AERA Open
January-March 2018, Vol. 4, No. 1, pp. 1–16
DOI: 10.1177/2332858417750376

Prevention and Treatment of Problem Behaviors in Young Children: Clinical Implications From a Randomized Controlled Trial of BEST in CLASS

Maureen A. Conroy
University of Florida
Kevin S. Sutherland
Virginia Commonwealth University
James Algina
Brittany Werch
Crystal Ladwig
University of Florida

The prevention science approach to emotional/behavioral disorders (EBD) focuses on early intervention targeting risk and resilience factors (e.g., early problem behaviors, teacher-child interactions, classroom climate). The current study investigates the effectiveness of BEST in CLASS, a classroom-based indicated preventive intervention targeting young children at risk for EBD, in terms of producing clinically meaningful reductions of problem behaviors and improvements in social skills. A total of 185 early childhood teachers (92 in the treatment condition) and 462 young children (230 in treatment) participated in the study. Teachers completed rating scale measures of problem behaviors and social skills. The clinical significance of the findings indicated that the BEST in CLASS condition had a lower percentage of children with clinical or borderline range scores at posttest when compared with children in the control classrooms. The BEST in CLASS prevention model, clinically significant intervention outcomes, and prevention of EBD are discussed.

Keywords: *early childhood, at risk students, instructional practices*

Given the national focus on the importance of the early years of a child's development as they relate to school readiness, it is increasingly critical that children be given high-quality early learning experiences to ensure that they can be successful when entering school. *School readiness* refers to children's level of preparation for learning in the school setting, as determined by their physical, cognitive, academic, social, and emotional development (U.S. Department of Health and Human Services, 2015). With the expansion of federally and state-funded early childhood programs, the promotion of young children's learning and development in early childhood school settings has received increased attention. Although high-quality learning experiences related to academic development, such as literacy and mathematics, are typically emphasized, social-emotional learning (SEL) is also critical to school readiness (Zins, Bloodworth, Weissberg, & Walberg, 2007). SEL includes skills that contribute to emotion regulation, positive interactions with teachers and peers, and appropriate behavior, which are important for school success (Collaborative for Academic, Social, and Emotional Learning, 2013).

Prevalence and Associated Factors

Unfortunately, the prevalence of young children who enter school with SEL difficulties and significant problem behavior seems to be increasing. An estimated 5% to 26% of children demonstrate serious social, emotional, and behavioral learning challenges (Brauner & Stephens, 2006). In addition, preschool children are suspended and expelled for demonstrating problem behaviors at three times the rate of K–12 students (Gilliam, 2005). The short- and long-term outcomes associated with problem behavior are poor. Early childhood problem behavior contributes to negative relationships with teachers and peers (Bulotsky-Shearer, Bell, & Dominguez, 2012; Bulotsky-Shearer, Dominguez, & Bell, 2012; Carter et al., 2010), learning and academic achievement difficulties (Hamre & Pianta, 2001), and later identification of emotional/behavioral disorders (EBD; Fanti & Henrich, 2010). According to research, problem behaviors develop early, and if they are not addressed when they are initially displayed, they worsen over time, require more

services and resources, and increase the likelihood for long-term negative outcomes (Dunlap et al., 2006).

The prevalence and outcomes of problem behaviors are even more concerning among young children who experience a range of child, family, and environmental risk factors. For example, rates of externalizing problem behavior are especially elevated for children living in poverty (Barbarin, 2007; Feil et al., 2005; Qi & Kaiser, 2003). In addition, negative child temperament, child adjustment problems, poor family functioning, and maternal depression are predictive of problem behaviors (Nelson, Stage, Duppong-Hurley, Synhorst, & Epstein, 2007). Young children who experience adverse relationships with their teachers and a negative classroom climate are also at a higher risk for problem behavior (Curby, Rimm-Kaufman, & Ponitz, 2009; Myers & Pianta, 2008; Pianta & Stuhlman, 2004; Pianta et al., 2005). When young children are exposed to a greater number and combination of these risk factors, their likelihood for future development of EBD, such as conduct disorder and oppositional defiant disorder, increases (Qi & Kaiser, 2003; Webster-Stratton, 1997). Research indicating the origins and risk factors associated with problem behavior highlights the need to examine early intervention, aimed at preventing these deleterious behaviors and lessening the impact of risk factors on children's school success.

While noted risk factors raise the likelihood of problem behavior, factors leading to resilience among young children play an important part in the prevention of problem behavior and the promotion of school readiness. Resilience (or protective) factors include child, family, and environmental characteristics that have a positive impact on a child's life, despite level of risk (National Research Council and Institute of Medicine, 2009a). For example, positive teacher-child interactions serve as a resilience factor influencing future social-emotional, behavioral, and academic success in school (Baker, 2006; Bulotsky-Shearer, Manz, et al., 2012; McCabe & Altamura, 2011; O'Connor & McCartney, 2007; Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010). Even when controlling for child and family factors, the quality of teacher-student relationships has demonstrated an effect on students' academic achievement (O'Connor & McCartney, 2007). In addition, the early development of social, emotional, and behavioral competence may prevent the onset of persistent problem behaviors (Bornstein, Chun-Shin, & Haynes, 2010). Therefore, addressing resilience factors, such as children's SEL and behavioral needs, within the context of the early childhood classroom is essential for the prevention of problem behaviors and the promotion of school readiness.

Prevention Science

Despite research on the risk factors associated with problem behaviors and the interventions based on tiered systems of support, effective interventions are not typically well integrated into early childhood programs to prevent the development of risk among young children and to promote their resilience. Traditional approaches to social-emotional and behavioral interventions for children who need additional support are also not often preventively employed in practice; rather, they focus on the treatment of individual children once identified and diagnosed—specifically, to reduce indicators (National Research Council and Institute of Medicine, 2009b). In contrast, prevention science aims to prevent problem behavior in the early stages of development as services are offered to children who are at risk for EBD, to reduce their likelihood of developing the disorder and needing treatment (National Research Council and Institute of Medicine, 2009a). The prevention science approach to problem behavior connects research on risk and resilience factors.

To illustrate, risk and resilience factors associated with problem behavior have been targeted with varying levels of intervention practices and programs, including *universal* (Tier 1), *selective* (Tier 2), and *indicated* (Tier 3) preventive positive behavioral support interventions (Sugai et al., 2000). Given that the goal of prevention science is to address problem behavior prior to or in the early stages of development, early childhood programs represent an important context for implementing these preventive interventions and addressing risk and resilience factors. The Teaching Pyramid (Fox, Carta, Strain, Dunlap, & Hemmeter, 2010; Hemmeter, Fox, & Snyder, 2013; Hemmeter, Ostrosky, & Fox, 2006) is a comprehensive multitiered model of support aimed at the social and behavioral competence needs of all children in early childhood classrooms. Through professional development activities (i.e., training and practice-based coaching), the Teaching Pyramid supports early childhood teachers' use of effective strategies to support children's SEL and development (see Hemmeter, Snyder, Kinder, & Artman, 2011; Snyder, Hemmeter, & Fox, 2015). On a Tier 2 level, teachers provide focused social-emotional supports for children who are at risk for EBD and have been unresponsive to universal supports at the Tier 1 level. Finally, Tier 3 levels of support are provided to children who continue to engage in persistent and ongoing problem behaviors and who have not responded to a Tier 1 or 2 level of support. Research has documented the efficacy of the Teaching Pyramid in ameliorating the needs of young children who are at varying levels of risk (see Hemmeter, Snyder, Fox, & Algina, 2016).

In addition to a comprehensive tiered approach—which includes an integrated system of levels of support based on need—some stand-alone, classroom-based early childhood universal (Tier 1) preventive interventions targeting SEL outcomes have been examined. These universal preventive interventions are typically implemented with all children from a selected population, regardless of individual risk factors. For example, Incredible Years (Webster-Stratton, Reid, & Hammond, 2004) is a universal curriculum for preventing

emotional and behavioral problems by enhancing social-emotional competence and reducing classroom risk factors among young children who come from socioeconomically disadvantaged backgrounds (Webster-Stratton, Reid, & Stoolmiller, 2008). Another example is Second Step (Committee for Children, 1991), a universal program aimed at preventing aggressive behavior among children and adolescents by increasing their social skills (Frey, Hirschstein, & Guzzo, 2000). Although universal preventive interventions aim to generally promote resilience and reduce risk associated with problem behavior and SEL difficulties, approximately 20% of children need additional supports at the Tier 2 or 3 level.

