Research Report

Effects of BEST in CLASS–Web on Teacher Outcomes: A Preliminary Investigation

Maureen A. Conroy[1] (iD)**, Kevin S. Sutherland[2],
Kristen L. Granger[2], Katerina M. Marcoulides[1]** (iD)**,
Edward Feil[3], Jessica Wright[2], Mayra Ramos[2],
and Alexandra Montesion[1]**

Abstract

This study examined the effects of the BEST in CLASS intervention professional development component when delivered online in comparison with in person and a control group. A total of 29 early childhood teachers serving young children demonstrating challenging behaviors were included as participants. Data were collected on teachers' classroom quality, self-efficacy, implementation of the BEST in CLASS practices, acceptability of the intervention, and costs of delivery. Positive outcomes on several dimensions of classroom quality were found for the BEST in CLASS–Web condition when compared with the BEST in CLASS and control conditions. Alternatively, teachers' sense of self-efficacy was greater when BEST in CLASS was delivered in person when compared with teachers who received online delivery or those in the control condition. Finally, teachers in both treatment conditions increased their use of the practices in comparison with control teachers. Teachers in both treatment conditions rated BEST in CLASS with high levels of acceptability. Implications and future research directions are discussed.

Keywords

early childhood, online professional development, practice-based coaching, social-emotional intervention

Years of research suggest that high-quality early childhood care and education provides the foundation for positive school outcomes and future school success (e.g., Schindler et al., 2015). However, almost a quarter of young children who attend early childhood programs display chronic challenging behaviors that place them at future risk for the development of emotional/behavioral disorders (EBD; Bulotsky-Shearer et al., 2012; McCabe & Altamura, 2011) and later

[1]University of Florida, Gainesville, USA
[2]Virginia Commonwealth University, Richmond, USA
[3]Oregon Research Institute, Eugene, USA

Corresponding Author:
Maureen A. Conroy, Anita Zucker Center of Excellence in Early Childhood Studies, University of Florida, Box 117050, Gainesville, FL 32611, USA.
Email: mconroy@coe.ufl.edu

school failure (Blair & Raver, 2012; Yoshikawa et al., 2012). Regrettably, many young children do not have access to quality early learning programs (Cloney et al., 2016) and many practitioners within these programs do not have the knowledge to address the socio-emotional and behavioral needs of young children attending these programs (Sutherland et al., 2013). To make quality early learning experiences a reality for *all* children, the field needs more teachers skilled in delivering evidence-based programs and practices that address social-emotional learning and behavioral well-being.

Research indicates that many early childhood teachers lack access to systematic and sustainable high-quality professional development opportunities necessary to further their knowledge and skills (Becker & Domitrovich, 2011; Bruder et al., 2009). Programs need to provide high-quality professional development and support teachers' implementation of evidence-based practices with children who demonstrate chronic challenging behaviors (Artman-Meeker & Hemmeter, 2012; Lloyd & Modlin, 2012). Fortunately, a number of evidence-based early intervention programs and practices that target the social-emotional needs of young children exist (e.g., Teaching Pyramid; Fox et al., 2003). Typically, these programs include a professional development component to support teachers' use of specific social-emotional learning practices with children in their classrooms, who are in need. Professional development models that engage teachers as active participants and include focused observation, practice, feedback, and reflection provided through coaching are effective in promoting teachers' implementation and sustained use of effective practices in their classrooms (Joyce & Showers, 2002; Snyder et al., 2012). Practice-based coaching (see Snyder et al., 2015, 2022) is one such model that is widely used and has been demonstrated to be effective in changing early childhood teachers' use of practices across a variety of content areas, including social and behavioral skills (e.g., see Conroy, Sutherland, Algina, et al., 2019; Fox et al., 2011; Hemmeter et al., 2016).

While high-quality professional development, including coaching, has resulted in teachers' increased use of evidence-based practices, until recently much of the delivery has occurred in person. Given the advances in technology and need for flexibility in the delivery of professional development activities, the feasibility of and need for online professional development activities has increased (e.g., see Artman-Meeker et al., 2014; Downer et al., 2011). While more research is currently underway in relation to web-based delivery of professional development, this study is among one of the few to compare outcomes across delivery conditions (e.g., in-person, web-based, control). BEST in CLASS–Web (Conroy & Sutherland, 2011) was designed to address a critical need in the field because it provides web-based professional development, including practice-based coaching, to increase teachers' focused use of evidence-based practices with young children who demonstrate chronic challenging behaviors.

Web-based professional development initiatives are increasingly becoming a more viable and economical option to a larger number of early childhood programs (Dede et al., 2009). Web-based professional development allows accessibility to communities that might not have access (e.g., rural communities) and is considered a cost-efficient delivery mode that enhances scalability (Castle & McGuire, 2010; Otte et al., 2014). In addition, web-based professional development may reduce the barrier of cost associated with in-person coaching (Knight, 2012), allowing for a different type of professional development infrastructure. Finally, consistency of professional development may be facilitated by using a web-based platform with online modules that follow a standard protocol that allow better automation of training and coaching (Otte et al., 2014). Multimedia components such as live videoconferencing can afford coaches opportunities to tailor content and instructional methods to meet teachers' unique learning needs and context-specific goals. Web-based professional development also has many teacher-level benefits, including flexibility and increased compatibility with teachers' schedules (Reeves & Pedulla, 2011). When aligned with adult learning principles (Ke &

Xie, 2009; Knowles, 1984), teachers are able to work through web-based content in a more convenient location, work at their own pace to achieve objectives most relevant to their needs, while continuing to engage in coaching activities that promote learning (e.g., shared goal setting and reflection), and build alliance. On the contrary, in-person coaching may be a better match for teachers who require more "hands-on" demonstrations and support in the classroom, are not comfortable with technology, or do not have access. Simply put, "one size [delivery format of professional development and coaching] does not fit all" and there is a need to offer different modalities of professional development and coaching (e.g., in-person, web-based) to meet diverse needs.

