a particularly promising approach to improve the teaching and learning environment for both teachers and their children, and ultimately improving child social, emotional, and behavioral outcomes.

2. BEST in CLASS

BEST in CLASS was designed as a Tier 2 intervention focusing on reducing chronic problem behaviors and improving interactions and relationships between teachers and young children who attend early childhood programs and who are at risk for EBDs due to their display of elevated rates of problem behavior in the classrooms (see Conroy, Sutherland, Vo, Carr & Ogston, 2014; Sutherland, Conroy, Vo & Ladwig, 2015, for a further description of BEST in CLASS). Children are systematically screened for risk for EBDs and teachers are trained to focus their use of the BEST in CLASS practices with the identified children during ongoing classroom instruction. To effectively deliver the intervention, teachers are initially trained in the manualized program and then provided weekly practice-based coaching (see Method section for further description of the program) in order to increase their quantity and quality of use of practices with focal children. Based on three theoretical frameworks, BEST in CLASS integrates behavioral (Skinner, 1953), transactional (Sameroff, 1983, 1995), and ecological (Bronfenbrenner, 1979, 2005) theories. To illustrate, BEST in CLASS is comprised of key instructional practices based upon behavioral principles (i.e., Skinner) that promote effective teacher-child interactions, while recognizing the transactional nature of social interchanges and how affecting behavior and transactions (e.g., improving teacher-student interactions; i.e., Sameroff) can influence various subsystems, including both more proximal processes (e.g., improving teacher-child relationships; i.e., Bronfenbrenner) as well as the child's broader ecology (e.g., the classroom environment; i.e., Bronfenbrenner). While initial studies have highlighted the promise of the program on improving child (Conroy et al., 2014; Conroy et al., 2015) and teacher (Conroy et al., 2014; Conroy et al., 2015; Sutherland et al., 2015) outcomes, BEST in CLASS has yet to be tested in a more rigorous randomized controlled trial.

2.1. Study aims and hypotheses

The first aim of the current study was to examine the effect of BEST in CLASS on child problem behavior. Previous research (Conroy et al., 2014; 2015) with less rigorous designs suggested that BEST in CLASS

reduced observed child problem behavior; therefore, we hypothesized reductions in teacher-reported and observed child problem behavior in the current study.

The second aim of the study was to examine the effect of BEST in CLASS on teacher-child interactions. Research has demonstrated the importance of positive teacher-child interactions (e.g., Doumen et al., 2008; Early et al., 2006) and increasing positive interactions may be even more important for young children who exhibit problem behaviors (Raver & Knitzer, 2002). Previous research (Conroy et al., 2015) has shown that BEST in CLASS had positive effects on teacher-child interactions; therefore, we hypothesized reductions in negative teacher-child interactions and increased positive teacher-child interactions.

The final aim of the study was to examine the effect of BEST in CLASS on teacher-child relationships. Similar to teacher-child interactions, research has demonstrated the importance of positive teacher-child relationships (e.g., Driscoll & Pianta, 2010; Henricsson & Rydell, 2004) and the quality of teacher-child relationships may be even more salient for young children with problem behavior (Hamre & Pianta, 2001). Given the positive effect of BEST in CLASS on observed teacher-child interactions (Conroy et al., 2015) we hypothesized that BEST in CLASS teachers would report greater levels of closeness and reduced levels of conflict with children than their counterparts in comparison classrooms.

3. Method
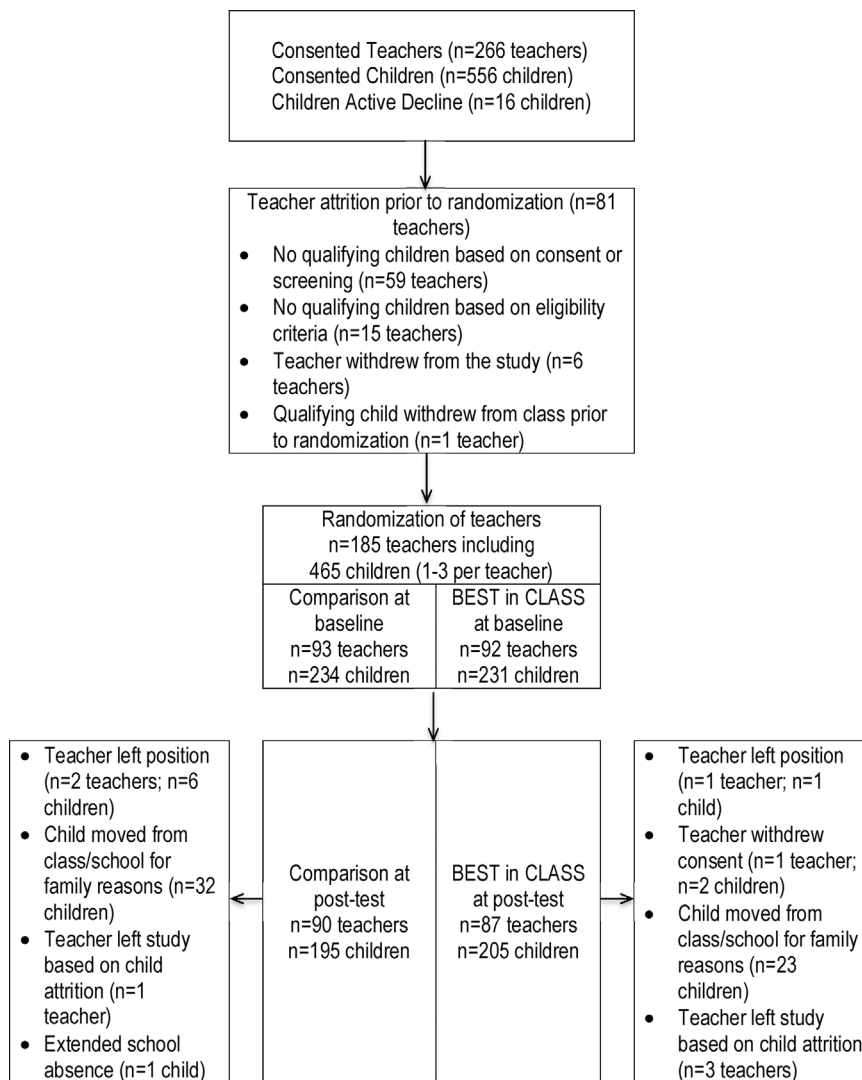
3.1. Design and experimental procedures

This study was a multi-site cluster randomized trial (Spybrook et al., 2011) conducted over four years in two southeastern states. Children were nested in teachers' classrooms and teachers were nested in sites ($n = 78$); each site was an early childhood program center or a school. For sites in which more than one teacher was consented (n = 43), teachers were randomly assigned from within their site (i.e., school or center) to the BEST in CLASS or to a business-as-usual (BAU) comparison condition. Teachers in sites in which one teacher was consented were assigned by simple random assignment to treatment or BAU. This process was repeated each of the four years with a new set of teachers and children participating each year. To minimize contamination across conditions, teachers were provided information about their roles in the study and the importance of not discussing participation with any other teachers in their school or program.

All study activities were approved by human subject protection boards. Obtaining teacher consent and screening of child participants began approximately one month after the beginning of school in order to allow the teachers to familiarize themselves with the children in their classrooms. Once teacher consent was obtained and children were screened to determine eligibility for participation and caregiver consent was received, pretest measures were completed. Following completion of pretest measures, randomization occurred (see Fig. 1 for study flow). Teachers in the BEST in CLASS condition received the BEST in CLASS teacher training, and the following week began receiving weekly practice-based coaching (see Method). In April of each year posttest measures were completed with all participating teachers and children. Although each classroom included a lead teacher and a teacher's assistant, only lead teachers in each classroom participated in the study and all teachers and children only participated for one year.

3.2. Setting and participants

To recruit teacher participants, the researchers approached early childhood programs in their region. All programs agreed to participate and assisted researchers in the recruitment of teachers. The majority of the early childhood programs (96%) participating in the study were federally or state-funded serving children who were from under-represented populations and income eligible (e.g., Head Start). The

```
┌─────────────────────────────────────────┐
│ Consented Teachers (n=266 teachers)     │
│ Consented Children (n=556 children)     │
│ Children Active Decline (n=16 children) │
└─────────────────────────────────────────┘
```

Teacher attrition prior to randomization (n=81 teachers)
- No qualifying children based on consent or screening (n=59 teachers)
- No qualifying children based on eligibility criteria (n=15 teachers)
- Teacher withdrew from the study (n=6 teachers)
- Qualifying child withdrew from class prior to randomization (n=1 teacher)

Randomization of teachers
n=185 teachers including
465 children (1-3 per teacher)

Comparison at baseline	BEST in CLASS at baseline
n=93 teachers	n=92 teachers
n=234 children	n=231 children

Teacher left position (n=2 teachers; n=6 children)
- Child moved from class/school for family reasons (n=32 children)
- Teacher left study based on child attrition (n=1 teacher)
- Extended school absence (n=1 child)

Comparison at post-test	BEST in CLASS at post-test
n=90 teachers	n=87 teachers
n=195 children	n=205 children

- Teacher left position (n=1 teacher; n=1 child)
- Teacher withdrew consent (n=1 teacher; n=2 children)
- Child moved from class/school for family reasons (n=23 children)
- Teacher left study based on child attrition (n=3 teachers)

remaining 4% of programs (*n* = 4) were locally or privately funded (e.g., YMCA). Classrooms were located within local school district elementary schools or early childhood education centers. The mean number of children per classroom was 17.55, with the number of adults per classroom averaging 2.21.