Unlike universal (Tier 1) preventive interventions, indicated (Tier 2) preventive interventions are aimed at high-risk children who have some signs or symptoms of a social, emotional, or behavioral disorder but do not currently meet diagnostic criteria (National Research Council and Institute of Medicine, 2009a). Indicated prevention of EBD commonly includes risk- and resilience-focused interventions for young children who have elevated aggression and behavioral difficulties. One such example is the FAST Track program (Conduct Problems Prevention Research Group, 1992), a prevention science approach targeting children at risk for conduct disorder and the developmental influences contributing to this elevated risk (Conduct Problems Research Prevention Group, 2002). The results of the efficacy trial showed that the FAST Track intervention reduced the lifetime prevalence of conduct disorder, attention-deficit/hyperactivity disorder, and oppositional defiant disorder among those at the highest initial risk (Conduct Problems Prevention Research Group, 2002; Dodge & Conduct Problems Prevention Research Group, 2007). In addition, First Steps to Success (Feil et al., 2014) is an intervention designed to reduce young children's problem behaviors in early childhood settings. In a recent efficacy trial, children who received the Preschool First Step to Success intervention demonstrated higher social skills and significantly fewer problem behaviors. These results provide evidence for the success of indicated preventive interventions implemented in early childhood school settings to reduce problem behaviors and the prevalence of EBD.

Finally, indicated (Tier 3) interventions are needed for those children who continue to engage in persistent and significant problem behaviors that further place them at elevated risk for EBD. Researchers found that a systematic process of providing individualized positive behavior supports is often required for children who need Tier 3 interventions (Dunlap & Fox, 2009; Dunlap et al., 2006). In a recent article, Dunlap and colleagues (Dunlap, Lee, Joseph, & Strain, 2015) provided a description of the "prevent-teach-reinforce for young children" model, designed to employ functional assessment and individualized positive supports for those needing Tier 3 interventions.

BEST in CLASS: An Indicated Preventive Intervention

Similar to other indicated interventions, BEST in CLASS (i.e., Behavioral, Emotional, Social Training: Competent Learners Achieving School Success; Conroy & Sutherland, 2008) represents a prevention science model at the Tier 2 level in that it systematically identifies and targets young children who engage in elevated chronic problem behaviors that place them at risk for the future identification of EBD. Unlike other Tier 2 interventions (e.g., FAST Track), BEST in CLASS is implemented by teachers during ongoing classroom instruction, which provides them with the flexibility to use the practices when identified children most need them. Theoretically, BEST in CLASS leverages the impact of the social transactions between teachers and children (Sameroff, 2009) within the early childhood classroom ecology (Bronfenbrenner, 1979, 2005) in terms of their social, emotional, and behavioral development. In practice, BEST in CLASS is based on a combination of high-quality professional development training and practice-based coaching to support the use of effective teacher instructional practices in early childhood classrooms that promote positive teacher-child interactions, enhance child engagement, increase learning opportunities, and decrease the occurrence of problem behaviors among young children (Conroy & Sutherland, 2008). Given that a negative classroom climate raises the risk for problem behaviors (Curby et al., 2009; Myers & Pianta, 2008; Pianta & Stuhlman, 2004; Pianta et al., 2005), BEST in CLASS promotes a positive classroom climate by focusing on enhanced child academic and social engagement and high-quality learning opportunities. In addition, BEST in CLASS promotes a higher level of positive teacher-child interactions and relationships, which enhance SEL and school readiness (Baker, 2006; Bulotsky-Shearer, Manz, et al., 2012; McCabe & Altamura, 2011; O'Conner & McCartney, 2007; Vandell et al., 2010).

BEST in CLASS is designed to equip early childhood teachers with the skills necessary to address the needs of children who are considered at risk for the development of EBD. The goal of BEST in CLASS is to facilitate teachers' fidelity of key instructional practices targeting specific children who are demonstrating problem behaviors across various contexts in the classroom. This is accomplished through the delivery of the following three core professional development components: a 1-day teacher training workshop, a teacher manual, and weekly practice-based coaching occurring over 14 weeks (see the appendix for a description of the BEST in CLASS professional development components).

Studies have been conducted on the BEST in CLASS intervention examining teacher outcomes (Conroy et al., 2017) and child outcomes (Conroy et al., 2015; Sutherland et al., 2018), as well as teacher implementation fidelity (Conroy et al., 2015) and coach implementation fidelity (Sutherland, Conroy, Vo, & Ladwig, 2015). Prior research

studies conducted on the BEST in CLASS intervention showed that teachers are able to implement the practices with high levels of adherence and competence (Conroy et al., 2017; Sutherland et al., 2018). Furthermore, BEST in CLASS decreases child problem behaviors, increases child engagement and social skills, and improves teacher-child interactions and relationships, including reductions in negative teacher-child interactions and raise the level of positive teacher-child interactions (see Sutherland et al., 2018). These results suggest that BEST in CLASS is effective in decreasing problem behaviors. However, no research has been conducted that examines the clinical significance of BEST in CLASS and the potential impact that it has on the reduction of risk for EBD.

Clinical Efficacy of Preventive Interventions

As discussed, the efficacy of various social-emotional and behavioral preventive interventions, including BEST in CLASS, is often substantiated by an overall increase in social skills and a decrease in problem behavior from pre- to post-intervention. However, few efficacy studies move beyond the *statistical* significance of findings and examine the *clinical* significance of findings, or the practical benefit of the intervention to the life of the child (Kazdin, 2003). That is, few studies determine whether, from pre- to post-intervention, children in the intervention group moved toward a healthier range of functioning in social, emotional, and behavioral competence and well-being as compared with children in the comparison group.

Examining efficacy in terms of clinically meaningful reductions in externalizing problem behavior and improvements in social skills is especially important for indicated preventive interventions that target children who demonstrate such levels outside the typical range of functioning, placing them at risk for EBD. Additionally, examining clinically meaningful changes can help researchers interpret findings from efficacy studies to highlight what interventions might be more or less effective for individual children (i.e., what works for whom).

Several studies of early childhood preventive interventions for problem behavior examined the clinical significance of findings. One such study investigated the clinical significance of the Early Risers Program, implemented by neighborhood family resource centers with urban kindergarten and first-grade children screened for aggressive behavior (August, Lee, Bloomquist, Realmuto, & Hektner, 2003). Clinical significance was determined through examination of pre- and post-intervention mean scores within the normative and nonnormative ranges on the Behavior Assessment System for Children (i.e., a teacher report form; Reynolds & Kamphaus, 2004) between intervention and control groups. Although differences between the experimental and control groups were not found for children who were mildly aggressive, clinically important differences were found for children who demonstrated severe aggression. Children with severe aggression in the intervention and control groups improved during the first year of the program. However, during the second year, children who demonstrated severe aggression in the intervention group maintained their gains, and the children in the control group worsened. In another study, Morrison and Bratton (2010) investigated the clinical significance of findings following Child-Teacher Relationship Training, which focused on preschool children in Head Start with identified behavior problems. According to teacher report data from the Caregiver-Teacher Report Form (C-TRF; Achenbach & Rescorla, 2000), 9 of the 15 children in the experimental group who scored in the borderline or clinical range prior to the intervention had improved scores in the normal range following the intervention. Last, Walker et al. (1998) investigated the clinical significance of the First Step to Success intervention, aimed at kindergarteners displaying early signs of antisocial behavior patterns (e.g., aggression, defiance). Although a number of efficacy studies conducted on the First Step to Success intervention yielded statistically significant child outcomes (e.g., see Feil et al., 2014; Sumi et al., 2012; Walker et al., 2009), Walker et al. found that the average scores for both cohorts on the Child Behavior Checklist Aggression subscale (Achenbach & Rescorla, 2000) went from marginally at risk (pre-intervention) to average (post-intervention). In addition, children scoring in the clinical range before the intervention had less favorable profiles across the dependent measures following the intervention than those children scoring in the nonclinical range before the intervention. With a limited number of studies, there is a clear need for prevention science research to examine the clinical significance of findings to determine the practical impact of preventive interventions on the development of problem behavior and EBD.