Given the utility of high-quality professional development, including practice-based coaching, to increase teachers' use of effective practices to support children's social-emotional learning needs and address challenging behaviors, there is a need to examine the initial promise of web-based delivery models in comparison with in-person professional development. In this investigation, we conducted a preliminary investigation of the BEST in CLASS intervention delivered over the web (i.e., BEST in CLASS–Web; Conroy & Sutherland, 2011) in comparison with delivery in person (i.e., BEST in CLASS) and a control group. This investigation addressed the following research questions:

Research Questions 1 (RQ1): Is there equivalence with regard to coaching fidelity, coaching dosage, and teacher implementation of intervention practices between the BEST in CLASS–Web and BEST in CLASS conditions?

Research Questions 2 (RQ2): Does BEST in CLASS–Web improve teacher outcomes and is the intervention effect of BEST in CLASS–Web equivalent to that of BEST in CLASS?

Research Questions 3 (RQ3): What is the intervention effect of the BEST in CLASS–Web and BEST in CLASS in comparison with a control group?

Research Questions 4 (RQ4): Do teachers rate BEST in CLASS–Web and BEST in CLASS as acceptable?

Research Questions 5 (RQ5): Is BEST in CLASS–Web less costly than BEST in CLASS?

Method

Study Design

This preliminary study was designed as a small, multisite pre- and posttest cluster randomized controlled trial (Spybrook & Raudenbush, 2009). Teachers were randomly assigned to one of two conditions: BEST in CLASS–Web or BEST in CLASS. A third group of teachers were randomly selected from a pool of control teachers who participated in a previous efficacy study examining the BEST in CLASS delivered in person to a "business as usual" (BAU) control condition (see Conroy, Sutherland, & Feil, 2019).

Setting

Research activities occurred in federal and state early childhood programs serving young children in two Southeastern states. A total of 19 classrooms located within early childhood centers or local elementary schools were included in this study with 73.68% ($n = 14$) representing federally funded programs (i.e., Head Start) and 26.32% ($n = 5$) representing state-funded early childhood programs. On average, there were 16.21 ($SD = 3.75$) children in each classroom. Control teachers were randomly selected from a previous efficacy trial on BEST in CLASS (see Conroy, Sutherland, Algina, et al., 2019), which represented the same geographical location and similar programs and classroom size.

Table 1. Coach Demographics.

	Total	
Demographics	N = 6	%
Racial background		
African American/Black	0	0
Asian/Pacific Islander	1	16.7
White	5	83.3
Hispanic/Latino	0	0
Native American/American Indian	0	0
Other	0	0
Gender: Women	6	100
Age range		
18–25 years	2	33.3
26–35 years	1	16.7
36–45 years	1	16.7
46–55 years	1	16.7
Above 55 years	0	0
Prefer not to report	1	16.7
Highest level of education		
Bachelor's degree	1	16.7
Master's degree	4	66.7
Education specialist	1	16.7
Doctoral degree	0	0

Note. Percentages may total greater than 100 due to rounding.

Participants

Coaches. Six women coaches delivered the practice-based coaching to teachers (see Table 1 and supplemental materials for a description of coaching demographics). A minimum of a bachelor's degree was the only requirement for selection of coaches. The age of the coaches was reported to be between 18 and 55 years. A third of the coaches reported their age to be between 18 and 25 years. The age range for the remaining coaches varied, with one coach declining to report age. Three of these four coaches had previous teaching experience in a classroom; the other coach had no previous teaching experience, but recently graduated with a bachelor's degree in developmental psychology. The coaches who reported having previous teaching experience reported an average of 12.70 ($SD = 7.46$) years of experience in classrooms, with an average of 2.40 ($SD = 3.36$) years of teaching specifically in early childhood classrooms. Of the six coaches, the majority ($n = 4$; 66.7%) coached teachers in both conditions. Due to location constraints, one coach (20%) coached teachers in the BEST in CLASS–Web condition only and another coach (20%) coached teachers in the BEST in CLASS condition only. The majority of coaches self-identified as White (80%), with one coach (20%) self-identifying as Asian/Pacific. Four of the six (66.7%) coaches had previous coaching experience. In addition, four (66.7%) coaches were enrolled in a graduate degree program in early childhood or early childhood special education. Four coaches held a teaching license or certificate. All coaches were trained to fidelity prior to beginning coaching sessions with teachers.

Teachers. To recruit teacher participants for the study, research staff met with lead early childhood teachers to provide an overview of the study and collect informed consent. Consented teachers identified a maximum of five children in their classroom who demonstrated elevated

Table 2. Teacher Demographics.

Demographics	BEST in CLASS–Web (n = 10)		BEST in CLASS (n = 9)		Control (n = 10)	
	N	%	N	%	N	%
Gender						
Men	1	10.00	—	—	—	—
Women	9	90.00	9	100.00	10	100.00
Age range						
18–25 years	4	40.00	1	11.11	1	10.00
26–35 years	2	20.00	1	11.11	1	10.00
36–45 years	3	30.00	3	33.33	5	50.00
46–55 years	—	—	—	—	2	20.00
Above 55 years	—	—	4	44.44	1	10.00
Prefer not to say	1	10.00	—	—	—	—
Racial background						
African American	4	40.00	3	33.33	4	40.00
Asian	—	—	2	22.22	—	—
Hispanic	1	10.00	2	22.22	—	—
White	5	50.00	2	22.22	6	60.00
Highest level of education						
High school	—	—	—	—	2	20.00
Associate's	4	40.00	2	22.22	—	—
Bachelor's	2	20.00	4	44.44	5	50.00
Master's	3	30.00	2	22.22	3	30.00
Other	1	10.00	1	11.11	—	—
Years teaching			M (SD)			
Total	6.80	6.91	18.67	13.00	9.3	9.91
Preschool	4.20	4.54	11.89	11.50	5.40	6.59

rates of challenging behavior for potential participation in the study. Upon caregiver consent, the child was screened for eligibility in the study. Pretest data collection began once each participating teacher had a minimum of one child who was deemed eligible for participation. Child eligibility included (a) nominated by eligible classroom teacher, (b) aged between 3 and 5 years, (c) scored within the risk range on the Early Screening Project (ESP; Feil et al., 1995), (d) did not have an indication of a developmental delay based on results of the Battelle Developmental Inventory, Second Edition (BDI-2; Newberg, 2005), (e) spoke fluent English, and (f) caregivers provided consent to participate. Institutional review board (IRB) approval was obtained at each research site.