3.2.1. Teachers

A total of 185 teachers participated in the study, with 92 in the BEST in CLASS intervention group and 93 in the comparison group (see Table 1). Teachers were eligible for inclusion if they met the following criteria: 1) taught in an early childhood classroom serving children ages 3–5 years, 2) served at least one child identified as being at risk for EBDs, and 3) consented to participate. There were no significant differences between the BEST in CLASS intervention group and the comparison group on the teacher demographic variables in Table 1.

3.2.2. Children

A total of 465 children participated in the study, with 231 in the BEST in CLASS intervention group and 234 in the comparison group (see Fig. 1). Children who met the following criteria were eligible for participation: 1) enrolled in a participating teacher's classroom, 2) identified as at risk for an EBD through the *Early Screening Project* (*ESP*; Feil et al., 1995), 3) identified as functioning within the normal range in cognition according to the *Battelle Developmental Inventory, Second Edition Screener* (*BDI II Screener*, Newborg, 2005), 4) proficient in

English, and 5) caregiver consent. To determine eligibility for participation, teachers nominated five children in their classroom who engaged in chronic problem behavior. Caregiver consent was then obtained and systematic screening for risk for EBDs took place using the ESP Stages 1 and 2. Additionally, the *BDI II Screener* was administered at that time. Finally, if applicable, children were screened for proficiency in English. After screening, 1–3 children per classroom were selected to participate in the study, depending upon returned caregiver consents, outcomes on the BDI II and the most elevated scores on the ESP. There were no significant differences between the intervention and comparison groups on the child demographic variables in Table 1. In addition there were no significant differences on the child demographic variables between children who completed the study and those who withdrew.

3.3. Measures

3.3.1. Early screening project (ESP)

The ESP (Feil et al., 1995) is a teacher report screening tool designed to identify young children at risk for EBD. The ESP uses a multiple gating format; in Stage 1 teachers rank the top 5 children in their classroom by externalizing behaviors. Following caregiver consent, teachers rated children identified in Stage 1 in subscale areas (i.e., Critical Events Index, Aggressive Behavior Scale, Adaptive Behavior, and Maladaptive Behavior). Consented children with the highest risk

Table 1
Participant Demographics.

			Comparison	BEST in CLASS	Total
Teachers	N		93	92	185
	Race/Ethnicity	African American	47 (50.54%)	41 (44.56%)	88 (47.57%)
		Caucasian	42 (45.16%)	44 (47.83%)	86 (46.49%)
		Hispanic	2 (2.15%)	3 (3.26%)	5 (2.70%)
		Asian	1 (1.08%)	1 (1.09%)	2 (1.08%)
		Other	1 (1.08%)	2 (2.17%)	3 (1.62%)
		No Report	0	1 (1.09%)	1 (0.54%)
	Gender	Female	90 (96.77%)	91 (98.91%)	181 (97.84%)
		Male	3 (3.23%)	1 (1.09%)	4 (2.16%)
		No Report	0	0	0
	Education	High School	4 (4.30%)	1 (1.09%)	5 (2.70%)
		Associates	28 (30.11%)	28 (30.43%)	56 (30.27%)
		Bachelors	35 (37.63%)	37 (40.22%)	72 (38.92%)
		Masters	25 (26.88%)	23 (25%)	48 (25.95%)
		Doctoral	0	1 (1.09%)	1 (0.54%)
		Other	1 (1.08%)	2 (2.17%)	3 (1.62%)
		No Report	0	0	0
	Yrs Teaching	Mean (SD)	12.96 (9.99)	11.35 (8.97)	12.09 (9.49)
Children	N		234	231	465
	Race/Ethnicity	African American	156 (66.67%)	152 (65.81%)	308 (66.24%)
		Caucasian	42 (17.95%)	37 (16.02%)	79 (16.99%)
		Hispanic	10 (4.27%)	11 (4.76%)	21 (4.52%)
		Asian	1 (0.43%)	0	1 (0.22%)
		Native American	0	1 (0.43%)	1 (0.22%)
		Other	16 (6.84%)	16 (6.93%)	32 (6.88%)
		No Report	10 (4.27%)	16 (6.93%)	26 (5.59%)
	Gender	Male	154 (65.81%)	147 (63.64%)	301 (64.73%)
		Female	80 (34.19%)	84 (36.36%)	164 (35.27%)
		No Report	0	2 (0.87%)	2 (0.43%)
	Age at Entry	Mean (SD)	4.36 (.53)	4.29 (.53)	4.32 (.53)

scores on the Stage 2 subscale participated in the study. Feil et al. (1995) reported 6-month test-retest reliability coefficients for the four subscales ranging from 0.74 to 0.90, and the Stage 2 subscale also evinced concurrent validity, with correlations ranging between 0.44 and 0.88 with the externalizing scale of the Caregiver-Teacher Report Form (C-TRF; Achenbach & Rescorla, 2000).

3.3.2. Battelle developmental inventory, second edition screener (BDI-II)

The BDI-II (Newborg, 2005) is a developmental screener for children from birth to age 8 to identify children at risk for developmental delays. In the current study, the Adaptive, Communication, and Cognitive subscales of the BDI-II were administered to determine if children performed within the normal range. Reported reliability coefficients for a normative sample, ages ranging from 0 to 95 months, were calculated by the split-half method and corrected using the Spearman-Brown formula. Average coefficients for the Adaptive, Communication, and Cognitive were 0.90, 0.95, and 0.93 respectively. Test-retest stability was conducted on two age groups, 2 and 4 year old children. Pearson correlations for Adaptive, Communication, and Cognitive were 0.89, 0.89, and 0.88 for the 2-year-old sample and 0.92, 0.88, and 0.87 for the 4-year-old sample (Newborg, 2005).

3.3.3. Caregiver teacher report form

The C-TRF (Achenbach & Rescorla, 2000) is a 100 item teacher-report instrument used to assess problem behavior in children ages one and a half to five. Items are scored on a 3-point Likert scale with responses ranging from *never true (0)* to *often true* (2) and used to inform three key subscales (i.e., externalizing behavior, internalizing behavior, and total problems). In the current sample, Cronbach's alpha demonstrated a high level of internal consistency for all subscales, ranging between 0.86 and 0.95 for pretest and above 0.90 for posttest for the externalizing, internalizing, and total problems subscales.

3.3.4. Social skills improvement system-rating scale

The SSIS-RS (Gresham & Elliott, 2008) is a 76 item teacher-report

measure, allowing for the evaluation of social skills and problem behaviors of children. Each item on the SSIS-RS is rated on a 4-point frequency scale, with responses ranging from 0 (*Never*) to 3 (*Almost Always*). Items are grouped into two subscales: social skills (e.g., completes tasks without bothering others) and problem behaviors (e.g., talks back to adults). For the current sample, internal consistency was acceptable with Cronbach's alpha equal to 0.94 and 0.95 for Social Skills at pre- and posttest waves, respectively and 0.90 and 0.92 for Problem Behavior.

3.3.5. Individualized classroom assessment scoring system

The inCLASS (Downer, Booren, Lima, Luckner, & Pianta, 2010; Downer, Sabol, & Hamre, 2010) is an observational measure assessing children's interactions with teachers, peers, and tasks in the classroom within three domains: Teacher Interactions (positive engagement with teacher, teacher communication, and teacher conflict), Peer Interactions (peer sociability, peer communication, peer assertiveness, and peer conflict) and Task Orientation (engagement within tasks, self-reliance, and behavior control). Observations were conducted for each focal child via three 15 min cycles. Scores for each dimension, ranging from one to seven, are determined by observations of specific child behavior present during each cycle. Final scores for each domain are calculated by taking the mean of each domain across all three cycles. Reliability coefficients for each of the inCLASS domains range from 0.72 to 0.92 (Downer, Booren et al., 2010; Downer, Sabol et al., 2010). Prior to administration all observers obtained a reliable score on a certification exam with a minimum standard of 80% mastery across all indicators. In addition, all observers were recertified annually and recalibrated semi-annually to minimize observer drift. Inter-observer agreement estimates were collected on 21.6% of all inCLASS observations using a secondary observer. The mean reliability estimate for all years of the study was 93.26%.

Table 2
BEST in CLASS Instructional Practice Modules.

Module/Strategy	Content
1. Rules	Provides information on how to create and implement high quality rules and use them intentionally, more frequently, and in a targeted fashion with focal children throughout activities during the school day.
2. Precorrection	Provides information on how to set and communicate high expectations to focal children to prevent challenging behaviors before they occur during routine classroom activities.
3. Opportunities to Respond	Provides information on how to engage focal children by providing them with frequent high quality opportunities during instructional activities
4. Behavior Specific Praise	Provides information on how to increase the frequency and targeted use of specific praise with focal children during instructional activities.
5. Corrective Feedback	Provides information on how to increase the use of feedback to an incorrect response or behavior demonstrated by the focal child with the intention of teaching the correct/appropriate behavior or response to the child.
6. Instructive Feedback	Provides information on how to expand the focal child's engagement by delivering additional specific instruction and information following child's correct answer or response.
7. Linking and Mastery	Provides information on how to combine and use instructional practices sequentially while interacting with focal child.

3.4. Teacher-child interactions direct observation system

The TCIDOS-RV2.1 (Conroy et al., 2015; Sutherland, Conroy, Abrams, Vo & Ogston, 2013) is a partial-interval observational coding system developed by the authors to record and code teacher behavior, child behavior and interactions of teachers and children during the implementation of BEST in CLASS. Using this system, trained coders (coaches and research assistants) observed teachers' use of the BEST in CLASS practices, teacher-child interactions, and behaviors displayed by the focal children, although results in this study only focused on teacher-child interactions (positive and negative), child disruptive behaviors [disruption/aggression/defiance (DAD)], and child engagement. See Conroy et al. (2015) and Sutherland et al. (2013) for operational definitions of codes.