Study Purpose

Despite the importance of earlier prevention efforts, the majority of school-based prevention programs have focused on children in first through 12th grade (Humphries & Keenan, 2006). The current study contributes to this prevention science research gap by investigating a social, emotional, and behavioral preventive intervention that teachers can deliver with fidelity in authentic early childhood settings. In addition, this study adds to the limited research investigating the clinical significance of early childhood preventive interventions for problem behaviors. BEST in CLASS is an indicated preventive intervention that targets risk and resilience factors for the problem behaviors of young children identified as being at risk for EBD. The goal of this study was to determine the efficacy of BEST in CLASS in producing clinically meaningful reductions in problem behaviors and improvements in social skills. The

primary research question centered on determining how many children receiving BEST in CLASS improved from pretest clinical or borderline ranges to posttest normal ranges for social skills and problem behavior as compared with the control classrooms.

Method

The current work was conducted as part of a 4-year efficacy study of the BEST in CLASS intervention (see Sutherland et al., 2018). Randomization into treatment and comparison groups occurred at the teacher level via a multisite cluster randomized design (Spybrook et al., 2011).[1] Teachers were blocked at the school level and randomly assigned to a BEST in CLASS intervention group or a business-as-usual comparison group. Teachers in both groups continued to implement their program-approved early childhood curricula throughout the study. Institutional review boards at two universities independently approved all research activities.

Setting and Participants

Research was conducted in 78 early childhood program settings across five school districts and three Head Start programs in two southeastern states. All programs, regardless of location, were federally and/or state-funded early childhood programs serving children 3 to 5 years of age (e.g., Head Start, state-funded prekindergarten).

Teachers. Teachers were deemed eligible for the study if they were currently employed as a preschool teacher in an early childhood program, had at least one child who met criteria for inclusion in the study, and consented to participate. A total of 185 teachers participated in the study, with 92 teachers in the BEST in CLASS condition and 93 teachers in the comparison condition (see Table 1 for teacher demographics). The majority of teachers held an associate degree (30.3%), a bachelor degree (39.5%), or a master degree (25.9%).

Children. Children who met the following four eligibility criteria were included in the study: (1) caregiver consent obtained, (2) classified as *at risk* based on teacher report on the Early Screening Project (Feil, Severson, & Walker, 1998), (3) no identified developmental delay as measured by the Battelle Developmental Inventory–Second Edition Screener (Newborg, 2005), and (4) fluent in English. A total of 462 children participated in the study, with 230 children in the BEST in CLASS condition and 232 children in the comparison condition. The sample was 64.7% male and 35.3% female. The mean age for children in the BEST in CLASS and comparison conditions was 4.28 and 4.36 years, respectively. The majority (66.2%) of children were African American. In addition, 17.1% of children were Caucasian, 4.5% Hispanic, and 12.1% other or unidentified.

Data Collection and Measures

Data were collected at pretest in the fall of the academic school year and again at posttest in the spring of the year upon completion of the BEST in CLASS intervention. Pretest data were collected following teacher consent, child caregiver consent, and eligibility determination. Throughout the study, implementation fidelity data were collected, including coaching and dosage, as well as the level and quality of the key practices constituting the BEST in CLASS intervention. For a detailed description of the BEST in CLASS data collection, including coaching and teacher fidelity procedures, refer to the appendix, and see Sutherland et al. (2018), Sutherland et al. (2015), and Sutherland, McLeod, Conroy, Abrams, and Smith (2014).

Child outcome measures. Teacher report measures were utilized to assess children's social skills and levels of problem behavior. The Social Skills Improvement System–Rating Scales (SSIS-RS; Gresham & Elliott, 2008) and the C-TRF (Achenbach & Rescorla, 2000) were administered at pre- and posttest.

The SSIS-RS (Gresham & Elliott, 2008) is a teacher report measure designed to assess the social skills and problem behaviors of young children. The instrument includes 76 items rated on a 4-point scale from 0 (*never*) to 3 (*almost always*) that can be grouped into two subscales: Social Skills and Problem Behavior. On the Problem Behavior subscale, a standard score >115 is considered above average and >130, well above average. On the Social Skills subscale, a standard score <85 is considered below average and <70, well below average. Gresham and Elliott (2008) found strong levels of internal consistency and test-retest reliability. For the current study, internal consistency was calculated, resulting in Cronbach's alphas of .94 (pretest) and .95 (posttest) for the Social Skills subscale and .90 (pretest) and .92 (posttest) for the Problem Behavior subscale.

The C-TRF (Achenbach & Rescorla, 2000) is a teacher report measure designed to assess the problem behaviors of young children aged 18 months through 5 years. The 100 items are scored on a 3-point Likert scale, with responses ranging from 0 (*never true*) to 2 (*often true*). The three subscales include Externalizing Problems, Internalizing Problems, and Total Problems. A standard score in the range of 60 to 63 is considered borderline and ≥64, clinical. For the current study, internal consistency was high, with Cronbach's alphas across all subscales ranging from .86 (pretest) to .95 (posttest).

Fidelity measures. Throughout the BEST in CLASS intervention, data were obtained to assess the fidelity of

TABLE 1
Participant Characteristics

Demographics	Teachers, *n* (%)[a]		
	BEST in CLASS (*n* = 92)	Comparison (*n* = 93)	Total (*n* = 185)
Gender			
Female	91 (98.9)	90 (96.8)	181 (97.8)
Male	1 (1.1)	3 (3.2)	4 (2.2)
Race/ethnicity			
African American	41 (44.6)	47 (50.5)	88 (47.6)
Caucasian	44 (47.8)	42 (45.1)	86 (46.5)
Hispanic	3 (3.3)	2 (2.2)	5 (2.7)
Other	3 (3.3)	2 (2.2)	5 (2.7)
No report	1 (1)	0 (0)	1 (0.5)
Years teaching			
Mean	11.35	12.86	12.11
SD	8.97	9.99	9.48
Education level			
High school	2 (2.2)	4 (4.3)	6 (3.2)
Associate	28 (30.4)	28 (30.1)	56 (30.3)
Bachelor	37 (40.2)	36 (38.7)	73 (39.5)
Master	23 (25)	25 (26.9)	48 (25.9)
Doctoral	1 (1.1)	0 (0)	1 (0.5)
Other	1 (1.1)	0 (0)	1 (0.5)
	Children, *n* (%)[a]		
	BEST in CLASS (*n* = 230)	Comparison (*n* = 232)	Total (*n* = 462)
Gender			
Male	146 (63.5)	153 (65.9)	299 (64.7)
Female	84 (36.5)	79 (34.1)	163 (35.3)
Age			
Mean	4.28	4.36	4.32
SD	0.53	0.53	0.53
Race/ethnicity			
African American	151 (65.7)	155 (66.8)	306 (66.2)
Caucasian	37 (16.1)	42 (18.1)	79 (17.1)
Hispanic	11 (4.8)	10 (4.3)	21 (4.5)
Other	17 (7.4)	16 (6.9)	33 (7.1)
No report	14 (6.1)	9 (3.9)	23 (5.0)

[a]Values are presented as *n* (%) unless noted otherwise.

implementation by participating teachers and BEST in CLASS coaching staff. These included measures of coaching as well as teachers' use of the BEST in CLASS practices. Teacher implementation fidelity was evaluated in BEST in CLASS and comparison classrooms with the BEST in CLASS Adherence and Competence Scale (Sutherland et al., 2014). Staff in both states were trained in the use of the scale through a manual and collaborative training. A total of 756 observations were conducted at pre- and posttest across both classrooms: BEST in CLASS (*n* = 389) and comparison (*n* = 367). Interobserver agreement data were obtained on 27% (*n* = 204) of those observations. Intraclass correlations were calculated for each item on each subscale: .69 (*SD* = .08; range, .60–.82) for adherence and .44 (*SD* = .12; range, .27–.59) for competence.

The mean adherence for teacher implementation of the BEST in CLASS intervention practices in the business-as-usual comparison condition was 2.35 (*SD* = 1.50; range, 1.41–4.77) at pretest and 2.43 (*SD* = 1.61; range, 1.36–4.66) at posttest. The mean adherence for teacher implementation

of intervention practices in the BEST in CLASS condition was 2.47 (*SD* = 1.42; range, 1.52–4.69) at pretest and 4.53 (*SD* = 0.98; range, 3.72–5.99) at posttest. The mean competence of teacher implementation of BEST in CLASS intervention practices in the business-as-usual comparison condition was 3.82 (*SD* = 0.45; range, 3.32–4.44) at pretest and 4.30 (*SD* = 0.42; range, 3.95–5.04) at posttest. The mean competence of teacher implementation of intervention practices in the BEST in CLASS condition was 3.78 (*SD* = 0.37; range, 3.31–4.37) at pretest and 5.51 (*SD* = 0.33; range, 5.13–5.97) at posttest.