Twenty-five teachers were recruited and they consented for participation in this study (see Table 2). Prior to randomization, three teachers withdrew their consent. Therefore, a total of 10 teachers were assigned to the BEST in CLASS condition and 12 teachers to the BEST in CLASS–Web condition. Due to logistical reasons, one teacher did not complete the BEST in CLASS condition and two teachers did not complete the BEST in CLASS–Web condition, leaving a total sample of nine teachers in the BEST in CLASS condition and 10 teachers in the BEST in CLASS–Web condition, respectively (see Table 2 and supplemental materials). There were several reasons teachers did not complete the study, including scheduling conflict, personal time conflict, and time constraints due to the end of the school year. To compare the outcomes for teachers in the intervention conditions (i.e., BEST in CLASS and BEST in CLASS–Web) with

teachers who did not receive the BEST in CLASS intervention, a group of control (i.e., BAU) teachers who participated in a previous study examining the efficacy of BEST in CLASS were randomly selected to serve as the control group for this investigation. A total of 10 control teachers served in a control for this investigation. The characteristics of the control teachers had similar demographics and characteristics to this sample of teachers (see Table 2) and the data collection procedures in the previous study mirrored the procedures in this study (for a description of the efficacy study, see Conroy, Sutherland, Algina, et al., 2019). Minimal differential attrition (6.67%) for teacher participants occurred over the course of the study.

Fidelity Measures

Fidelity was assessed using measures across intervention components (i.e., training, coaching, and teacher implementation), dimensions of fidelity (i.e., dosage, adherence, and competence) and reporters (i.e., coach report, observation) to provide a comprehensive picture of fidelity. The measures used to assess fidelity in all conditions of the study are as follows.

Teacher training workshop and online module completion. To measure whether teachers completed the teacher training component of the intervention, we collected teacher attendance at the BEST in CLASS teacher training. In addition, data were collected on teacher completion of each of the online modules in the BEST in CLASS–Web condition.

Coaching dosage and fidelity

Dosage. Coaches in the BEST in CLASS and BEST in CLASS–Web conditions self-reported their coaching dosage after each coaching session. Data captured on the coaching dosage form included the following: date of meeting, duration of meeting, coaching supports provided, and data (i.e., video clips, graphical display, and anecdotal notes) reviewed with the teacher during the coaching meeting. Coaches in both conditions also tracked the type and frequency of follow-up contact with their teacher between meetings. The percentage of sessions where required components of the coaching protocol (e.g., direct observation) were completed was also recorded. Teachers who completed the BEST in CLASS intervention (through in-person or web-based coaching) were expected to receive approximately two coaching sessions per practice module.

Fidelity. Once coaching began, data were collected on the BEST in CLASS coaching implementation process using a direct observation coaching fidelity form. Coaches in both the in-person and web conditions recorded and uploaded videos of their coaching sessions onto a secure server for review by a trained observer. Trained coders completed fidelity checks on 25% ($n = 30$) of BEST in CLASS and 23% ($n = 29$) of BEST in CLASS–Web randomly selected videos.

The BEST in CLASS coaching fidelity form includes five domains: (a) reflection and feedback, (b) practice instruction/review, (c) shared goal setting and decision-making, (d) general items, and (e) quality collaboration. Each domain includes between two and six quality indicators, with scores ranging from 0 (*not observed*) to 4 (*exceeds standards*). Coaching fidelity observers were trained and practiced coding coaching videos until 80% interobserver agreement (IOA) was attained on each domain. Coaching fidelity data were collected for 30 in-person coaching meetings, with 27% being double coded for IOA and 92.96% of all domains scoring 2 (acceptable) or better. Coaching fidelity data were collected for 29 web-based coaching meetings, with 28% of them being double coded for IOA and 99.26 of all domains scoring 2 (acceptable) or better (see Table 3).

Teacher practice fidelity. The BEST in CLASS Adherence and Competence Scale (BICACS; Sutherland et al., 2014) is an observational tool used to assess the teachers' delivery of the BEST in CLASS practices in each condition, including the control condition. The BICACS assesses the

Table 3. Means and Standard Deviations of Indicators on Coaching Fidelity Form.

Domain	BEST in CLASS–Web	BEST in CLASS
	M (SD) *N* = 29	*M (SD)* *N* = 30
Reflection and feedback	3.32 (0.57)	3.20 (0.91)
Practice review	3.33 (0.73)	2.87 (1.00)
Goal/decision-making	3.03 (0.62)	2.91 (0.69)
General	3.33 (0.46)	3.24 (0.63)
Quality Collaboration	3.76 (0.37)	3.66 (0.44)
Overall *M*	3.37 (0.60)	3.19 (0.79)

Note. The range of scores on the BEST in CLASS coaching fidelity form is 0 (*not observed*) to 4 (*exceeds standards*). A score of 2.0 and above is considered acceptable. The "General" domain aims to address whether coaches included BEST in CLASS specific components during their coaching sessions with their teachers (e.g., accuracy in responses to teacher questions, use of BEST in CLASS language during the coaching session).

teachers' adherence to and competence of delivery of the BEST in CLASS practices. Adherence measures the thoroughness and frequency of delivery of the practices as intended, and competence measures the skillfulness (e.g., quality of use) of teacher delivery of the BEST in CLASS practices. The adherence and competence scales both comprise six items, each representing the BEST in CLASS practices.