Observational sessions were fifteen minutes in length and occurred at three time points during each year of the study: baseline, posttest, and maintenance (approximately four weeks after posttest completion). Observation intervals were 10-s, followed by a 5-s interval for recording. Lily software (Tapp, 2010) preprogrammed with all TCIDOS codes was utilized during observational sessions. Following each session, files were analyzed using Intman software (Tapp, 1996) to reveal percentages of child behavior and interactions as well as inter-observer agreement. Prior to conducting observations, observers were trained and obtained reliable scores on each code (i.e., minimum of 80%). Following training, observers completed five master coded videos and were considered reliable when achieving 80% reliability on all codes across three videos. Additionally, observers recalibrated mid-year and prior to posttest to minimize observer drift. A total of 1202 observations were conducted over the course of the study, with 23.9% coded independently by a second observer. The mean estimates of interobserver agreement for disruptive behavior, engagement, positive interaction, and negative interaction were 96.21%, 94.10%, 94.76%, and 95.70%, respectively.

3.4.1. Student teacher relationship scale

Designed to measure teachers' perceptions of their relationships with students, the STRS (Pianta, 1993) is a teacher-report measure with subscales assessing domains of closeness (the degree of warmth, positive emotions, and open communication between teacher and child) and conflict (the degree of negative interactions and emotions involving teacher and child). Items are rated on a 5-point Likert-type scale, with responses ranging from 1 (*definitely does not apply*) to 5 (*definitely applies*) and scores summed for each domain. For the current sample, internal consistency was acceptable with Cronbach's alpha equal to 0.82 and 0.81 for Closeness at pre- and posttest, respectively and 0.86 and 0.88 for Conflict at pre- and posttest, respectively.

3.5. Treatment and comparison conditions

3.5.1. BEST in CLASS

Children in the BEST in CLASS condition were exposed to business as usual (BAU) with the addition of BEST in CLASS. BEST in CLASS is considered a Tier-2 intervention, because teachers systematically identify specific focal children in their classrooms who continue to display chronic problem behavior and are systematically identified as elevated risk for EBDs and then increase their use of targeted instructional practices with these children during instructional time. The intention of BEST in CLASS is to increase a teacher's use of effective instructional practices with focal children that promote positive teacher-child interactions and child engagement while simultaneously decreasing the occurrence of child problem behaviors (e.g., disruption, defiance). This is accomplished through the implementation of three key components delivered to teachers: (1) the BEST in CLASS Teacher Manual, (2) the BEST in CLASS Teacher Workshop (a 1-day didactic teacher training), and (3) the BEST in CLASS Practice-based Coaching (14 weeks of one-on-one practice-based coaching with performance feedback).

The BEST in CLASS Teacher Manual provides an overview of the intervention components and implementation process. The manual is comprised of seven modules; six of the modules provide instructional support on the BEST in CLASS practices (i.e., rules, precorrection, opportunities to respond, behavior specific praise, corrective feedback, and instructive feedback) and the concluding module focuses on how to link these practices together (see Table 2 for an overview of each practice). Teachers are introduced to the manual at the BEST in CLASS workshop, but continue to use the manual throughout the 14 weeks of practice-based coaching as they master each module. The BEST in CLASS workshop, led by university faculty and BEST in CLASS coaches, follows the same general course as the manual. There, teachers receive detailed information about the use of each BEST in CLASS practice.

During the study, a total of 26 coaches (24 female; 19 Caucasian, 3 Hispanic/Latino, 4 African-American) worked with teachers who participated in the BEST in CLASS condition. The BEST in CLASS coaching process begins approximately one week after the teacher workshop. The coaching process is cyclical and includes the following three steps. First, teachers and coaches meet each week to collaboratively develop an action plan for implementing BEST in CLASS practices in the classroom with the focal children. Next, the teacher implements the plan while the coach observes. In the final step, the coach and the teacher meet to reflect on the implementation plan and during that meeting, the coach provides performance-feedback. Coaches and teachers spend two weeks on the implementation of each practice. Ultimately, coaching is intended to help further the teachers' understanding of the practice and become competent in using the practice with the identified focal children in their classroom. For example, during the first several weeks, coaches collaborate with teachers to learn how to use rules effectively in their classroom with the focal children. Teachers plan with the coach when and how they will use classrooms rules more extensively with the focal child. Typically, they remind the focal child of the rules prior to an instructional activity (e.g., circle time) and then provide corrective feedback to the focal child if the rule was broken or behavior specific praise if the rule was followed. The teacher and coach collaborate

weekly to develop action plans for each of the practices. Next, the coaches' conduct an observation while the teacher implements the plan, which lasts approximately 15 min per teacher-child dyad. During this time, the coach observes and video records the teacher implementing the action plan with focal children. The coach also takes anecdotal notes and uses the TCIDOS to collect observational data on the teacher's use of the BEST in CLASS practices, teacher-child interactions, child engagement and problem behaviors using the TCIDOS. These data are used during the final step of the coaching process as a performance feedback tool to provide teacher with feedback on the implementation of the action plan. During the feedback meeting, the teacher also reflects and evaluates implementation of the plan. This three step process is repeated each week (2 weeks per practice) for 14-week of practice-based coaching. See Conroy et al. (2015) and Sutherland et al. (2015) for more detailed description of the practice-based coaching component of BEST in CLASS.

3.5.2. Business as usual

Children in the comparison group experienced a BAU condition only. The BAU condition consisted of daily instructional activities typically offered by early childhood teachers in classrooms. Typical of most federally or state-funded programs, teachers in both conditions reported using a variety of manualized early childhood curricula to guide their overall daily instruction. Teaching Strategies Gold (Heroman, Burts et al., 2010), the Creative Curriculum (Heroman, Trister Dodge et al., 2010) and High Scope (High Scope Educational Research Foundation, 2014) were the most commonly reported. Curricula and instructional strategies reported by teachers in both conditions to address social-emotional learning and manage behavior included manualized interventions (e.g., Al's Pals, (Geller, 1999); Second Step, (Committee for Children, 2011)) and specific behavioral intervention practices, including time out, token economy, response cost, and sending children to the director or the principal's office.

3.6. Treatment fidelity

Teacher implementation of the BEST in CLASS practices was assessed on two observed dimensions (adherence and competence) using the BEST in CLASS Adherence and Competence Scale (BiCACS; Sutherland, McLeod, Conroy, Abrams & Smith, 2014) in both the BEST in CLASS and BAU conditions. Adherence refers to the extent to which the program practices were delivered as intended, while competence refers to the level of skill and degree of responsiveness demonstrated by a teacher when delivering the practices (Carroll & Nuro, 2002). Raters assessed teachers' extensiveness (i.e., adherence) and quality (i.e., competence) of delivery using a 7-point Likert-type scale. Anchors on the adherence scale range from "Not at all" to "Very extensive" while anchors on the competence scale range from "Very poor" to "Excellent." Coaches conducted 389 observations in BEST in CLASS classrooms at pretest and posttest, and trained observers conducted 367 observations at pretest and posttest in BAU classrooms. A secondary observer conducted reliability checks during 204 observations (27% of the total observations). Single measure intraclass correlation coefficients, ICC (2,1), were 0.74 and 0.53 for the Adherence and Competence scales, respectively.

Mean adherence at pretest was 2.35 ($SD = 1.50$, *range* 1.41–4.77) for BAU and 2.47 ($SD = 1.42$, *range* 1.52–4.69) for BEST in CLASS with an effect size of -0.08 and at posttest was 2.43 ($SD = 1.61$, *range* 1.36–4.66) for BAU and 4.53 ($SD = 0.98$, *range* 3.72–5.99) for BEST in CLASS with an effect size of 1.30. Mean competence at pretest was 3.82 ($SD = 0.45$, *range* 3.32–4.44) for BAU and 3.78 ($SD = 0.37$, *range* 3.31–4.37) for BEST in CLASS with an effect size of 0.09 and at posttest was 4.30 ($SD = 0.42$, *range* 3.95–5.04) for BAU and 5.51 ($SD = 0.33$, *range* 5.13–5.97) for BEST in CLASS with an effect size of 2.88.

3.7. Attrition and missing data

We conducted an attrition analysis following procedures presented in WWC (U.S. Department of Education, 2016). With an overall teacher attrition rate of 4.86% and a differential attrition rate of 1.13%, teacher attrition was low according to the WWC standards presented in Table III.1 of WWC (U.S. Department of Education, 2013). We applied the WWC procedures to child outcomes variable by variable because in addition to attrition there were some cases with missing data on each posttest variable. Total missing data percentages were 10.91% for CTR-F and SSIS-S scales, 12.0% and 12.2% for STRS scales, 15.4% for T-CIDOS scales, and 16.0% to 16.2% for inCLASS scales. Differential attrition rates were 5.7% for CTR-F and SSIS-SR, 6.5% and 6.9% for STRS, 5.1% to 5.7% for inCLASS and 6.5% for TCIDOS. Missing data rates are considered low under a liberal assumption about the relationship of posttest variables and the propensity to complete the study. According to WWC (U.S. Department of Education, 2013) this may be a reasonable assumption when attrition arises, as it does in the current study, from the movement of young children in and out of school districts due to family mobility and from random absences on the days that assessments are conducted.