Data Analysis

To address our research questions, we used a transitional model (Agresti, 2007, section 9.4). In a transitional model for two time points, the outcome variable is status, which is defined for the current study as falling into (a) the clinical or borderline range or (b) within the normal range at Time 2 (i.e., posttest). Status was measured separately for the three C-TRF subscales and the two SSIS-SR subscales. The independent variables for the current investigation were condition (BEST in CLASS vs. business-as-usual comparison), pretest status (clinical or borderline vs. normal range), and state (A vs. B). In addition, the two-way interactions among the independent variables and the three-way interaction were included in the model. Condition, pretest status, and state were coded –.5 or .5, with BEST in CLASS, clinical or borderline, and State B coded .5. We included state because the study was conducted in two southeastern states. However, in reporting the results, we focus on the effect of treatment and the relationship of pre- to posttest status.

To account for nesting of children in teachers and teachers in schools, an approach developed in survey sampling was used (Heeringa, West, & Berglund, 2010). As applied in the current study, a single-level logistic regression model was conducted, but the standard errors were adjusted for nesting. We used this approach rather than a multilevel analysis to estimate a marginal model rather than a conditional model.[2] The logistic regression results were obtained via the complex procedure in M*plus* 7.4, which provides maximum likelihood estimates and cluster-robust standard errors (McNeish, Stapleton, & Silverman, 2016) for hypothesis testing.[3]

Results

Table 2 provides frequencies (percentages) and logits for posttest status on the C-TRF subscales Externalizing Problems, Internalizing Problems, and Total Problems. The marginal model represents the percentages in the clinical or borderline range (or, equivalently, the logits for the clinical or borderline range) and therefore accounts for the relationships of the percentages to condition, pretest status, and state. Results for Externalizing Problems in Table 2 indicate that none of the children with a pretest score in the normal range who resided in State B and were in BEST in CLASS had a clinical or borderline status at posttest. Due to the zero frequency, it is not possible to calculate the logit for this group or to validly estimate the C × P × S interaction based on the model. Inspection of the percentages in the clinical or borderline range at posttest suggests (a) little or no treatment effect in either state for children in the normal range at pretest and (b) a treatment effect of similar magnitude in both states for children in the clinical or borderline range at pretest. Therefore, we removed the C × P × S interaction from the model for Externalizing Problems, which permitted estimating the logistic regression equation.

Table 3 presents the results of the logistic regression analysis for C-TRF. For all three C-TRF subscales, the effect of condition was significant and negative, indicating that the BEST in CLASS intervention reduced the percentage of children in the clinical or borderline range. We computed the success rate difference (SRD; Kraemer & Kupfer, 2006) as an effect size for condition for each of the four combinations of pretest status and state. SRD is the difference in the proportion of children in the posttest clinical or borderline range for BEST in CLASS and the comparison condition. To quantify the effect size for the main effect of condition, we report the SRD averaged over the four combinations. The effect size for Externalizing Problems was –.119, indicating that, as averaged over pretest status and states, the percentage of children in the clinical or borderline range was lower by 11.9 percentage points for the BEST in CLASS condition in comparison with the control condition. For Internalizing Problems and Total Problems, the effect sizes were –.194 and –.173, respectively.

There was also a significant positive relationship of pre- and posttest status, indicating that, when compared with children who scored in the normal range at pretest, children who scored in the clinical or borderline range at pretest were more likely to be in the clinical or borderline range at posttest. We adapted SRD to quantify the strength of the relationship between pre- and posttest status by calculating the difference in the proportion of children in the posttest clinical or borderline range as compared with (a) children who were in the pretest clinical or borderline range and (b) those who were not. We refer to this effect size as the *clinical or borderline rate differential* (CBRD). The CBRDs for pretest status, as averaged across combinations of condition and states, were .493, .388, and .382 for Externalizing Problems, Internalizing Problems, and Total Problems, respectively. The relationship to pre- and posttest status was modified by a significant P × S interaction for both Externalizing Problems and Internalizing Problems. These interactions indicated that the pretest-posttest relationship effect was larger in State B. For Externalizing Problems, the CBRD was .563 in State B and .423 in State A. For Internalizing

TABLE 2
Frequencies (Percentages) and Logits for Posttest Status on C-TRF Subscales by State, Pretest Status, and Condition

C-TRF subscale		Posttest status, *n* (%)		
State: Pretest status	Condition	Normal	C or B	Logit
Externalizing Problems				
A: Normal	Comparison	16 (76.19)	5 (23.81)	−1.16
	BEST in CLASS	22 (75.86)	7 (24.14)	−1.15
A: C or B	Comparison	19 (25.00)	57 (75.00)	1.10
	BEST in CLASS	28 (42.42)	38 (57.58)	0.31
B: Normal	Comparison	28 (90.32)	3 (9.68)	−2.23
	BEST in CLASS	29 (100.00)	0 (0.00)	ND
B: C or B	Comparison	19 (28.36)	48 (71.64)	0.93
	BEST in CLASS	40 (49.38)	41 (50.62)	0.02
Internalizing Problems				
A: Normal	Comparison	41 (80.39)	10 (19.61)	−1.41
	BEST in CLASS	53 (89.83)	6 (10.17)	−2.18
A: C or B	Comparison	19 (41.30)	27 (58.70)	0.35
	BEST in CLASS	28 (77.78)	8 (22.22)	−1.25
B: Normal	Comparison	60 (89.55)	7 (10.45)	−2.15
	BEST in CLASS	61 (93.85)	4 (6.15)	−2.72
B: C or B	Comparison	8 (25.81)	23 (74.19)	1.06
	BEST in CLASS	24 (53.33)	21 (46.67)	−0.13
Total Problems				
A: Normal	Comparison	18 (64.29)	10 (35.71)	−0.59
	BEST in CLASS	30 (85.71)	5 (14.29)	−1.79
A: C or B	Comparison	22 (31.88)	47 (68.12)	0.76
	BEST in CLASS	32 (53.33)	28 (46.67)	−0.13
B: Normal	Comparison	36 (87.80)	5 (12.20)	−1.97
	BEST in CLASS	33 (86.84)	5 (13.16)	−1.89
B: C or B	Comparison	17 (29.82)	40 (70.18)	0.86
	BEST in CLASS	41 (56.94)	31 (43.06)	−0.28

Note. C-TRF, Caregiver-Teacher Report Form; Normal = normal range; C or B = clinical or borderline range; ND = not defined.

Problems, the CBRD was .521 for State B and .256 for State A.

Table 4 presents the frequencies (percentages) and logits for posttest status on the SSIS-RS Problem Behavior and Social Skills subscales. Table 5 presents the results of the logistic regression analysis on the relationships of the percentages of children in the clinical or borderline range at posttest to condition, state, and pretest status (i.e., classified in the normal vs. clinical or borderline range at pretest). The condition estimate was negative and significant for Problem Behavior, indicating that treatment reduced the percentage of children in the clinical or borderline range (SRD = −.186). There was a significant Condition × Pretest × State interaction for social skills. A simple interaction analysis resulted in a significant Condition × Pretest interaction for State A ($z = -2.104$, $p = .035$), with a significant condition effect for children in the clinical or borderline range at pretest ($z = -3.112$, $p = .002$, SRD = −.277). The simple Condition × Pretest interaction was not significant for State B. The condition effect for State B was significant ($z = -2.361$, $p = .018$, SRD = −.161). These analyses suggest that BEST in CLASS reduced in both states (A and B) the percentage of children who were in the clinical or borderline range at pretest. Additionally, pretest status was positively and significantly related to posttest status for both SSIS-RS subscales, indicating that children who scored in the clinical or borderline range at pretest had a tendency to be in the clinical or borderline range at posttest. CBRD effect sizes were .328 and .338 for Problem Behavior and Social Skills, respectively.