In all three conditions (BEST in CLASS, BEST in CLASS–Web, and control) trained observers completed 10- to 15-min observations of teachers during an instructional activity in the classroom when the teachers were working with focal children. The BICACS adherence scale was measured on a 7-point Likert-type scale ranging from 1 (*not at all*) to 7 (*very extensive*). The competence scale was measured on a 7-point Likert-type scale ranging from 1 (*very poor*) to 7 (*excellent*). Higher rankings on the adherence scale indicate greater thoroughness in delivery of intervention practices and higher rankings in the competence scale indicate greater skillfulness in implementation of BEST in CLASS practices. Observers and coaches were provided with a BICACS manual and attended trainings that provided instructions on the procedures for administering and scoring the BICACS. In total, 99 teacher–child observations were completed across pre- and posttest. The mean intraclass correlation coefficients (ICCs) were .74 (adherence) and .60 (competence).

Teacher Outcome Measures

Classroom Assessment Scoring System (CLASS). The CLASS (Pianta et al., 2008) was used to assess classroom quality. CLASS scores were calculated from a total of four observation cycles (i.e., 10- to 20- min per observation). The quality of the classroom was assessed along 10 dimensions using a scale from 1 to 7 (i.e., 7 indicated highest level quality). Scores in each dimension were averaged to calculate the mean composite scores across three domains: Emotional Support, Classroom Organization, and Instructional Support. Certified observers conducted CLASS observations at pre- and posttest. Due to the relatively small sample size for this study, the internal consistency calculations were derived from a previous BEST in CLASS efficacy study (see Conroy, Sutherland, Algina, et al., 2019). The internal consistency was acceptable, with Cronbach's alpha equal to .88 for Emotional Support, .89 for Classroom Organization, and .85 for Instructional Support.

Teachers' Sense of Efficacy Scale (TSES). The TSES (Tschannen-Moran & Hoy, 2001) is a 24-item Likert-type type question survey given at both pre- and posttest. The TSES includes three

subscales that capture a teacher's perspectives on Student Engagement, Instructional Strategies, and Classroom Management. Questions are scaled, ranging from 1 (*nothing*) to 9 (*a great deal*). The internal consistency calculations derived from the previous BEST in CLASS efficacy study for this measure was acceptable, with Cronbach's alpha equal to .86 for Student Engagement, .90 for Instructional Strategies, and .89 for Classroom Management.

Teacher–Child Interaction Direct Observation System (TCIDOS). The TCIDOS−Research, Version 3.1 (TCIDOS-RV3.1) is a researcher-developed partial interval-based recording system used to measure teacher implementation of the BEST in CLASS practices. Trained coders conduct 15-min observations, focusing on teacher use of practices with a focal child in the classroom alternating between a 10-s observation interval and a 5-s recording interval. Using Lily software (Tapp, 2010), coders record the teacher's use of practices during each observation interval. Observations were conducted during authentic large and small group activities in the classroom and observers coded teacher behaviors occurring with each focal child, separately, within those activities. Coders were trained to a gold standard, ensuring IOA to a minimum of 80% on each code. Training sessions on the TCIDOS-RV3.1 included (a) didactic training on detection of teachers' delivery of BEST in CLASS practices, (b) TCIDOS-RV3.1 manual that included the coding system and operational definitions codes, and (c) video exemplars of each practice illustrating a continuum of adherence and competence in using the practice. Coders were also trained on how to use the Lily software (Tapp, 2010). All coders checked out prior to pretest data collection and recalibration occurred prior to posttest.

Teacher Intervention Acceptability

The BEST in CLASS Intervention Acceptability Scale uses 12 items to obtain teachers' perspectives of the acceptability of the intervention, using a 5-point Likert-type scale ranging from 1 (*not at all*) to 5 (*very much*). At the end of the intervention, teachers were asked to complete the rating scale and data were collected examining how acceptable and useful the BEST in CLASS intervention was for teachers who received in-person or web-based delivery.

BEST in CLASS, BEST in CLASS–Web, and Control Conditions

BEST in CLASS. The BEST in CLASS condition in this study followed the same protocol established in previous BEST in CLASS research (for a description, see Conroy, Sutherland, Algina, et al., 2019). BEST in CLASS is a Tier 2 intervention designed to support teachers' use of evidence-based practices with young children in their classrooms who demonstrate challenging behaviors. BEST in CLASS consists of three manualized components: (a) a 6-hr in-person teacher training workshop, (b) a teacher resource manual containing modules on the six practices (i.e., rules, precorrection, opportunities to respond, behavior specific praise, corrective feedback, and instructive feedback, plus linking and mastery), and (c) 14 weeks of practice-based coaching (adapted from Snyder et al., 2015) during which coaches support teachers' use of each practice with fidelity and competence. Project staff, including the coaches, delivered the in-person training workshop. During the teacher workshop, teachers were provided with a BEST in CLASS teacher resource manual. The teacher training workshop included exemplar videos, interactive activities, and knowledge checks. Following the teacher training workshop, coaches began meeting weekly with their teachers to implement the practice-based coaching component. The practice-based coaching component follows the model developed by Snyder et al. (2015, 2022) and begins with coaches and teachers creating shared goals and action plans for implementation of the BEST in CLASS practices. This meeting is followed by the coach conducting a focused observation of the teachers' implementation of the BEST in CLASS practice with focal children

in the classroom. This step is followed by an in-person coaching meeting where the coach and the teacher reflect on the implementation of the practice and the coach provides performance feedback to the teacher on implementation and child behavior. This practice-based coaching cycle repeats each week for 14 weeks.