For each of the 15 outcome variables, pretest means were compared for groups with and without missing posttest scores. These comparisons were conducted using the Benjamini-Hochberg false discovery rate (BHFDR) procedure with a false discovery rate of 0.05. Controlling the BHFDR over all 15 variables, none of the tests were statistically significant.

3.8. Data analyses

The design of the study was a multi-site cluster randomized trial (Spybrook et al., 2011), with children nested in teachers and teachers nested in sites (i.e., schools or centers). We used a multilevel model to analyze the data on each of the scales within CTR-F, inCLASS, SSIS-RS, STRS subscales, and TCIDOS. The fixed effects included treatment, state, the Treatment by State (TxS) interaction, a coefficient for a child-level covariate, a coefficient for a teacher-level covariate, and Covariate x Treatment (CxT) interactions at the child- and the teacher- level. The child-level covariate was the class-mean centered pretest and the teacher-level-covariate was the grand mean centered teacher-level mean pretest. The treatment factor was coded -0.5 for the BAU control group and 0.5 for the BEST in CLASS group and the state factor was coded -0.5 for state A and 0.5 for state B. Random effects for child and teacher were included in the model. Estimation and hypotheses testing were conducted using M*plus* with the two-level complex procedure, which provides full information maximum likelihood estimates. The standard errors provided by two-level complex account for non-normality and for clustering at both the teacher and school level. When either interaction was non-significant, it was deleted from the model and data were reanalyzed. For each analysis the covariate was brought into the model and therefore the sample size for an analysis was all cases with pretest data, posttest data, or both. Sample sizes were N = 462 for analysis of C-TRF, SSIS-RS, and inCLASS scales, n = 461 for all TCIDOS variables and STRS Closeness, and n = 460 for STRS Conflict. Hypothesis tests were conducted using the BHFDR procedure with a false discovery rate of 0.05. The procedure was applied to the set of scales within each of C-TRF, inCLASS, SSIS-RS, and STRS, and T-CIDOS. Unadjusted probability (*p*) values are reported in the results section. Footnotes to Tables 3 and 5 indicate statistical tests that are significand by the BHFDR procedure. Wald test statistics are computed reported in M*plus*; these are reported a z in the text and in Tables 3 and 5. Estimated parameters are reported in the *b* columns and standard errors are reported in the in the SE columns in Tables 3 and 5.

Table 3
Multilevel Analysis Results: Teacher Report Measures

	b	SE	z	b	SE	z	b	SE	z
Child Teacher Report Form									
Effect	Externalizing			Internalizing			Total		
Treatment (T).	−3.82	0.85	−4.49*	−2.53	0.96	−2.64*	−3.56	0.92	−3.86*
State (S)	−0.21	0.75	−0.28	0.51	0.94	0.54	0.38	0.86	0.44
TxS	−0.36	1.76	−0.21	−1.50	1.94	−0.77	−1.44	1.87	−0.77
Child Covariate (L1)	0.61	0.07	8.14*	0.60	0.08	7.20*	0.64	0.11	5.68*
Teacher Covariate (L2)	0.68	0.07	9.38*	0.65	0.07	9.76*	0.69	0.07	9.58*
L1xT	–	–	–	–	–	–	–	–	–
L2xT	–	–	–	−0.38	0.13	−2.84†	−0.36	0.14	−2.52†
Social Skills Improvement System: RS									
	Prob. Behavior						Social Skills		
Treatment (T)	−6.68	1.40	−4.78*				5.43	1.15	4.73*
State (S)	−1.30	1.55	−0.84				−0.12	1.43	−0.08
TxS	−0.89	2.84	−0.31				−0.36	2.29	−0.16
Child Covariate (L1)	0.47	0.08	6.15*				0.50	0.05	10.62*
Teacher Covariate (L2)	0.55	0.06	9.90*				0.57	0.06	9.59*
L1xT	–	–	–				–	–	–
L2xT	−0.24	0.11	−2.23†				–	–	–
Student-Teacher Relationship Scales									
	Closeness						Conflict		
Treatment (T)	0.16	0.06	2.78*				−0.29	0.08	−3.48*
State (S)	−0.03	0.06	−0.52				0.01	0.10	0.15
TxS	0.16	0.11	1.38				−0.02	0.17	−0.14
Child Covariate (L1)	0.36	0.07	5.47*				0.50	0.07	7.55*
Teacher Covariate (L2)	0.53	0.06	9.36*				0.71	0.07	10.43*
L1xT	–	–	–				–	–	–
L2xT	–	–	–				–	–	–

Note. *Significant effect by BHFDR; †$p < 0.05$ and significant by BHFDR in model including both interactions; –Effect not included in final analysis model.

4. Results

4.1. Baseline equivalence

Comparisons of pretest means for the BEST in CLASS and BAU groups were conducted for the entire sample. None of these comparisons were statistically significant when the BHFDR was controlled over all 15 variables. Comparison of the BEST in CLASS and the BAU groups were also conducted using the BHFDR for children who did not withdraw from the study. None of the comparisons were statistically significant.

4.2. Child problem behavior

4.2.1. Caregiver-teacher report form

Results for the Externalizing scale of the C-TRF indicated that the effect of treatment was significant, $b_T = -3.82$, $z = -4.49$, $p < 0.001$ (See Table 3). Adjusted means, averaged over the states for each treatment, were 61.81 for BAU condition and 57.99 for BEST in CLASS (see Table 4 for means and standard deviations by condition). Cohen's effect size (*d*) was −0.42. Therefore, BEST in CLASS resulted in lower externalizing behaviors. The treatment effect for the Internalizing scale was also significant, $b_T = -2.53$, $z = -2.64$, $p = 0.008$ with average adjusted means equal to 54.91 for BAU and 52.38 for BEST in CLASS and $d = -0.23$. The teacher-level CxT interaction was significant for the Internalizing scale, $b_{L2xT} = -0.38$, $z = -2.84$, $p = 0.004$. Thus BEST in CLASS resulted in lower levels of internalizing than did the BAU condition and had a larger effect on internalizing problems for classes where participating focal children had higher

average levels of internalizing at pretest. For the Total Problems scale, the treatment effect was significant, $b_T = -3.56$, $z = -3.86$, $p < 0.001$ with adjusted means equal to 59.77 for BAU and 56.21 for BEST in CLASS and $d = -0.37$. The teacher-level CxT interaction was significant, $b_{L2xT} = -0.36$, $z = -2.52$, $p = 0.011$. In comparison to the BAU group, BEST in CLASS resulted in lower levels of total problems and had a larger effect on total problems for classes where participating focal children had higher average levels of total problems at pretest.

4.2.2. Social skills improvement system-rating system

For the Problem Behavior scale, the treatment effect was significant, $b_T = -6.80$, $z = -4.78$, $p < 0.001$. The adjusted means were 117.75 for BAU and 111.07 for BEST in CLASS and d was −0.42. The teacher-level CxT interaction was significant, $b_{L2xT} = -0.24$, $z = -2.23$, $p = 0.026$. Therefore, BEST in CLASS resulted in lower levels of problem behavior with a larger effect on problem behavior in classes where participating focal children had higher average levels of problem behavior at pretest. The treatment effect for the Social Skills scale was significant, $b_T = 5.43$, $z = 4.73$, $p < 0.001$, with adjusted means equal to 83.87 for BAU and 89.30 for BEST in CLASS ($d = 0.42$).

4.2.3. Teacher-child interactions direct observation system

There was a significant treatment, $b_T = 0.04$, $z = 4.37$, $p < 0.001$, $d = 0.44$, state, $b_S = -0.04$, $z = -3.57$, $p < 0.001$, $d = -0.41$, and TxS, $b_{TS} = 0.04$, $z = 2.14$, $p = 0.032$, $d = 0.45$ effect for engagement (see Table 5). Adjusted means were 0.96 and 0.90 for the BAU groups at States A and B, respectively, and 0.98 and 0.96, for the BEST in CLASS groups. Occurrence of child engagement was high in both intervention

Table 4

Means and Standard deviations for CTRF, SSIS-RS, STRS, inCLASS, and TCIDOS.