Discussion

With a focus on prevention science, the overall goal of the current study was to examine the influence of BEST in CLASS in producing clinically meaningful reductions in externalizing problem behaviors and improvements in social skills among young children identified as being at risk for

TABLE 3
Results of Logistic Regression Analysis for C-TRF Subscales

	C-TRF subscales											
	Externalizing Problems				Internalizing Problems				Total Problems			
	Estimate	*SE*	*z*	*p*	Estimate	*SE*	*z*	*p*	Estimate	*SE*	*z*	*p*
C	−0.710	0.321	−2.214	.027	−1.034	0.303	−3.418	.001	−0.786	0.259	−3.041	.002
P	2.668	0.308	8.670	.000	2.121	0.279	7.595	.000	1.861	0.256	7.272	.000
S	−1.050	0.334	−3.141	.002	0.135	0.275	0.491	.624	−0.383	0.245	−1.561	.119
C × P	−0.268	0.654	−0.409	.682	−0.725	0.505	−1.436	.151	−0.455	0.506	−0.901	.368
C × S	−0.349	0.431	−0.810	.418	0.303	0.605	0.501	.617	0.524	0.517	1.014	.311
P × S	1.677	0.611	2.743	.006	1.554	0.558	2.782	.005	0.716	0.512	1.399	.162
C × P × S					0.223	1.010	0.221	.825	−1.534	1.011	−1.517	.129

Note. C-TRF = Caregiver-Teacher Report Form; C = condition; P = pretest; S = state.

TABLE 4
Frequency (Percentages) and Logits for Posttest Status on SSIS-RS Subscales by Research State, Pretest Status, and Treatment

SSIS-SR subscale		Posttest status		
State: Pretest status	Condition	Normal	C or B	Logit
Problem Behavior				
A: Normal	Comparison	25 (67.57)	12 (32.43)	−0.73
	BEST in CLASS	23 (69.70)	10 (30.30)	−0.83
A: C or B	Comparison	11 (18.64)	48 (81.36)	1.47
	BEST in CLASS	30 (49.18)	31 (50.82)	0.03
B: Normal	Comparison	33 (70.21)	14 (29.79)	−0.86
	BEST in CLASS	41 (87.23)	6 (12.77)	−1.92
B: C or B	Comparison	18 (35.29)	33 (64.71)	0.61
	BEST in CLASS	38 (60.32)	25 (39.68)	−0.42
Social Skills				
A: Normal	Comparison	27 (90.00)	3 (10.00)	−2.20
	BEST in CLASS	18 (78.26)	5 (21.74)	−1.28
A: C or B	Comparison	18 (27.27)	48 (72.73)	0.98
	BEST in CLASS	39 (54.93)	32 (45.07)	−0.20
B: Normal	Comparison	23 (69.70)	10 (30.30)	−0.83
	BEST in CLASS	38 (88.37)	5 (11.63)	−2.03
B: C or B	Comparison	31 (47.69)	34 (52.31)	0.09
	BEST in CLASS	41 (61.19)	26 (38.81)	−0.46

Note. SSIS-RS, Social Skills Improvement System–Rating Scales; Normal = normal range; C or B = clinical or borderline range.

EBD. The current investigation extends our previous research on the BEST in CLASS intervention and was conducted as part of a 4-year efficacy study (see, Sutherland et al., 2018), which found statistically significant differences between children in the BEST in CLASS and control conditions, with better outcomes for children who received the intervention. Overall findings from the current investigation also found positive outcomes for children who received the BEST in CLASS intervention as compared with the control group. Specifically, the percentage of children receiving the BEST in CLASS intervention who scored in the clinical or borderline range—as measured by the Externalizing Problems, Internalizing Problems, and Total Problems subscales of the C-TRF and the Problem Behavior and Social Skills subscales of the SSIS-RS—was lower at posttest as compared with children who did not receive the intervention. However, for social skills, inferential tests suggest that the smaller percentage was not realized for children in State A who were not in the clinical or borderline range at pretest.

TABLE 5
Results of Logistic Regression Analysis for SSIS-RS Subscales

	SSIS-RS subscales							
	Problem Behavior				Social Skills			
	Estimate	*SE*	*z*	*p*	Estimate	*SE*	*z*	*p*
C	−0.907	0.285	−3.181	.001	−0.501	0.313	−1.604	.109
P	1.510	0.259	5.837	.000	1.690	0.304	5.554	.000
S	−0.633	0.285	−2.219	.027	−0.132	0.313	−0.423	.672
C × P	−0.651	0.517	−1.258	.208	−0.724	0.608	−1.190	.234
C × S	−0.275	0.570	−0.482	.630	−0.740	0.625	−1.184	.236
P × S	−0.053	0.517	−0.103	.918	−0.882	0.608	−1.449	.147
C × P × S	1.381	1.035	1.335	.182	2.742	1.217	2.253	.024

Note. SSIS-RS, Social Skills Improvement System–Rating Scales; C = condition; P = pretest; S = state.

Limited research has examined the clinical significance of preventive early childhood interventions aimed at young children who engage in significant problem behaviors that place them at risk for EBD. The current study adds to this important but overlooked area of research. Results of this study suggest that the BEST in CLASS intervention—an indicated preventive intervention that targets resilience factors related to problem behaviors among young children identified as being at risk for EBD—produces clinically significant outcomes. Researchers found that problem behaviors that interfere with young children's learning and development emerge early and persist well into the school years (NICHD Early Child Care Research Network, 2002; Vandell et al., 2010). Kazdin and Weiz (2003) suggested that the successful prevention and intervention of young children's problem behaviors depend on early treatment, when developmental trajectories are still malleable. Fortunately, when we intervene at a young age before these behaviors have become well established in children's repertoires, we may change their trajectory and prevent many of the lifelong negative consequences manifested through their ongoing existence (Bornstein et al., 2010).

Findings from the current study are important for another reason. Unlike other Tier 2 interventions that found positive effects on clinical change (e.g., Early Risers), BEST in CLASS is delivered by teachers during naturally occurring classroom activities. The BEST in CLASS intervention focuses on increasing the use of key instructional practices during activities in which children are engaged throughout their school day, thus decreasing problem behaviors during these activities and improving interactions and relationships between teachers and young children as they naturally interact. Research suggested that improving teacher-child interactions (Early et al., 2006) and relationships (O'Connor & McCartney, 2007) can have positive long-term effects on child functioning. Results from the current study suggest that BEST in CLASS may have some influence on helping children move from a clinical or borderline category to the normal range of functioning. For example, as reported in Table 2, 42% to 78% of children in the BEST in CLASS condition with clinical or borderline scores on the C-TRF at pretest had normal scores at posttest, and as seen in Table 4, 49% to 61% of the children in the BEST in CLASS condition with clinical or borderline scores on the SSIS-RS at pretest had normal posttest scores. These findings highlight the potential importance of teacher-child interactions and relationships as a protective factor in promoting positive child outcomes. In addition, feasible teacher-delivered Tier 2 interventions that can help move children from a clinical or borderline range to a more normal developmental range, such as BEST in CLASS, can be particularly helpful in identifying nonresponders that may benefit from more intensive Tier 3 interventions.

Our results are similar to and support findings from previous research indicating that children who received intervention and scored in the clinical or borderline range at pretest raised their performance to within the normal range at posttest (e.g., Morrison & Bratton, 2010; Walker et al., 1998). However, previous research and our findings suggest that a number of additional factors may influence clinical outcomes. For example, some researchers found that the intensity of problem behaviors affects the clinical significance of the findings. August et al. (2003) examined the clinical significance of the Early Risers program and found differences between the experimental and treatment groups for children who demonstrated significant aggression but not for children who demonstrated mild aggression. Additionally, researchers found that some children's behavior actually continues to worsen without treatment (e.g., August et al., 2003). Another factor that might influence the outcomes of the current investigation is the intensity or dosage of the intervention. The BEST in CLASS intervention targets children at risk for EBD and is delivered throughout the school day. In contrast, some universal, or Tier 1, interventions

focus on all children (e.g., Frey et al., 2000) and are delivered during a series of short small group intervention sessions (e.g., Walker et al., 1998); therefore, they might not deliver a sufficient dosage of intervention to influence clinically significant outcomes for individual children at elevated risk. Unfortunately, these types of analyses were beyond the scope of the current investigation. However, there has been a recent emphasis on authentic change agents (i.e., teachers) implementing interventions within the context of the school day (e.g., Feil et al., 2014; Sutherland et al., 2018). Future research may be able to address these types of questions. Additionally, researchers will want to consider that, even though children may receive Tier 1 and 2 interventions, a small percentage of children continue to need additional supports (i.e., Tier 3). It may be that some children who require Tier 3 services are exposed to high-quality Tier 1 and 2 practices but require additional supports due to their individual developmental learning needs (e.g., Dunlap et al., 2015).

As noted, results differed across the two states involved in the study on some of the outcome variables. Differences in implementation of BEST in CLASS between States A and B are not a likely explanation for state differences in outcomes, because BEST in CLASS is a manualized intervention and no variation in implementation occurred between the states. A more likely explanation may be several demographic differences between participants in States A and B. For example, caregivers in State A were, on average, younger and had less education in comparison with caregivers in State B. Additionally, household income was lower in State A. All of these are risk factors that can contribute to EBD. Last, teachers in State B were more highly educated than teachers in State A.