BEST in CLASS–Web. BEST in CLASS–Web was adapted from the original BEST in CLASS and delivered through two online platforms, one to provide remote teacher training, and the other to implement the practice-based coaching component. The teacher training was developed by Technology-based Behavioral Interventions Delivery System (T-BIDS; see Feil et al., 2020) and occurred through a series of interactive modules that were accessible to teachers at any time following pretest data collection. The training included the same content as the in-person training, but was delivered through the web through the asynchronous modules. Online training modules included an introduction to BEST in CLASS and separate modules describing each BEST in CLASS practice. These practice-specific modules mirrored the content from the in-person training and contained introductory information about the practice, video exemplars, and interactive knowledge checks. A final module for linking and mastering the BEST in CLASS practices together followed these practice-specific modules. One difference between the in-person delivery and online delivery of practice-based coaching was rather than the coach observing the teachers' use of the practices in-person, the teachers used an iPad® and the TORSH Talent application (v.1.0.11) to video-record themselves during the instructional activity identified on the action plan as a time for implementing the practices with the focal children in the classroom. Once the recording period was complete, the video automatically uploaded to the www.TORSHTalent. com website. The coach then viewed the video, gathered data on the teacher's use of practices as planned, and embedded comments with feedback into the video using the TORSH Talent platform. During the subsequent coaching meeting which occurred over Zoom©, coaches and teachers reflected on the use of the practices and the coaches provided teachers performance feedback. In addition to the online training modules and the delivery of practice-based coaching online, teachers had access to resources on TORSH Talent, such as the BEST in CLASS teacher resource manual, exemplar videos, and forms used during coaching.

In addition to the online modules and practice-based coaching, the BEST in CLASS–Web teachers were trained by the coaches to use BEST in CLASS technology prior to beginning the intervention. Coaches visited the teacher's classroom to help set up the iPad® and train teachers in the use of the technology (e.g., iPad® use to video-record). Teachers received a hard copy of the BEST in CLASS teacher resource manual, a BEST in CLASS teacher technology user manual, an iPad®, instructions on how to use the technology (e.g., iPad®, TORSH Talent application), and hardware (e.g., tripod) for video-recording themselves in their classroom. Information was also shared on how to use the online platforms for coaching meetings. Prior to the beginning of coaching, coaches initiated virtual introductory meetings with their teachers over Zoom©. The meeting served as an opportunity for coaches and teachers to problem-solve any technology issues. These meetings were also used to familiarize the coach with the teacher's classroom and the focal children. After the introductory meeting, practice-based coaching sessions began between the teacher and the coach. The video-exemplar library was used during coaching meetings and was available for teachers to access independently as well.

Control Condition

Teachers in the control condition did not receive any professional development or coaching other than what their school or BEST in CLASS delivered to all teachers in their BEST in CLASS (e.g., training in a new curriculum adopted by the BEST in CLASS). They did not participate in any BEST in CLASS–Web or BEST in CLASS trainings, coaching, or receive any resource materials.

Data Analysis

To address the research questions, contrasts between matched group means were conducted on the basis of repeated measures analysis of variance (ANOVA) methods (i.e., for pre- and posttest score comparisons between repeated assessments on the same participants), whereas contrasts between independent group means were conducted on the basis of completely randomized ANOVA methods (i.e., for mean comparisons between assessments on participants from different groups at pretest and posttest). An examination of the assumptions for independence, normality (based on Lilliefors (1967), method of the Kolmogorov–Smirnov test), homogeneity of variance (based on Hartley (1940), *F-max* test), and sphericity (based on the Mauchley (1940) and Geisser and Greenhouse (1958) tests), all indicated that the data met the assumptions for using the ANOVA methods.

To statistically test the extent to which BEST in CLASS–Web was effective in improving teacher outcomes, we calculated effect sizes (Cohen's *d*) and statistical significance of differences (*p* values) between observed means at pre- and posttest. To test whether the treatment effect of BEST in CLASS–Web was equivalent to that of BEST in CLASS, we calculated effect sizes (Cohen's *d*) and statistical significance of differences (*p* values) between observed means at posttest for BEST in CLASS–Web, BEST in CLASS, and the control condition. In light of the ongoing debate concerning the use of significance testing versus effect size for the accurate determination of treatment effects, we elected to report both Cohen's *d* and statistical probability (*p*) values but focus on interpreting the obtained effect size measures (Patriota, 2017). We elected this approach because, in general, statistical probability (*p*) values may not be useful for evaluating the implications of the findings of a study, particularly when multiple comparisons are being conducted and the sample sizes are small (Thompson, 2002) as is the case in this study. In contrast, Cohen's *d* effect size can be used to examine the magnitude of differences between two (or more) groups on a given variable and it is not affected by the number of comparisons made or affected by dependent or independent comparison groups (Cohen, 1988; Nakagawa & Cuthill, 2007). As a general measure of treatment effects, Cohen's *d* effect size is particularly informative relative to the practical importance and magnitude of observed differences; it is also immensely informative when contrasting differences between small sample–sized groups (Nakagawa & Cuthill, 2007). Common standards for assessing effect size using Cohen's *d* are 0.20 for small, 0.50 for moderate, and 0.80 for large (Cohen, 1988).

Cost Comparison

A descriptive comparative analysis was conducted to examine the costs associated with the delivery of BEST in CLASS professional development components, including practice-based coaching, on the web in comparison with delivery on-site. Average costs were calculated for materials, coaches' time, and associated travel costs across research sites.

Results

Baseline Equivalence

To determine baseline similarity between BEST in CLASS–Web, BEST in CLASS, and the control group, mean differences on all study variables at pretest were tested using one-way ANOVAs. Several differences were found. In regard to teaching practices, results revealed a significant group effect on teachers' use of behavior-specific praise, $F(2, 50) = 6.63, p < .05$; Tukey's HSD (honestly significant difference) post hoc comparison tests revealed that the percentage of intervals mean score for the BEST in CLASS group was significantly higher than the BEST in CLASS–Web and control groups. Results also revealed a significant group effect on teacher