| Instrument | Scale | State | BEST in CLASS | | | | | | Comparison Group | | | | | |
| | | | Pretest | | | Posttest | | | Pretest | | | Posttest | | |
			n	M	SD	n	M	SD	n	M	SD	n	M	SD
CTRF	Externalizing	A	117	64.18	7.23	95	57.99	6.95	122	65.43	9.76	97	62.96	10.03
		B	113	64.63	8.14	110	58.09	9.48	110	63.50	8.15	98	61.06	9.67
	Internalizing	A	117	57.04	7.93	95	52.41	8.19	122	58.58	10.40	97	56.34	11.21
		B	113	57.77	9.17	110	52.52	9.80	110	55.28	9.40	98	54.08	12.74
	Total	A	117	61.97	6.70	95	56.14	7.69	122	63.75	9.42	97	61.13	10.74
		B	113	62.98	7.71	110	56.33	9.52	110	60.99	8.32	98	58.94	10.55
SSIS	Problem Behavior	A	116	122.71	16.49	95	112.54	12.46	121	122.06	17.72	97	120.92	17.15
		B	113	119.05	14.18	110	109.47	15.10	110	118.66	15.67	98	115.70	16.92
	Social Skills	A	116	74.72	11.95	95	87.27	12.63	121	78.35	12.37	97	83.34	12.78
		B	113	80.04	12.56	110	90.32	13.58	110	79.24	11.95	98	84.28	12.51
STRS	Closeness	A	113	3.96	0.75	93	4.25	0.60	119	4.02	0.64	93	4.23	0.56
		B	112	3.94	0.73	110	4.31	0.58	108	3.90	0.61	98	4.06	0.63
	Conflict	A	113	2.73	0.90	94	2.33	0.85	118	2.75	0.95	93	2.62	0.95
		B	113	2.88	1.00	110	2.45	1.06	107	2.82	1.04	98	2.67	1.05
inCLASS	Conflict	A	116	2.12	0.66	90	1.74	0.45	119	2.19	0.59	89	2.03	0.63
		B	113	1.98	0.74	103	1.65	0.53	110	1.98	0.65	94	1.83	0.60
	Peer Interaction	A	116	2.46	0.83	90	2.72	0.96	119	2.32	0.80	90	2.41	0.82
		B	113	2.21	0.83	103	2.20	0.72	110	2.31	0.84	94	2.23	0.72
	Task Orientation	A	115	3.85	0.75	90	4.37	0.86	119	3.67	0.92	90	4.02	0.81
		B	113	3.76	0.80	103	3.98	0.80	110	3.56	0.72	94	3.92	0.86
	Teacher Interaction	A	116	2.94	0.91	90	3.20	0.91	119	2.75	0.93	90	2.91	0.93
		B	113	2.67	0.87	103	2.92	0.82	110	2.72	0.80	94	2.67	0.82
TCIDOS	Positive Interaction	A	115	0.92	0.12	93	0.99	0.02	121	0.92	0.15	92	0.97	0.05
		B	112	0.88	0.18	103	0.96	0.07	111	0.83	0.19	92	0.90	0.14
	Engagement	A	115	0.92	0.09	93	0.98	0.05	121	0.92	0.14	92	0.96	0.06
		B	112	0.88	0.16	103	0.96	0.06	111	0.84	0.17	92	0.90	0.14
	Negative Interaction	A	115	0.05	0.07	93	0.01	0.02	121	0.06	0.11	92	0.02	0.05
		B	112	0.13	0.15	103	0.05	0.08	111	0.15	0.18	92	0.12	0.15
	Disruption,	A	115	0.05	0.09	93	0.00	0.01	121	0.03	0.08	92	0.02	0.04
	Aggression, and Defiance	B	112	0.11	0.13	103	0.04	0.07	111	0.13	0.17	92	0.12	0.15

Table 5

Multilevel Analysis Results: Direct Observations.

	b	SE	z	b	SE	z	b	SE	z	b	SE	z
	Individualized Classroom Assessment Scoring System											
Effect	Conflict			Peer Interaction			Task Orientation			Teacher Interaction		
Treatment (T)	−0.24	0.06	−4.24*	0.12	0.09	1.26	0.17	0.10	1.64	0.23	0.09	2.46*
State (S)	−0.13	0.06	−2.03	−0.31	0.08	−3.83*	−0.22	0.11	−1.89	−0.18	0.11	−1.68
TxS	0.11	0.12	0.98	−0.25	0.18	−1.38	−0.32	0.20	−1.61	0.04	0.19	0.20
Child Covariate (L1)	0.26	0.06	4.20*	0.17	0.07	2.30*	0.17	0.06	2.94*	0.14	0.08	1.81
Teacher Covariate (L2)	0.19	0.06	3.47*	0.26	0.08	3.51*	0.17	0.08	2.18*	0.30	0.05	5.76*
L1xT	−0.32	0.11	−2.81[†]	–	–	–	–	–	–	−0.43	0.15	−2.84[†]
L2xT	–	–	–	–	–	–	–	–	–	–	–	–

	Teacher-Child Interactions Direct Observation System											
	Positive Interaction			Engagement			Disruptive Behav.			Negative interaction		
Treatment (T)	0.04	0.01	4.72*	0.04	0.01	4.37*	−0.04	0.01	−4.65*	−0.04	0.01	−4.28*
State (S)	−0.05	0.01	−4.57*	−0.04	0.01	−3.57*	0.05	0.01	5.08*	0.05	0.01	5.55*
TxS	0.04	0.02	2.11*	0.04	0.02	2.14*	−0.05	0.02	−2.92*	−0.06	0.02	−3.12*
Child Covariate (L1)	0.02	0.05	0.51	0.04	0.06	0.64	0.07	0.05	1.27	0.09	0.06	1.43
Teacher Covariate (L2)	0.09	0.06	1.54	0.09	0.06	1.43	0.33	0.06	5.42*	0.24	0.08	3.14*
L1xT	–	–	–	–	–	–	–	–	–	–	–	–
L2xT	–	–	–	–	–	–	–	–	–	–	–	–

Note. *Significant effect by BHFDR; [†]$p < .05$ and significant by BHFDR in model including both interactions; –Effect not included in final analysis model.

groups; BEST in CLASS increased engagement, with a larger increase for State B.[1] For disruptive behavior, there were significant treatment, $b_T = -0.04$, $z = -4.66$, $p < 0.001$, $d = -0.46$, state, $b_S = 0.05$, $z = 5.08$, $p < 0.001$, $d = 0.47$, and TxS, $b_{TS} = -0.05$, $z = -2.92$, $p = 0.004$, $d = -0.57$ effects. Adjusted means were 0.03 and 0.10 for the BAU groups at States A and B, respectively, and 0.01 and 0.03,for the BEST in CLASS groups. BEST in CLASS decreased the occurrence of disruptive behavior, with a larger decrease in State B.[2]

4.3. Teacher-child interactions

4.3.1. Individualized classroom assessment scoring system

The treatment effect for the Conflict scale was significant, $b_T = -0.24$, $z = -4.24$, $p < 0.001$, with adjusted means equal to 1.94 for BAU and 1.70 for BEST in CLASS and $d = -0.43$. The child-level CxT interaction was significant, $b_{L1xT} = -0.32$, $z = -2.81$, $p = 0.005$.BEST in CLASS resulted in lower levels of conflict with larger effects for children who had higher pretest conflict scores compared to participating focal children in their class.The state effect was significant for the Peer Interaction scale, $b_S = -0.31$, $z = -3.83$, $p < 0.001$, $d = -0.38$.The treatment effect for the Teacher Interaction scale was significant $b_T = 0.23$, $z = 2.46$, $p = 0.014$ with adjusted means equal to 2.80 for BAU and 3.03 for BEST in CLASS and $d = 0.26$. The child-level CxT interaction was significant, $b_{L1xT} = -0.43$, $z = -2.84$, $p = 0.005$. BEST in CLASS resulted in higher levels of teacher interaction with larger effects for children who had lower pretest teacher interaction scores compared to participating children in their class.

4.3.2. Teacher-child interactions direct observation system

For positive interaction the following effects were significant: treatment, $b_T = 0.04$, $z = 4.72$, $p < 0.001$.$d = 0.45$, state, $b_S = -0.05$, $z = -4.57$, $p < 0.001$, $d = -0.53$, and TxS $b_{TS} = 0.04$, $z = 2.11$, $p = 0.035$, $d = 0.42$. Adjusted means were 0.97 and 0.90 for the BAU groups at States A and B, respectively, and 0.99 and 0.96, for the BEST in CLASS groups. Occurrence of positive interaction was high in both intervention groups; BEST in CLASS increased the rate of positive interaction, with a larger increase for State B.[3] For negative interaction, there were also significant treatment, $b_T = -0.04$, $z = -4.28$, $p < 0.001$, $d = -0.43$, state $b_S = 0.05$, $z = 5.55$, $p < 0.001$, $d = 0.54$, and TxS $b_{TS} = -0.06$, $z = -3.12$, $p = 0.002$, $d = -0.63$ effects. The adjusted means were 0.03 and 0.11 for the BAU groups at States A and B, respectively, and 0.02, and 0.04 for the BEST in CLASS groups. BEST in CLASS reduced the occurrence of negative interaction, with a larger effect for State B.[4]

---

[1] The distribution of positive interaction was substantially non-normal. Although the robust maximum likelihood method is designed to correct standard errors for non-normality, we also created a dichotomized version of positive interaction in which scores less than 1 were set equal to 0. This variable measures whether positive interaction was observed in fewer than all observation intervals or all observation intervals. We analyzed the variable using a two-level logistic regression model and the two-level complex procedure. The treatment effect was significant, $b_T = 1.33$, $z = 3.11$, $p = 0.001$ indicating that BEST in CLASS increased the rate of positive interaction. This result agrees with the result reported in the body of the paper.

[2] For engagement the same transformation was used and the treatment effect was significant, $b_T = 1.49$, $z = 3.09$, $p = 0.002$ indicating BEST in CLASS increased the rate of engagement. This result agrees with the result reported in the body of the paper.

[3] For the variable disruptive behavior scores greater than 0 were set equal to 1. The treatment effect was significant, $b_T = -1.52$, $z = -3.98$, $p < 0.000$ indicating BEST in CLASS decreased the rate of disruptive behavior. This result agrees with the result reported in the body of the paper.

[4] For negative interaction the transformation described in footnote 3 was used and the treatment effect was significant, $b_T = -1.33$, $z = -3.94$, $p < 0.000$ indicating BEST in CLASS decreased the rate of negative behaviors. This result agrees with the result reported in the body of the paper.

4.4. Teacher-student relationships

4.4.1. Student-teacher relationship scale

The treatment effect was significant for the Closeness Scale $b_T = 0.16$, $z = 2.78$, $p = 0.005$. The averaged adjusted means were 4.16 for BAU and 4.32 for BEST in CLASS, with $d = 0.26$. The treatment effect was also significant for the Conflict Scale $b_T = -0.29$, $z = -3.48$, $p < 0.001$. The averaged adjusted means were 2.65 for BAU and 2.36 for BEST in CLASS, with $d = -0.29$. BEST in CLASS resulted in higher closeness scores and lower conflict scores than did the BAU condition.