As the literature suggested, the development of EBD is progressive and influenced by early exposure to a number of child risk factors (e.g., temperament, development delays; Benson & Aman, 1999; Reid, 1993), family risk factors (e.g., poverty; Kaiser, Cai, Hancock, & Foster, 2002), and school risk factors (e.g., coercive teacher-child relationships, negative classroom climate; Curby et al., 2009). BEST in CLASS is designed to primarily target risk factors that might be associated with classroom settings, such as teacher-child interactions and classroom climate; thus, the training of teachers might have played a role in teachers' uptake of BEST in CLASS practices within their ongoing classroom instruction.

Limitations

The findings from the current study are promising, but several limitations should be considered when interpreting the findings. First, in this investigation, we report findings from teacher report measures rather than direct observations of children's behavior in their classrooms. Although our findings are supported by two standardized measures and although the C-TRF and the SSIS-RS are both psychometrically sound and common instruments used for research and clinical purposes, direct observations of children's problem behaviors and social skills would provide additional information and help to substantiate findings. Additionally, we did not include data from children's families. Problem behaviors often occur across home and school settings. Obtaining further information from families to examine the pre- and posttest differences in home settings might be important when interpreting findings related to clinical significance. Finally, an area of utmost importance to our findings that we are unable to report is longitudinal data. Therefore, we are unsure whether the improvements that children in the BEST in CLASS intervention group made between pre- and posttest would continue over time and whether these improvements affected children's future identification of EBD. Clearly, these limitations suggest that our findings be viewed with caution and are directions for future research.

Conclusion

In summary, these findings indicate that when children received the BEST in CLASS instructional practices intervention targeted at promoting engagement and positive teacher-child interactions, their SEL outcomes improved. As discussed earlier, a number of social-emotional and behavioral preventive interventions exist. However, few investigators have focused on the clinical and practical significance of the intervention outcomes (Kazdin, 2003). Findings from research often indicate that children's behavior statistically improves following an intervention. Still, the field has yet to emphasize research that examines whether these gains are meaningful from a clinical perspective, meaning that children's social, emotional, and behavioral competence skills are within the normal range of development following intervention, thereby potentially reducing their risk for developing EBD. Additional research should be conducted to examine the clinical significance of findings as well as various factors that might influence outcomes, such as the level and intensity of behaviors and dosage. Through the implementation of effective early intervention programs, we may be able to reduce the occurrence of young children entering school at risk for or later identified with EBD. As suggested by Hawkins and colleagues (2015), "prevention is the best investment we can make, and the time to make it is now" (p. 1).

Appendix

BEST in CLASS Professional Development Components

Teachers participate with their coaches in the BEST in CLASS teacher training workshop: a 1-day didactic training session to orient teachers to the specific concepts and

instructional practices that compose BEST in CLASS (i.e., rules, precorrection, opportunities to respond, behavior-specific praise, corrective feedback, and instructive feedback). The training includes modeling, exemplar videos, and interactive activities that help teachers begin to identify specific behaviors and contexts where problem behaviors most frequently occur in their classrooms.

The BEST in CLASS teacher manual includes seven modules: one for each instructional practice and a concluding module designed to facilitate teachers' ability to link multiple practices for greater impact. Each module includes a definition of the practice, practical examples and nonexamples to develop understanding of the practice, an explanation of how to implement the practice at a Tier 2 level with targeted children, and suggestions for sharing this information with families. Teachers receive the manual at the workshop and continue to utilize it as a resource throughout the coaching portion of the intervention.

BEST in CLASS practice-based coaching begins about 1 week after the teacher-training workshop. The coaching process is cyclical and includes several key components, including modeling and/or prompting, focused observation, teacher and coach reflection, facilitated instruction of new skills, and action planning. The process lasts 14 weeks, with 2 weeks spent on each module in the BEST in CLASS teacher manual. Each week, an individualized action plan is collaboratively created by the coach and the teacher to facilitate the teacher's use of one or more practices with targeted children in the contexts in which their problem behaviors are most likely to occur. For a more detailed description of the coaching process as well as how coaches are trained and evaluated, see Sutherland, Conroy, Vo, and Ladwig (2015).

Data Collection and Measures

Screening. Child screening measures included the Early Screening Project (ESP; Feil et al., 1998), the Battelle Developmental Inventory–Second Edition Screener (BDI-2 Screener; Newborg, 2005), and a program-designed language screener. The ESP is a multistep teacher report screening tool. Stage 1 of the ESP presents teachers with a list of externalizing and internalizing behaviors. Participating teachers were asked to select and rank (by severity) up to five children in their class who demonstrated externalizing behaviors. The teacher was then asked to complete the language screener for those children to ensure that each identified child met the criteria for fluency in English. Next, teachers completed Stage 2 of the ESP, which includes a series of checklists grouped into subscales, including the Critical Events Index, Aggressive Behavior Scale, Adaptive Behavior Scale, and Maladaptive Behavior Scale. Test-retest reliability coefficients for subscales of the ESP range from .74 to .90, and concurrent validity scores range from .44 to .88 (Feil et al., 1998), correlating to the Externalizing Problems subscale of the C-TRF (Achenbach & Rescorla, 2000).

Finally, the BDI-2 (Newborg, 2005) was individually administered to children by trained research staff. The BDI-2 is a screening tool that identifies children from birth through age 8 years who are at risk for developmental delay. In the current study, the Adaptive, Communication, and Cognitive subscales of the BDI-2 determined if the child was performing in these developmental areas within the normal range. Reliability coefficients for the normative sample (aged 0–95 months) were calculated via the split-half method and corrected with the Spearman-Brown formula. Pearson correlations were high for both a 2-year-old sample and a 4-year-old sample across subscales, with scores ranging from .88 to .89 for the former and between .87 and .92 for the latter. Following the BDI-2, the three children in each class with the highest level of risk, as identified on the ESP, who also met the other eligibility criteria were selected as participants for the study.

Fidelity measures. Throughout the BEST in CLASS intervention, data were obtained to assess the fidelity of implementation by participating teachers and BEST in CLASS coaching staff. These included measures of coaching as well as teachers' use of the BEST in CLASS practices. Coach training was manualized and included exemplar videos of teachers using the BEST in CLASS practices, as well as interactive activities and practice. Coaches from both states were trained according to the manualized procedures. Prior to implementing the coaching component, coaches completed a checkout process, which included role-playing, self-evaluation, and scoring of a video-recorded coaching session by a master coaching observer. Additionally, all coaches were observed during the coaching process, and fidelity data were obtained throughout implementation of the coaching component (for further details, see Sutherland et al., 2015). Dosage of coaching was self-recorded by the coach following each session.

Teacher implementation fidelity was evaluated in BEST in CLASS and comparison classrooms via the BEST in CLASS Adherence and Competence Scale (Sutherland et al., 2014). The scale is an observational tool designed to evaluate teachers' adherence and competence in delivery of the BEST in CLASS practices. The adherence form utilizes a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very extensive*) to measure the extensiveness of teachers' use of BEST in CLASS practices with focal children. Thoroughness, frequency, rate, and duration of teacher behaviors are considered for scoring purposes. The competence form is based on a 7-point Likert scale ranging from 1 (*very poor*) to 7 (*excellent*) to measure the quality of teachers' use of BEST in CLASS practices. Scores are based on skillfulness, responsiveness, timing, and appropriateness of teachers' use of the practices.

Results of Three-Level Analysis

In M*plus*, Bayesian estimation of a probit regression model must be used to conduct a three-level analysis of a dichotomous variable. We used default prior distributions in

TABLE A1
Results of Probit Regression Analysis for C-TRF Subscales

	C-TRF subscales								
	Externalizing Problems			Internalizing Problems			Total Problems		
	Estimate	LL	UL	Estimate	LL	UL	Estimate	LL	UL
C	−0.917	−1.869	−0.155	−1.165	−2.381	−0.307	−1.031	−1.964	−0.305
P	2.877	2.015	4.172	2.530	1.610	4.116	2.170	1.409	3.308
S	−1.180	−2.243	−0.362	0.250	−0.599	1.183	−0.350	−1.188	0.385
C × P	0.227	−1.266	1.896	−0.970	−2.803	0.563	−0.045	−1.476	1.423
C × S	−0.251	−1.593	1.013	0.568	−1.115	2.410	0.492	−0.985	1.968
P × S	1.996	0.435	4.100	2.343	0.831	4.441	0.762	−0.558	2.335
C × P × S				0.189	−2.866	3.305	−1.673	−4.702	1.006

Note. C-TRF = Caregiver-Teacher Report Form; LL, lower limit; UL = upper limit; C = condition; P = pretest; S = state. Lower and upper limits are based on a 95% credibility interval.