adherence to behavior-specific praise, $F(2, 50) = 6.76$, $p < .05$; Tukey's HSD post hoc comparison tests revealed that the mean score for the BEST in CLASS group was significantly higher than the BEST in CLASS–Web and control groups. Similarly, results revealed a significant group effect on teacher competence in delivery of behavior-specific praise, $F(2, 11) = 4.57$, $p < .05$; Tukey's HSD post hoc comparison tests revealed that the mean score for the BEST in CLASS group was significantly higher than the control group. In addition, results revealed a significant group effect on teacher competence in delivery of precorrection, $F(1, 5) = 13.06$, $p < .05$; Tukey's HSD post hoc comparison tests revealed that the mean score for the BEST in CLASS group was significantly higher than the control group. In regards to teacher self-efficacy, results revealed a significant group effect on teacher self-efficacy for student engagement, $F(2, 26) = 3.59$, $p < .05$; Tukey's HSD post hoc comparison tests revealed that the mean score for the BEST in CLASS–Web was significantly higher than the BEST in CLASS group. However, no significant baseline group differences were found for the remaining study variables, which included classroom quality (i.e., CLASS), the remaining two subscales for teacher self-efficacy, and teacher use (i.e., adherence and competence) of delivery of the remaining practices.

Treatment Fidelity

Teacher training workshop and online module completion. All teachers assigned to the BEST in CLASS condition completed the in-person teacher training. Teachers assigned to the BEST in CLASS–Web condition were provided access to the online training modules. Of the 10 teachers in the BEST in CLASS–Web condition, 4 teachers completed all of the online modules and 6 teachers completed, on average, 66% ($SD = 15.98$) of the modules, but did not complete all the modules.

Coaching dosage and fidelity

   BEST in CLASS–Web coaching dosage and fidelity. Eight of the 10 teachers (80%) received coaching on all seven modules over 13 to 14 weeks, with an average of 13.5 web-based coaching sessions per teacher. Two of the 10 teachers (20%) did not receive coaching on all of the modules due to scheduling constraints and end of the school year. Coaches' direct observation of teachers' use of the practices took place using teacher video-recordings, with observational data collected in 96.64% of those observations and anecdotal data (i.e., notes made by coach about quality examples of implementation or missed opportunities) recorded for 100% of video observations. On average, BEST in CLASS–Web coaching sessions with teachers lasted approximately 31.41 min (range = 8–65 min). After coaching sessions that occurred in the BEST in CLASS–Web condition, teachers received an average of 4.09 (range = 0–11) follow-up contacts from their coach. As part of the coaching process, coaches in both conditions follow up the coaching session by sending the teacher a note with the goal, action plan, and performance feedback discussed during the coaching meeting. In addition, coaches in the online condition followed up with additional reminders for teachers to conduct their video-recording and upload their videos, and a reminder about the upcoming coaching meeting. On average, when coaches implemented coaching sessions over the web, fidelity data indicate they implemented the coaching protocol with a mean fidelity of 3.37 ($SD = 0.60$), which exceeded the minimum criteria of 2.0 and above on a 4-point scale.

   BEST in CLASS coaching dosage and fidelity. Eight of the nine teachers (89%) received coaching on all seven modules over the 13 to 14 weeks of coaching, with an average of 13.9 in-person coaching sessions per teacher. One teacher did not receive the full coaching dosage due to scheduling constraints and the end of the school year. For the BEST in CLASS condition, teachers received modeling and prompting from the coach in 18% of coaching sessions. Direct observations of teachers occurred during 100% of coaching visits to the classroom, with video data collected in 99% of those observations and anecdotal data recorded for 100% of visits. On

Table 4. BEST in CLASS Adherence and Competence Scales.

	BEST in CLASS–Web			BEST in CLASS			Control		
BICACS	M	SD	ES	M	SD	ES	M	SD	ES
Adherence									
Pretest	2.38	1.02	1.02	2.41	1.17	1.24	2.42	1.26	−0.13
Posttest	3.22	1.51		3.60	1.72		2.29	1.60	
Competence									
Pretest	3.93	1.35	0.40	4.31	1.14	0. 66	3.39	1.24	0.64
Posttest	4.53	1.28		5.16	0.96		4.52	1.32	

Note. ES = effect size. BICACS = BEST in CLASS Adherence and Competence Scale.

average, in-person coaching meetings with teachers lasted 30.52 min (range = 14–60 min). After coaching meetings, teachers received an average of 2.66 (range = 0–11) follow-up contacts from their coach. As part of the coaching process in person, coaches follow up the coaching session by sending the teacher a note with the goal, action plans, and performance feedback attached. They also send a reminder about the observation and subsequent coaching meeting. On average, when coaches implemented coaching sessions in person, fidelity data indicate they implemented the coaching protocol with a mean fidelity of 3.19 ($SD = 0.79$), which exceeded the minimum criteria of 2.0 and above on a 4-point scale.

Teacher practice fidelity. Teachers in the BEST in CLASS and BEST in CLASS–Web conditions increased their adherence and competence from pretest to posttest (see Table 4). The effect sizes for adherence were 1.24 and 1.02 for BEST in CLASS and BEST in CLASS–Web, respectively. The effect sizes for competence were 0.66 and 0.40 for BEST in CLASS and BEST in CLASS–Web, respectively. Teachers in the control condition had a decrease in adherence from pretest to posttest (Cohen's $d = -0.13$) and an increase in competence from pretest to posttest (Cohen's $d = 0.64$).

Teacher Outcomes

Classroom quality. Results revealed a substantial change for the BEST in CLASS–Web intervention between pre- and posttest teacher scores on the Emotional Support and Classroom Organization domains of the CLASS (Pianta et al., 2008; Cohen's $d = 1.63$, $p < .05$ and Cohen's $d = 1.09$, $p = .028$, respectively; see Table 5). In addition, the posttest means for each of the domain scores under the BEST in CLASS–Web condition were, on average, larger than those for BEST in CLASS (Emotional Support: Cohen's $d = 1.60$, $p < .05$; Classroom Organization: Cohen's $d = 1.11$, $p < .05$; and Instructional Support: Cohen's $d = 1.09$, $p < .05$) and the control condition (Emotional Support: Cohen's $d = 2.01$, $p < .05$; Classroom Organization: Cohen's $d = 1.13$, $p < .05$; and Instructional Support: Cohen's $d = 0.99$, $p < .05$). In contrast, the posttest means for Classroom Organization, Emotional Support, and Instructional Support, when comparing the BEST in CLASS condition with the control condition, were negligible (see Table 6).