5. Discussion

The current study examined the efficacy of BEST in CLASS on focal children's problem behaviors and their interactions and relationships with their teachers. First, we examined the effects of BEST in CLASS on children's problem behavior as well as their social skills and engagement during early childhood classroom activities. Findings suggest that focal children in BEST in CLASS classrooms had reductions in problem behaviors, both teacher reported and observed, and improvements in teacher reported social skills and observed engagement during classroom activities in comparison to focal children in the BAU condition. While previous research investigating the outcomes of other interventions addressing behavior problems (e.g., see Domitrovich et al., 2007; Webster-Stratton et al., 2001) have found positive effects, the outcomes of the current study indicate slightly higher effects sizes than previous research. In the current investigation, we found significant effects sizes on teacher reported child problem behavior (*ES* ranged from −0.42 to −0.44) and social skills (*ES* = 0.42) within the medium range (Cohen, 1988). We also found significant increases in child engagement (*ES* = 0.44) and decreases in disruptive behavior (*ES* = −0.46) in the treatment condition when compared to BAU as measured by direct observation. In a *meta*-analysis on problem behavior targeted by intensive focused programming, Schindler et al. (2015) found mixed findings, with an average effect size of −0.26. One explanation for the higher effect sizes in the current study may be related to the level of intensity of teachers' delivery during activities throughout the day. Different than some other Tier 2 interventions (e.g., Banking Time; Driscoll & Pianta, 2010) that often use a "pull out approach" a unique feature of BEST in CLASS is the delivery of key evidence-based practices by the classroom teacher to specific focal children throughout the day during routine classroom activities, at times when interactions between teachers and children naturally occur (e.g., circle time, small group activities, transitions).

An important feature of the current study is the measurement of teacher delivery of BEST in CLASS practices via direct observation. Adherence data indicate that teachers in the BEST in CLASS condition increased their implementation of intervention practices from pretest to posttest (*ES* = 1.30), and competence data indicate that teachers in the BEST in CLASS condition increased their quality of delivery of intervention practices from pretest to posttest (*ES* = 2.88). The large effect sizes for competence of delivery may have been related to the practice-based coaching component of the BEST in CLASS model, with the feedback provided by coaches leading to greater quality of delivery of practices to focal children. While it remains unknown what level of integrity is necessary for program effects, recent work (e.g., ; Goncy, Sutherland, Farrell, Sullivan, & Doyle, 2015; Sutherland, Conroy, McLeod, Algina & Wu, 2017) suggests that competence of delivery may have a unique influence on child outcomes.

In addition to positive outcomes for child problem behavior, this study found positive effects on teacher-child interactions and relationships. Researchers have found that teachers' interactions with children in their classrooms who demonstrate problem behaviors tend to be more negative (e.g., see Barnett & Boocock, 1998); however, we found reductions in negative teacher-child interactions, increased positive

teacher-child interactions, and greater levels of closeness and reduced levels of conflict with focal children in the BEST in CLASS condition in comparison to their counterparts in the BAU classrooms. Although transactional models (Sameroff, 1983) originally focused on interactions between children and their caregivers, there is increasing emphasis on examining relationships and interactions between children and their teachers. In addition, research continues to emphasize the role positive teacher-child relationships have on short and long term educational outcomes for children, particularly when these relationships are established early (Girolametto, Weitzman, & Greenberg, 2003; NICHD, Early Child Care Research Network, 2003). As described earlier, BEST in CLASS uses high quality teacher training and practice-based coaching to increase teachers' use of evidence-based practices with children demonstrating chronic problem behavior in their classroom. Increasing teachers' use of these practices can positively influence moment-to-moment teacher-child exchanges, which result in reciprocal positive interactions between teachers and children. While not examined in the current study, over time these positive reciprocal interactions are likely to improve the teacher-child relationship and remain an important area for future research.

While there were small to moderate effect sizes for the Conflict (ES = 0.43) and Teacher Interactions (ES = 0.26) domains of the inCLASS, no effects were found for the Peer Interactions or Task Orientation domains. This is not surprising as the inCLASS measures dimensions within Teacher Interactions (e.g., positive engagement with teacher) and Conflict (e.g., teacher conflict) that BEST in CLASS targets via training and coaching in improving teacher interactions with focal children during ongoing classroom activities. While interactions with peers are clearly an important outcome for children, dimensions assessed by the Peer Interactions subscale of the inCLASS (e.g., peer sociability; peer communication) are not addressed by BEST in CLASS and therefore we would not expect a significant effect on this outcome. Non-significant findings for Task Orientation are a bit more surprising as we did find significant treatment effects for child engagement on the TCIDOS measure. In addition to engagement, the inCLASS domain of Task Orientation subscale also takes into account self-reliance. Although engagement is a target of the intervention, self-reliance is not addressed by BEST in CLASS, which might have impacted overall treatment effects on this subscale.

5.1. Limitations

Despite the promising findings, the current study has several limitations that should be considered when interpreting results. First, although blinded observations are preferred, it is important to point out that data collected on observed variables were not collected by observers blind to study condition. As described earlier in the manuscript, observers were trained and recalibrated to the gold standard for each instrument and acceptable reliability for each measure was obtained on at least 20% of the observations in both the pretest and posttest conditions. Inter-observer agreement results for observational measures were acceptable, ranging from 94.10% to 96.21% for the TCIDOS and 93.26% for the inCLASS, hopefully mitigating to some degree concern about observer bias. Second, several measures used teacher report (i.e., CTRF, SSIS-RS), which can be prone to informant bias. To address this potential concern, multiple measures, both teacher report and observational, were employed. Third, findings from the current study may not generalize to programs that serve children and families in higher resource communities. Fourth, it is not entirely clear that there was not some contamination (which would diminish the magnitude of the effect of BEST in CLASS) across BEST in CLASS and BAU classrooms that were in the same building, nor do we have specific dosage data for the universal curricula used in these classrooms. That said, the integrity data that was collected in both BEST in CLASS and BAU classrooms should mitigate this limitation to some degree, as these data show differences between the two conditions on teacher delivery of BEST in CLASS

practices. Finally, teachers who received the BEST in CLASS intervention increased their use of effective instructional practices, resulting in positive child outcomes; however, we did not examine whether similar levels of training and coaching on another intervention program would have resulted in changes in the teachers in the BAU condition (i.e., active control) or the sustainability of the outcomes of the BEST in CLASS teachers. Simply, we do not know if the results occurred because of the extensive level of training provided to the teachers in the intervention condition in general or the extensive level of training specifically in the BEST in CLASS and the long term effects of BEST in CLASS on teachers' use of the BEST in CLASS instructional practices and whether children's outcomes maintained over time.

5.2. Implications and conclusion

While the primary purpose of the current study was to examine the efficacy of BEST in CLASS on children's problem behavior and their interactions and relationships with their teachers, it is important for future research and practice to understand the mechanisms and processes contributing to children's positive outcomes. Transactional theory (see Sameroff, 1983; Sameroff & MacKenzie, 2003) provides an important framework for helping to explain our findings and provides direction for future work in this area. In particular, transactional models emphasize the reciprocal influences of interactions between adults and children and subsequent behavior. Additionally, this framework emphasizes the cumulative nature of these interactions. To examine these cumulative reciprocal interactions, future research should focus on examining how improved child outcomes may influence teachers' behavior and subsequent teacher-child interactions and relationships. For example, researchers have found that teachers are more likely to interact positively with young children who are engaged (Carr, Taylor, & Robinson, 1991); thus, future research should examine the relationship between improved child outcomes and both the proximal influence on teacher behavior as well as the cascading impact of positive reciprocal teacher-child interactions and relationships (Sutherland & Oswald, 2005). In addition, future work should investigate teacher-child interactions and teacher-child relationships as mediators of teacher-delivered interventions such as BEST in CLASS. Findings from the current study, which indicated improved child outcomes as well as increases in positive teacher-child interactions and teacher-reported closeness and reductions in teacher-reported conflict, provide a rationale for investigating these constructs as mediators. Unfortunately with the current dataset these variables were collected at the same time point as the child outcome variables, precluding mediation analyses.

As indicated by Hamre, Pianta, Mashburn, and Downer (2012) and Baker (2006), teachers play a critical role in improving the social and behavioral competence of young children in their classrooms. Research indicates a need for interventions that can be implemented in classrooms and provide additional intensive supports to positively address the social and behavioral skills of youngsters exhibiting chronic problem behavior (Qi & Kaiser, 2004; Schindler et al., 2015). Since many early childhood teachers lack the knowledge and skills to work with children who have chronic problem behavior (Hemmeter et al., 2009), programs such as BEST in CLASS using high quality teacher training and coaching to increase their knowledge, skills, and capacity of teachers to work effectively with these children in their classroom are promising. As teachers increase their capacity to deliver effective instructional practices with targeted children, their fluency and levels of delivery are likely to increase resulting in a greater impact on reduction of children's problem behavior. Future work should also examine the effect of programs such as BEST in CLASS on the broader classroom environment. While focused on children with chronic problem behavior, broad changes in teacher instructional behavior may have a positive impact on other children in the classroom and this question merits further inquiry. In addition, future research that uses an active

control group instead of BAU, including similar levels of training and coaching, may be helpful in strengthening the confidence in the efficacy of the training and coaching components of BEST in CLASS.