TABLE A2
Results of Probit Regression Analysis for SSIS-RS Subscales

	SSIS-RS subscales					
	Problem Behavior			Social Skills		
	Estimate	LL	UL	Estimate	LL	UL
C	−1.275	−2.269	−0.543	−0.520	−1.509	−0.021
P	1.935	1.168	3.138	1.487	0.945	2.192
S	−0.806	−1.688	−0.084	−0.119	−0.659	0.408
C × P	−0.567	−2.111	0.788	−0.480	−1.505	0.533
C × S	−0.400	−1.926	1.025	−0.600	−1.623	0.371
P × S	−0.351	−2.045	1.030	−0.627	−1.784	0.354
C × P × S	1.491	−1.323	4.581	1.977	−0.060	4.065

Note. SSIS-RS = Social Skills Improvement System–Rating Scales; LL, lower limit; UL = upper limit; C = condition; P = pretest; S = state. Lower and upper limits are based on a 95% credibility interval.

M*plus*. We included variance component at the teacher and school levels, a School × Condition variance component, and the covariance of the school and the School × Condition random effects. We also included the same fixed effects used in the logistic regression models. Results for the C-TRF variables are presented in Table A1 and include estimates and lower and upper limits of a 95% credibility interval. A credibility interval that does not contain zero indicates support for the belief that the parameter being estimated is not equal to zero. For all three C-TRF variables, there was support for the belief that the condition effects are not zero and that pretest status is related to posttest status. In addition, evidence indicates a nonzero Pretest × State interaction for the Internalizing Problems subscale. Table A2 contains results for the SSIS-RS variables. For both variables, there was support for the belief that the condition effects are not zero and that pretest status is related to posttest status. Results also indicate a nonzero state effect for problem behavior. The Bayesian results for the three-level model were in agreement with the design-based results for the single-level model, with the exceptions that the three-level model indicated a condition effect for social skills and did not indicate a P × S interaction for the Internalizing Problems subscale or a Condition × Pretest × State interaction for the Social Skills subscale.

Authors' Note

The opinions expressed in this article are those of the authors and may or may not represent views of the U.S. Department of Education, Institute of Education Sciences, U.S. Department of Health and Human Services, and the U.S. National Institutes of Health.

Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A110173 to the University of Florida.

Notes

1. In the multisite cluster randomized design used in the current research, *site* refers to schools.

2. A marginal model estimates the effect of treatment averaged across all participants. In a conditional model, the fixed effects parameters in the multilevel model are defined conditionally on the random effects in the model (see Agresti, 2007).

3. We also estimated a three-level model and report the results for this model in the appendix.

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.

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). New York, NY: Wiley.

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

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

Barbarin, O. A. (2007). Mental health screening of preschool children: Validity and reliability of ABLE. *American Journal of Orthopsychiatry*, *77*, 402–418.

Benson, B. A., & Aman, M. G. (1999). Disruptive behavior disorders in children with mental retardation. In H. Quay, & A. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 559–578). New York, NY: Springer.

Bornstein, M. H., Chun-Shin, H., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Development and Psychopathology*, *22*, 717–735. doi:10.1017/S0954579410000416

Brauner, C. B., & Stephens, C. B. (2006). Estimating the prevalence of early childhood serious emotional/behavioral disorders: Challenges and recommendations. *Public Health Reports*, *121*(3), 303–310.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U. (2005). *Making human beings human: Bioecological perspectives on human development*. Thousand Oaks, CA: Sage.

Bulotsky-Shearer, R. J., Bell, E. R., & Dominguez, X. (2012). Latent profiles of problem behavior within learning, peer, and teacher contexts: Identifying subgroups of children at academic risk across the preschool year. *Journal of School Psychology*, *50*, 775–798.

Bulotsky-Shearer, R. J., Dominguez, X., & Bell, E. R. (2012). Preschool classroom behavioral context and school readiness outcomes for low-income children: A multilevel examination of child- and classroom-level influences. *Journal of Educational Psychology*, *104*, 421–438.

Bulotsky-Shearer, R. J., Manz, P. H., Mendez, J. L., McWayne, C. M., Sekino, Y., & Fantuzzo, J. W. (2012). Peer play interactions and readiness to learn: A protective influence for African American preschool children from low-income households. *Child Development Perspectives*, *6*(3), 225–231.

Carter, A. S., Wagmiller, R. J., Gray, S. O., McCarthy, K. J., Horwitz, S. M., & Briggs-Gowan, M. J. (2010). Prevalence of *DSM-IV* disorder in a representative, healthy birth cohort at school entry: Sociodemographic risks and social adaptation. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 686–689.

Collaborative for Academic, Social, and Emotional Learning. (2013). *CASEL guide: Effective social and emotional learning programs—Preschool and elementary school edition*. Chicago, IL: Author.

Committee for Children. (1991). *Second Step: A violence prevention curriculum preschool-kindergarten*. Seattle, WA: Committee for Children.

Conduct Problems Prevention Research Group. (1992). A developmental and clinical model for the prevention of conduct disorder: The FAST Track Program. *Development and Psychopathology*, *4*(4), 509–527.

Conduct Problems Prevention Research Group. (2002). Using the Fast Track randomized prevention trial to test the early-starter model of the development of serious conduct problems. *Development and Psychopathology*, *14*(4), 925–943.

Conroy, M. A., & Sutherland, K. S. (2008). *Promoting social, emotional, and behavioral competence in young high-risk children: A preventative classroom-based early intervention model*. Unpublished document, Department of Special Education and Disability Policy, Virginia Commonwealth University, Richmond, VA.

Conroy, M. A., Sutherland, K. S., Algina, J. J., Ladwig, C. N., Werch, B., Martinez, J., . . . Gyure, M. (2017). *BEST in CLASS: A professional development intervention fostering high quality classroom experiences for young children with problem behavior*. Manuscript submitted for publication.

Conroy, M. A., Sutherland, K. S., Algina, J. J., Wilson, R. E., Martinez, J. R., & Whalon, K. J. (2015). Measuring teacher implementation of the BEST in CLASS intervention program and corollary child outcomes. *Journal of Emotional and Behavioral Disorders*, *23*(3), 144–155.

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), 912–925.

Dodge, K. A., & Conduct Problems Prevention Research Group. (2007). Fast Track randomized controlled trial to prevent externalizing psychiatric disorders: Findings from Grades 3 to 9. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(10), 1250–1262.

Dunlap, G., & Fox, L. (2009). Positive behavior support and early intervention. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of positive behavior support* (pp. 49–71). New York, NY: Springer.

Dunlap, G., Lee, J. K., Joseph, J. D., & Strain, P. (2015). A model for increasing the fidelity and effectiveness of interventions for challenging behaviors: Prevent-teach-reinforce for young children. *Infants & Young Children*, *28*(1), 3–17.

Dunlap, G., Strain, P. S., Fox, L., Carta, J., Conroy, M., Smith, B., . . . Sowell, C. (2006). Prevention and intervention with young children's challenging behavior: A summary of current knowledge. *Behavioral Disorders*, *32*, 29–45.

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.

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*(5), 1159–1175.

Feil, E. G., Frey, A., Walker, H. M., Small, J. W., Seeley, J. R., Golly, A., & Forness, S. R. (2014). The efficacy of a home-school intervention for preschoolers with challenging behaviors: A randomized controlled trial of preschool First Step to Success. *Journal of Early Intervention*, *36*(3), 151–170. doi:10.1177/1053815114566090.

Feil, E. G., Severson, H. H., & Walker, H. M. (1998). Screening for emotional and behavioral delays: The Early Screening Project. *Journal of Early Intervention*, *21*, 252–266. doi:10.1177/105381519802100306

Feil, E. G., Small, J. W., Forness, S. R., Serna, L. A., Kaiser, A. P., Hancock, T. B., . . . Lopez, M. L. (2005). Using different measures, informants, and clinical cut-off points to estimate prevalence of emotional or behavioral disorders in preschoolers: Effects on age, gender, and ethnicity. *Behavioral Disorders*, *30*, 375–391.