Teacher self-efficacy. Inspection of the means on the TSES (Tschannen-Moran & Hoy, 2001) indicates small to large changes for BEST in CLASS–Web between pre- and posttest teacher scores on the Student Engagement (Cohen's $d = 0.15$, $p = .50$), Instructional Strategies (Cohen's $d = 0.68$, $p = .022$), and Classroom Management (Cohen's $d = 0.57$, $p = .18$) domains (see Table 5). In addition, inspection of Cohen's d effect sizes comparing conditions shows that the posttest

Table 5. Means and Effect Sizes for Study Variables.

	BEST in CLASS–Web			BEST in CLASS			Control		
	Pretest M	Posttest M	Cohen's d	Pretest M	Posttest M	Cohen's d	Pretest M	Posttest M	Cohen's d
CLASS									
Emotional support	5.34	5.88	1.63*	5.30	5.35	0.07	4.83	5.22	0.40
Classroom organization	4.64	5.27	1.09*	4.82	4.64	−0.17	4.57	4.63	0.06
Instructional support	2.86	2.92	0.06	2.88	2.16	−1.06	2.11	2.23	0.14
TSES									
Student engagement	7.39	7.54	0.15	5.85	7.67	1.36*	6.95	6.98	0.03
Instructional strategies	7.31	7.89	0.68*	6.08	7.89	1.32*	7.09	7.18	0.10
Classroom management	6.76	7.41	0.57	5.76	7.75	1.49*	6.64	6.79	0.18
TCIDOS—RV3.1									
Rules	0.01	0.21	0.78*	0.02	0.18	0.66	0.01	0.01	−0.11
Behavior-specific praise	0.00	0.19	0.72*	0.01	0.16	0.65	0.00	0.00	0.32
Precorrection	0.00	0.17	0.64	0.00	0.14	0.56	0.00	0.00	−0.08
Instructive feedback	0.00	0.17	0.63	0.00	0.14	0.55	0.01	0.01	0.12
Corrective feedback	0.02	0.18	0.58	0.01	0.14	0.55	0.02	0.02	0.15
Opportunity to respond	0.46	0.53	0.31	0.52	0.60	0.38	0.40	0.38	−0.15

Note. Standard deviations are not reported as they can be readily computed using the mean, sample size, and resulting effect size measure. CLASS = Classroom Assessment Scoring System; TSES = Teachers' Sense of Efficacy Scale; TCIDOS-RV3.1 = Teacher–Child Interaction Direct Observation System–Research Version 3.1.
*$p < .05$.

means scores for the Student Engagement, Instructional Strategies, and Classroom Management domains under the BEST in CLASS–Web condition versus the BEST in CLASS condition were not different. However, scores for teachers exposed to BEST in CLASS–Web were higher than those in the control condition (Student Engagement: Cohen's $d = 0.80$, $p < .05$; Instructional Strategies: Cohen's $d = 1.02$, $p < .05$; and Classroom Management: Cohen's $d = 0.89$, $p < .05$). Scores for teachers exposed to BEST in CLASS were also higher than those in the control condition (Student Engagement: Cohen's $d = 1.0$, $p < .05$; Instructional Strategies: Cohen's $d = 1.02$, $p < .05$; and Classroom Management: Cohen's $d = 1.07$, $p < .05$; see Table 6).

Teachers' implementation of practices. Teachers' use of rules, behavior-specific praise, precorrection, instructive feedback, corrective feedback, and opportunities to respond were assessed with the TCIDOS-RV3.1. Inspection of the means indicates that, for the BEST in CLASS–Web, there was substantial change from pre- to posttest for the use of rules (Cohen's $d = 0.78$, $p = .031$) and behavior-specific praise (Cohen's $d = 0.72$, $p = .045$). For instructive feedback, corrective feedback, and precorrection, the changes from pre- to posttest were moderate, while for opportunity to respond, the changes were small (see Table 5). When comparing teachers across the BEST in CLASS–Web condition with those in the BEST in CLASS and the control conditions, results show small but not significant differences on the posttest means compared with those for the BEST in CLASS (see Table 6). However, BEST in CLASS–Web showed significant change in use of rules (Cohen's $d = 0.97$, $p < .05$), behavior-specific praise (Cohen's $d = 0.82$, $p < .05$), precorrection (Cohen's $d = 0.74$, $p < .05$), instructive feedback (Cohen's $d = 0.70$, $p < .05$), corrective feedback (Cohen's $d = 0.77$, $p < .05$), and opportunities to respond (Cohen's $d = 0.77$, $p < .05$) when compared with the control condition. Similarly, BEST in CLASS showed significant change in use of rules (Cohen's $d = 0.72$, $p < .05$) and opportunities to respond (Cohen's $d = 0.98$, $p < .05$) when compared with the control condition (see Table 6). The IOA

Table 6. Summary of Effects on Posttest Scores Across Treatment Conditions.

Instrument	BEST in CLASS–Web vs. BEST in CLASS	BEST in CLASS–Web vs. control	BEST in CLASS vs. control
	Cohen's *d*	Cohen's *d*	Cohen's *d*
CLASS			
Emotional support	1.60*	2.00*	0.19
Classroom organization	1.11*	1.13*	0.01
Instructional support	1.09*	0.99*	−0.10
TSES			
Student engagement	−0.14	0.80*	1.0*
Instructional strategies	0.00	1.02*	1.02*
Classroom management	−0.38	0.89*	1.07*
TCIDOS-RV3.1			
Rules	0.15	0.87*	0.72*
Behavior-specific praise	0.13	0.82*	0.69
Precorrection	0.15	0.74*	0.59
Instructive feedback	0.16	0.70*	0.54
Corrective feedback	0.14	0.77*	0.53
Opportunity to respond	−0.29	0.77*	0.98*

Note. CLASS = Classroom Assessment Scoring System; TSES = Teachers' Sense of Efficacy Scale; TCIDOS-RV3.1 = Teacher–Child Interaction Direct Observation System−Research Version 3.1.
*$p < .05$.

data were collected in 30.49% of observation sessions. Mean IOA for rules was 0.98 (range = 0.93–1.0), for precorrection it was 0.99 (range = 0.97–1.0), for opportunities to respond it was 0.86 (range = 0.75–0.94), for behavior-specific praise it was 0.98 (range = 0.93–1.0), for corrective feedback it was .99 (range = 0.93–1.0), and for instructive feedback it was 0.99 (range = 0.95–1.0).