However, helping teachers sustain their delivery of high quality instructional practices is a challenge for the field. While there is accumulating evidence that coaching teachers in authentic settings to deliver effective instructional practices has promise for improving delivery of evidence-based programming (e.g., Joyce & Showers, 2002; Snyder et al., 2012; Sutherland et al., 2015), the feasibility of providing systematic coaching may be cost prohibitive for many programs, particularly those serving low-resource communities. Exploring ways to provide coaching and professional development in cost effective manners is an important goal for the field, as is learning more about the potential influence of teacher training and experience on implementation, and models using on-site personnel as coaches or online models of training (e.g., Miller, Fitzgerald, Koury, Mitchem, & Hollingsead, 2007; Mitchem et al., 2009) may be promising directions for future work.

In sum, the findings of the current investigation make an important contribution to the literature reinforcing the critical role of teachers' increased use of evidence-based practices with young children who demonstrate chronic problem behavior, which result in positive social/ emotional and behavioral outcomes of children as well as improved teacher-child interactions and relationships. Results from this study suggest that BEST in CLASS can be implemented by teachers with fidelity in a randomized controlled trial, and that children benefit in the short-term from their teacher's participation in the program. At the same time, targeting increased use of evidence-based practices by teachers with identified children who exhibit early patterns of problem behavior appears to have great promise at impacting both proximal (child behavior; teacher-child interactions) and distal (teacher-child relationships) outcomes for this vulnerable population.

Acknowledgements

References

Achenbach, T. M., & Rescorla, L. A. (2000). *CBCL/1, 5-5 & C-TRF/1, 5-5 profiles*. Burlington : VT: Research Center for Children, Youth, and Families.

APA (2000). *Diagnostic and statistical manual of mental disorders DSM* (4th ed.). Washington: DC: American Psychiatric Association American Psychiatric Association (2003) (text revision).

Baker, J. A. (2006). Contributions of teacher-child relationships to positive school adjustment during elementary school. *Journal of School Psychology*, 44(3), 211–229. http://dx.doi.org/10.1016/j.jsp.2006.02.002.

Barnett, W. S., & Boocock, S. S. (Eds.). (1998). Early care and education for children in poverty: promises, programs, and long-term results. (1998). *Choice Reviews Online*, 35 (11), 35–6364. doi:10.5860/choice.35-6364.

Benson, B. A., & Aman, M. G. (1999). Disruptive behavior disorders in children with mental retardation. In H. C. Quay, & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 559–578). New York: Springer. http://dx.doi.org/10.1007/978-1-4615-4881-2_26.

Beyer, T., Postert, C., Muller, J. M., & Furniss, T. (2012). Prognosis and continuity of children mental health problems from preschool to primary school: Results of a four-year longitudinal study. *Child Psychiatry and Human Development*, 43, 533–543.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Making human beings human: Bioecological perspectives on human development. In U. Bronfenbrenner (Ed.), Thousand Oaks, CA: Sage Publications, Inc.

Burchinal, M., Vandergrift, N., Pianta, P., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, 25(2), 166–176.

Buyse, E., Verschueren, K., Verachtert, P., Van Damme, J., & Leuven, K. U. (2009). Predicting school adjustment in early elementary school: Impact of teacher-child relationship quality and relational classroom climate. *The Elementary School Journal*, 110(2), 119–141.

Campbell, S. B., Spieker, S., Burchinal, M., & Poe, M. D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry*, 47(8), 791–800. http://dx.doi.org/10.1111/j.1469-7610.2006.01636.x.

Carr, E. G., Taylor, J. C., & Robinson, S. (1991). The effects of severe behavior problems in children on the teaching behavior of adults. *Journal of Applied Behavior Analysis*, 24(3), 523. http://dx.doi.org/10.1901/jaba.1991.24-523.

Carroll, K. M., & Nuro, K. F. (2002). One size cannot fit all: A stage model for psychotherapy manual development. *Clinical Psychology: Science and Practice*, 9, 396–406. http://dx.doi.org/10.1093/clipsy/9.4.396.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates, 20–26.

Committee for Children (2011). *Second step: Social-emotional skills for early learning*. Seattle: Author.

Conroy, M. A., Sutherland, K. S., Vo, A. K., Carr, S. E., & Ogston, P. (2014). Early childhood teachers' use of effective instructional practices and the collateral effects on young children's behavior. *Journal of Positive Behavior Interventions*, 16, 81–92. http://dx.doi.org/10.1177/1098300713478666.

Conroy, M. A., Sutherland, K. S., Wilson, R. E., Martinez, J., Whalon, K. J., & Algina, J. (2015). Measuring teacher implementation of the BEST in CLASS intervention program and corollary child outcomes. *Journal of Emotional and Behavioral Disorders*, 23, 144–155. http://dx.doi.org/10.1177/1063426614532949.

Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher-child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, 101(4), 912. http://dx.doi.org/10.1037/a0016647.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). THE EARLY YEARS: Preschool program improves cognitive control. *Science*, 318(5855), 1387–1388. http://dx.doi.org/10.1126/science.1151148.

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

Doumen, S., Verschueren, K., Buyse, E., Germeijs, V., Luyckx, K., & Soenens, B. (2008). Reciprocal relations between teacher-child conflict and aggressive behavior in Kindergarten: A three-wave longitudinal study. *Journal of Clinical Child & Adolescent Psychology*, 37(3), 588–599. http://dx.doi.org/10.1080/15374410802148079.

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly*, 25(1), 1–16. http://dx.doi.org/10.1016/j.ecresq.2009.08.004.

Downer, J., Sabol, T. J., & Hamre, B. (2010). Teacher–child interactions in the classroom: Toward a theory of within-and cross-domain links to children's developmental outcomes. *Early Education and Development*, 21(5), 699–723. http://dx.doi.org/10.1080/10409289.2010.497453.

Driscoll, K. C., & Pianta, R. C. (2010). Banking time in Head Start: Early efficacy of an intervention designed to promote supportive teacher-child relationships. *Early Education and Development*, 21, 38–64. http://dx.doi.org/10.1080/10409280802657449.

Dunlap, G., Strain, P. S., Fox, L., Carta, J. J., Conroy, M., Smith, B. J., ... Sowell, C. (2006). Prevention and intervention with young children's challenging behavior: Perspectives regarding current knowledge. *Behavioral Disorders*, 32(1), 29–45. Retrieved from http://www.jstor.org/stable/23890617.

Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., ... Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, 21(2), 174–195. http://dx.doi.org/10.1016/j.ecresq.2006.04.004.

Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: Findings from the national institute of child health and human development study of early child care. *Developmental Psychology*, 46, 1159–1175. http://dx.doi.org/10.1037/a0020659.

Fox, L., Carta, J., Strain, P. S., Dunlap, G., & Hemmeter, M. L. (2010). Response to intervention and the Pyramid Model. *Infants and Young Children*, 23(1), 3–13.

Geller, S. (1999). *Al's pals: Kids making healthy choices*. Richmond,VA: Wingspan LLC.

Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. New York, NY: Foundation for Child Development.

Girolametto, L., Weitzman, E., & Greenberg, J. (2003). Training day care staff to facilitate children's language. *American Journal of Speech-Language Pathology*, 12(3), 299–311. http://dx.doi.org/10.1044/1058-0360(2003/076).

Goncy, L. A., Sutherland, K. S., Farrell, A. D., Sullivan, T. N., & Doyle, S. T. (2015). Measuring teacher implementation in delivery of a bullying prevention program: The impact of instructional and protocol adherence and competence on student responsiveness. *Prevention Science*, 16, 440–450. http://dx.doi.org/10.1007/s11121-014-0508-9.

Graves, S. L., & Howes, C. (2011). Ethnic differences in social-emotional development in preschool: The impact of teacher child relationships and classroom quality. *School Psychology Quarterly*, 26, 202–214.

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

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72(2), 625–638. http://dx.doi.org/10.1111/1467-8624.00301.

Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2012). Promoting young children's social competence through the preschool PATHS curriculum and MyTeachingPartner professional development resources. *Early Education & Development*, 23(6), 809–832. http://dx.doi.org/10.1080/10409289.

2011.607360.

Hamre, B. K. (2014). Teachers' daily interactions with children: An essential ingredient in effective early childhood programs. *Child Development Perspectives*, *8*(4), 223–230. http://dx.doi.org/10.1111/cdep.12090.

Hemmeter, M. L., Fox, L., & Snyder, P. (2009). Evaluating the potential efficacy of the Teaching Pyramid Intervention. *Poster presented at the meeting of the Institute of Education Sciences.*

Henricsson, L., & Rydell, A. M. (2004). Elementary school children with behavior problems: Teacher-child relations and self-perception. A prospective study. *Merrill-Palmer Quarterly*, *50*(2), 111–138. http://dx.doi.org/10.1353/mpq.2004.0012.

Heroman, C., Burts, D. C., Berke, K., & Bickart, T. S. (2010). *Teaching Strategies GOLD objectives for development & learning: Birth through kindergarten.* Bethesda, MD: Teaching Strategies, LLC.

Heroman, C., Trister Dodge, D., Berke, K., Bickart, T. S., Colker, L., Jones, C., ... Dighe, J. (2010). *The creative curriculum for preschool* (5th ed.). Bethesda, MD: Teaching Strategies, LLC.

Hetzner, N., Johnson, A. D., & Brooks-Gunn, J. (2010). Poverty, effects of on social and emotional development. *International Encyclopedia of Education*, 643–652. http://dx.doi.org/10.1016/b978-0-08-044894-7.00617-5.

High Scope Educational Research Foundation (2014). *High scope curriculum.* Ypsilanti, MI: High Scope Educational Research Foundation.