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

Frey, K. S., Hirschstein, M. K., & Guzzo, B. A. (2000). Second Step: Preventing aggression by promoting social competence. *Journal of Emotional and Behavioral Disorders*, *8*(2), 102–112.

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

Gresham, F., & Elliott, S. N. (2008). *Social Skills Improvement System (SSIS) 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. doi:10.1111/1467-8624.00301

Hawkins, J. D., Jenson, J. M., Catalano, R., Fraser, M. W., Botvin, G. J., Shapiro, V., . . . Rotheram-Borus, M. J. (2015). *Unleashing the power of prevention*. Washington, DC: National Academy of Medicine.

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). *Applied survey analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Hemmeter, M. L., Fox, L., & Snyder, P. (2013). A tiered model for promoting social-emotional competence and addressing challenging behavior. In V. Buysse, & E. Peisner-Feinberg (Eds.), *Handbook of response to intervention in early childhood* (pp. 85–101). Baltimore, MD: Brookes.

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

Hemmeter, M. L., Snyder, P. A., Fox, L., & Algina, J. (2016). Evaluating the implementation of the pyramid model for promoting social-emotional competence in early childhood classrooms. *Topics in Early Childhood Special Education*, *36*, 133–146.

Hemmeter, M. L., Snyder, P. A., Kinder, K., & Artman, K. (2011). Impact of performance feedback delivered via electronic mail on preschool teachers' use of descriptive praise. *Early Childhood Research Quarterly*, *26*, 96–109.

Humphries, M. L., & Keenan, K. E. (2006). Theoretical, developmental and cultural orientations of school-based prevention programs for preschoolers. *Clinical Child and Family Psychology Review*, *9*(2), 135–148.

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.

Kazdin, A. E. (2003). Clinical significance: Measuring whether interventions make a difference. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (3rd ed., pp. 691–710). Washington, DC: American Psychological Association.

Kazdin, A. E., & Weiz, J. R. (Eds.). (2003). *Evidence-based psychotherapies for children and adolescents*. New York, NY: Guilford Press.

Kraemer, H. C., & Kupfer, D. J. (2006). Size of treatment effects and their importance to clinical research and practice. *Biological Psychiatry*, *59*, 990–996. doi:10.1016/j.biopsych.2005.09.014

McCabe, P. C., & Altamura, M. (2011). Empirically valid strategies to improve social and emotional competence of preschool children. *Psychology in the Schools*, *48*(5), 513–540.

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2016). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. doi:10.1037/met0000078

Morrison, M. O., & Bratton, S. C. (2010). Preliminary investigation of an early mental health intervention for Head Start programs: Effects of child teacher relationship training on children's behavior problems. *Psychology in the Schools*, *47*(10), 1003–1017.

Myers, S. S., & Pianta, R. C. (2008). Developmental commentary: Individual and contextual influences on student-teacher relationships and children's early problem behaviors. *Journal of Clinical Child & Adolescent Psychology*, *37*(3), 600–608.

National Research Council and Institute of Medicine. (2009a). Defining the scope of prevention. In M. E. O'Connell, T. Boat, & K. E. Warner (Eds.), *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities* (pp. 59–70). Washington, DC: National Academies Press.

National Research Council and Institute of Medicine. (2009b). Introduction. In M. E. O'Connell, T. Boat, & K. E. Warner (Eds.), *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities* (pp. 15–32). Washington, DC: National Academies Press.

Nelson, J. R., Stage, S., Duppong-Hurley, K., Synhorst, L., & Epstein, M. H. (2007). Risk factors predictive of the problem behavior of children at risk for emotional and behavioral disorders. *Exceptional Children*, *73*(3), 367–379.

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

NICHD Early Child Care Research Network. (2002). Early child care and children's development prior to school entry: Results from the NICHD study of early childhood care. *American Educational Research Journal*, *39*(1), 133–164.

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.

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. doi:10.1207/s1532480xads0903_2

Pianta, R. C., & Stuhlman, M. W. (2004). Teacher-child relationships and children's success in the first years of school. *School Psychology Review*, *33*(3), 444–458.

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*, 188–216.

Reid, J. B. (1993). Prevention of conduct disorder before and after school entry: Relating interventions to developmental findings. *Development and Psychopathology*, *5*(1–2), 243–262.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior assessment scale for children*. Bloomington, MN: Pearson Assessments.

Sameroff, A. (Ed.). (2009). *The transactional model of development: How children and contexts shape each other*. Washington, DC: American Psychological Association.

Snyder, P. A., Hemmeter, M. L., & Fox, L. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Topics in Early Childhood Special Education*, *35*, 133–143.

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*. New York, NY: William T. Grant Foundation.

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, . . . Wilcox, B. (2000). Applying positive behavioral support and functional behavioral assessment in schools. *Journal of Positive Behavioral Interventions*, *2*, 131–143.

Sumi, W. C., Woodbridge, M. W., Javitz, H. S., Thornton, S. P., Wagner, M., Rouspil, K., . . . Severson, H. H. (2012). Assessing the effectiveness of First Step to Success: Are short-term results the first step to long-term behavioral improvements? *Journal of Emotional & Behavioral Disorders*, *21*(1), 66–78.

Sutherland, K. S., Conroy, M. A., Algina, J., Ladwig, C., Jessee, G., & Gyure, M. (2018). Reducing child problem behaviors and improving teacher-child interactions and relationships: A randomized controlled trial of BEST in CLASS. *Early Childhood Research Quarterly*, *42*, 31–43. doi.org/10.1016/j.ecresq.2017.08.001

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*(1), 21–33.

Sutherland, K. S., McLeod, B. D., Conroy, M. A., Abrams, L., & Smith, M. M. (2014). Preliminary psychometric properties of the BEST in CLASS Adherence and Competence Scale. *Journal of Emotional and Behavioral Disorders*. *22*(4), 249–259. doi:10.1177/1063426613497258

U.S. Department of Health and Human Services. (2015). *Head Start early learning outcomes framework: Ages birth to five*. Washington, DC: Administration for Children and Families Office of Head Start.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development*, *81*(3), 737–756.

Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Severson, H. H., & Feil, E. G. (1998). First Step to Success an early intervention approach for preventing school antisocial behavior. *Journal of Emotional and Behavioral Disorders*, *6*(2), 66–80.

Walker, H. M., Seeley, J. R., Small, J., Severson, H. H., Graham, B. A., Feil, E. G., . . . Forness, S. R. (2009). A randomized controlled trial of the First Step to Success early intervention: Demonstration of program efficacy outcomes within a diverse, urban school district. *Journal of Emotional and Behavioral Disorders*, *17*(4), 197–212. doi:10.1177/1063426609341645

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, MD: Brookes.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, *33*, 105–124. doi:10.1207/S15374424JCCP3301_11

Webster-Stratton, C., Reid, M. J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the Incredible Years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry*, *49*(5), 471–488.

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2007). The scientific base linking social and emotional learning to school success. *Journal of Educational and Psychological Consultation*, *17*, 191–210.

Authors

MAUREEN A. CONROY is the Anita Zucker Endowed Professor in Early Childhood Studies and codirector of the Anita Zucker Center for Excellence in Early Childhood Studies, University of Florida, PO Box 117050, Gainesville, FL 32611; mconroy@coe.ufl.edu. Her research focuses on the prevention and amelioration of social-emotional learning skills among young children at risk for school failure.

KEVIN S. SUTHERLAND is a professor in the School of Education, Virginia Commonwealth University, PO Box 842020, Richmond, VA 23284; kssuther@vcu.edu. His research focuses on intervention development and implementation issues for students with or at risk for emotional behavioral disorders.

JAMES ALGINA is an emeritus professor of research and evaluation methodology, University of Florida, PO Box 117047, Gainesville FL 32611; algina@ufl.edu. He is a fellow of the American Educational Research Association and Division 5 of the American Psychological Association. His research interests include classical test theory, item response theory, multilevel modeling, and structural equation modeling.

BRITTANY WERCH is a postdoctoral affiliate in the Anita Zucker Center for Excellence in Early Childhood Studies, University of Florida, PO Box 117050, Gainesville, FL; bwerch@ufl.edu. Her research focuses on early intervention, implementation science, and improving outcomes for children at risk for disabilities.

CRYSTAL LADWIG is the BEST in CLASS project coordinator, Anita Zucker Center for Excellence in Early Childhood Studies, University of Florida, PO Box 117050, Gainesville, FL; cladwig2@cox.net. Her research interests include intervention strategies that support children with social-emotional learning difficulties and professional development for early childhood teachers.