Howes, C., & Smith, E. W. (1995). Relations among child care quality, teacher behavior, children's play activities, emotional security, and cognitive activity in child care. *Early Childhood Research Quarterly*, *10*(4), 381–404. http://dx.doi.org/10.1016/0885-2006(95)90013-6.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Kaiser, A. P., Cai, X., Hancock, T. B., & Foster, E. M. (2002). Teacher-reported behavior problems and language delays in boys and girls enrolled in Head Start. *Behavioral Disorders*, *28*, 23–39.

Ladd, G. W., & Burgess, K. B. (1999). Charting the relationship trajectories of aggressive, withdrawn, and aggressive/withdrawn children during early grade school. *Child Development*, *70*, 910–929. http://dx.doi.org/10.1111/1467-8624.00066.

Marchant, M., Young, K. R., & West, R. P. (2004). The effects of parental teaching on compliance behavior of children. *Psychology in the Schools*, *41*(3), 337–350. http://dx.doi.org/10.1002/pits.10165.

Markowitz, J., Carlson, E., Frey, W., Riley, J., Shimshak, A., Heinzen, H., ... Hyunshik, L. (2006). *Preschoolers with disabilities: Characteristics, services, and results. wave 1 overview report from the pre-elementary education longitudinal study (PEELS). NCSER 2006–3003.* National Center for Special Education Research.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947. http://dx.doi.org/10.1037/0012-1649.43.4.947.

Miller, K., Fitzgerald, G., Koury, K., Mitchem, K., & Hollingsead, C. (2007). KidTools: Self-management, problem solving, organizational and planning tools for children and teachers. *Intervention in School and Clinic*, *43*(1), 12–19.

Mitchem, K., Koury, K., Fitzgerald, G., Hollingsead, C., Miller, K., Tsai, H. H., & Zha, S. (2009). The effects of instructional implementation on learning with interactive multimedia case-based instruction. *Teacher Education and Special Education*, *32*, 297–318.

Morris, P., Mattera, S. K., Castells, N., Bangser, M., Bierman, K., & Raver, C. (2014). *Impact findings from the head start CARES demonstration: national evaluation of three approaches to improving preschoolers' social/emotional competence. executive summary. OPRE report 2014-44.* MDRC.

National Institute of Child Health and Human Development (NICHD) (2003). Does quality of child care affect child outcomes at age 4 ½? *Developmental Psychology*, *39*, 451–469.

Newborg, J. (2005). *Battelle developmental inventory* (2nd ed.). Itasca, IL: Riverside.

O'Connor, E., & McCartney, K. (2007). Examining teacher-child relationships and achievement as part of an ecological model of development. *American Educational Research Journal*, *44*(2), 340–369. http://dx.doi.org/10.3102/0002831207302172.

Patterson, G. R., Reid, J. B., & Eddy, J. M. (2002). A brief history of the Oregon model. In J. B. Reid, G. R. Patterson, & J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 3–21). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10468-001 pp. 3–20.

Phillips, D. A., & Shonkoff, J. P. (Eds.), (2000). *From neurons to neighborhoods: The science of early childhood development.* National Academies Press.

Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. http://dx.doi.org/10.1207/s1532480xads0903_2.

Pianta, R. C. (1993). *The student-teacher relationship scale.* Charlottesville, VA: University of Virginia.

Qi, C. H., & Kaiser, A. P. (2003). Behavior problems of preschool children from low-income families review of the literature. *Topics in Early Childhood Special Education*, *23*(4), 188–216. http://dx.doi.org/10.1177/02711214030230040201.

Qi, C. H., & Kaiser, A. P. (2004). Problem behaviors of low-income children with language delays. *Journal of Speech Language and Hearing Research*, *47*(3), 595. http://dx.doi.org/10.1044/1092-4388(2004/046).

Raver, C. C., & Knitzer, J. (2002). *Ready to enter: What research tells policymakers about*

strategies to promote social and emotional school readiness among three- and four-year-olds. St. Louis: Federal Reserve Bank of St. Louis. Retrieved from http://proxy.library.vcu.edu/login?url=http://search.proquest.com.proxy.library.vcu.edu/docview/1698411792?accountid=14780.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Smallwood, K., & Sardine, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly*, *23*(1), 10–26.

Sameroff, A. J., & MacKenzie, M. J. (2003). Research strategies for capturing transactional models of development: The limits of the possible. *Development and Psychopathology*, *15*(03), 613–640. http://dx.doi.org/10.1017/s0954579403000312.

Sameroff, A. J. (1983). Developmental systems: Contexts and evolution. In P. H. Mussen (Gen. Ed.), & W. Kessen (Vol. Ed.) (Eds.), *Handbook of child psychology: Vol 1. History, theory, and methods* (pp. 237–294). (4th ed.). New York: Wiley.

Schindler, H. S., Kholoptseva, J., Oh, S. S., Yoshikawa, H., Duncan, G. J., Magnuson, K. A., & Shonkoff, J. P. (2015). Maximizing the potential of early childhood education to prevent externalizing behavior problems: a meta-analysis. *Journal of School Psychology*, *53*(3), 243–263. http://dx.doi.org/10.1016/j.jsp.2015.04.001.

Scott-Little, M. C., & Holloway, S. D. (1992). Child care providers' reasoning about misbehaviors: Relation to classroom control strategies and professional training. *Early Childhood Research Quarterly*, *7*(4), 595–606. http://dx.doi.org/10.1016/0885-2006(92)90125-i.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Snyder, P., Hemmeter, M. L., Meeker, K. A., Kinder, K., Pasia, C., & McLaughlin, T. (2012). Characterizing key features of the early childhood professional development literature. *Infants & Young Children*, *25*(3), 188–212.

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., Raudenbush, S., & TO, A. (2011). *Optimal design plus empirical evidence: Documentation for the Optimal Design software.* William T. Grant Foundation Retrieved on November, 5, 2012.

Stormont, M., Beckner, R., Mitchell, B., & Richter, M. (2005). Supporting successful transition to kindergarten: General challenges and specific implications for students with problem behavior. *Psychology in the Schools*, *42*(8), 765–778. http://dx.doi.org/10.1002/pits.20111.

Sutherland, K. S., & Oswald, D. P. (2005). The relationship between teacher and student behavior in classrooms for students with emotional and behavioral disorders: Transactional processes. *Journal of Child and Family Studies*, *14*(1), 1–14. http://dx.doi.org/10.1007/s10826-005-1106-z.

Sutherland, K. S., Conroy, M. A., Abrams, L., Vo, A., & Ogston, P. (2013). An initial evaluation of the Teacher-Child Interaction Direct Observation System: Measuring teacher-child interaction behaviors in classroom settings. *Assessment for Effective Intervention*, *39*, 12–23. http://dx.doi.org/10.1177/1534508412463814.

Sutherland, K. S., Conroy, M. A., Vo, A., & Ladwig, C. (2015). Implementation integrity of practice-based coaching: Preliminary results from the BEST in CLASS Efficacy Trial. *School Mental Health*, *7*, 21–33. http://dx.doi.org/10.1007/s12310-014-9134-8.

Sutherland, K. S., Conroy, M. A., McLeod, B. D., Algina, J., & Wu, E. (2017). Teacher competence of delivery of BEST in CLASS as a mediator of treatment effects. *School Mental Health.*. http://dx.doi.org/10.1007/s12310-017-9224-5.

Tapp, J. T. (1996). *Intman: Interval manager [Computer software].* Nashville, TN: Vanderbilt Kennedy Center.

Tapp, J. T. (2010). *Lily data collector [Computer software].* Nashville, TN: Vanderbilt Kennedy Center.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse (2013). *What works clearinghouse: Procedures and standards handbook (version 3.0).* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf. (2013, March).

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse (2016). *What works clearinghouse: Cluster standards.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_cluster_standards_030416.pdf. (2016, March).

Van Acker, R., Grant, S. H., & Henry, D. (1996). Teacher and student behavior as a function of risk for aggression. *Education and Treatment of Children*, *19*(3), 316–334.

Vancraeyveldt, C., Verschueren, K., Wouters, S., Van Craeyevelt, S., Van den Noortgate, W., & Colpin, H. (2015). Improving teacher-child relationship quality and teacher-rated behavioral adjustment amongst externalizing preschoolers: Effects of a two-component intervention. *Journal of Abnormal Child Psychology*, *43*, 243–257. http://dx.doi.org/10.1007/s10802-014-9892-7.

Wahler, R. G., & Dumas, J. E. (1986). Maintenance factors in coercive mother-child interactions: The compliance and predictability hypotheses. *Journal of Applied Behavior Analysis*, *19*(1), 13–22. http://dx.doi.org/10.1901/jaba.1986.19-13.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in Head Start. *Journal of Clinical Child Psychology*, *30*(3), 283–302. http://dx.doi.org/10.1207/s15374424jccp3003_2.

Webster-Stratton, C. (1997). Early intervention for families of preschool children with conduct problems. In M. J. Guralnick (Ed.), *The effectiveness of early intervention* (pp. 429–453). Baltimore: Brookes.

Wehby, J. H., Symons, F. J., & Shores, R. E. (1995). A descriptive analysis of aggressive behavior in classrooms for children with emotional and behavioral disorders. *Behavioral Disorders*, *20*(2), 87–105.

Williford, A. P., Wolcott, C. S., Whittaker, J. V., & LoCasale-Crouch, J. (2015). Program and teacher characteristics predicting the implementation of Banking Time with preschoolers who display disruptive behaviors. *Prevention Science*, *16*(8), 1054–1063. http://dx.doi.org/10.1007/s11121-015-0544-0.