Intervention acceptability. All teachers receiving BEST in CLASS rated their overall experience as *very good*. On average, teachers rated BEST in CLASS as easy to implement, improved their classroom atmosphere, and was not disruptive to their classrooms. Overall, teachers agreed that the BEST in CLASS intervention included familiar strategies that they learned to implement more effectively and reported that they would continue to use the BEST in CLASS practices after the study. There was a slight difference in the overall ratings between those who received BEST in CLASS and BEST in CLASS–Web. BEST in CLASS–Web teachers indicated that the intervention was slightly more time-intensive than those who received BEST in CLASS (see Table 7).

Cost Comparison

When examining the total costs associated with delivering BEST in CLASS on-site, in comparison with over the web, outcomes indicate that BEST in CLASS−Web is costlier than BEST in CLASS delivered in person (see Table 8). The increase in costs of the web-based delivery is primarily related to the additional support needed for technology equipment and technology assistance as needed. The cost for coaches' time to deliver BEST in CLASS was relatively the same. Not surprisingly, the costs related to coach travel was greater when BEST in CLASS were delivered in person in comparison with delivered over the web.

Table 7. Teacher Intervention Acceptability.

Survey questions	BEST in CLASS–web (N = 10) M (SD)	BEST in CLASS (N = 8) M (SD)	Cohen's d
1. How comfortable were you with implementing?	4.50 (0.53)	4.63 (0.74)	0.20
2. How time-intensive was it for you to implement?	3.40 (0.84)	2.63 (0.74)	0.97*
3. How difficult was it for you to implement?	1.80 (0.92)	1.38 (0.52)	0.56
4. How disruptive was it to your classroom?	1.90 (0.88)	1.50 (0.76)	0.48
5. How comfortable were you with the amount of training?	4.60 (0.52)	4.50 (0.76)	0.15
6. How useful to improving classroom atmosphere?	4.70 (0.48)	5.00 (0.00)	0.88
7. How useful to improving behavior?	4.50 (0.71)	4.75 (0.71)	0.35
8. To what extent are there disadvantages?	4.30 (0.67)	4.38 (0.92)	0.09
9. Does the BEST in CLASS intervention include new or different strategies that you have not used previously?	3.20 (1.23)	4.13 (0.99)	0.88*
10. Does the BEST in CLASS intervention include new or different strategies that you have learned how to implement better or more effectively?	4.40 (0.84)	4.75 (0.46)	0.51
11. Will you continue to use the BEST in CLASS intervention?	4.90 (0.32)	4.88 (0.35)	0.06
12. Overall, how would you rate the BEST in CLASS intervention?	4.70 (0.48)	5.00 (0.00)	0.88

Note. Items 2, 3, and 4 were reverse scaled, where lower scores indicate greater intervention acceptability.
*$p < .05$.

Table 8. Cost Comparison of BEST in CLASS Versus BEST in CLASS–Web.

Category	BEST in CLASS–Web	BEST in CLASS
Materials		
Implementation materials including technology devices and resources	US$12,904	US$2,140
Time		
Coach training and coaching meetings	US$7,340	US$7,976
Travel		
Travel to BEST in CLASS sites	US$99	US$745
Total	US$20,343	US$10,861

Discussion

The aims of this preliminary study were to (a) examine the equivalence with regard to coaching fidelity, coaching dosage, and teacher implementation of intervention practices of BEST in CLASS–Web in comparison with in-person delivery, and (b) examine teacher outcomes of BEST in CLASS–Web in comparison with BEST in CLASS and a control condition. Finally, we examined the acceptability of the intervention and cost of the intervention. In the following, we discuss these findings, limitations of this study, and implications for future research.

Fidelity of coaching implementation and coaching dosage revealed that, in general, the delivery of the practice-based coaching component was implemented with fidelity across both intervention conditions (i.e., web and in person). In addition, teachers in both conditions received

similar amounts of coaching across the intervention; however, there were differences between conditions in the completion of the teacher training. Teachers who received the teacher training in person completed the training in a single day, whereas the teachers who had access to the teacher training using asynchronous web-based modules only completed, on average, two thirds of the teacher training. The differences in teachers completing the teacher training may be a result of how the training was delivered across conditions. The teachers who received the training in person attended a full-day workshop at a school location. The teachers who received the training online completed the training modules on their own throughout the 14 weeks of coaching. Regardless of whether teachers received the full teacher training, observational data from the BICACS suggest that teachers in both the BEST in CLASS–Web and BEST in CLASS conditions delivered the practices more extensively (adherence) and with greater quality (competence). The BICACS data also suggest that the coaching supports that teachers receive may contribute to the quality of delivery of BEST in CLASS practices in both conditions. As suggested in the professional development literature, trainings that target increased knowledge of an intervention, such as the workshop and modules in this study, are not as likely to result in changes in teacher's practice in a classroom setting (Joyce & Showers, 2002). At the same time, the coaching fidelity data may provide further insight into differences in classroom quality data favoring the web condition. That is, although coaching fidelity scores were high for both conditions, slight differences were noted favoring the BEST in CLASS–Web condition. For example, four of the five coaching fidelity indicators were scored at 100% for the BEST in CLASS–Web condition, whereas only one indicator on the coaching fidelity checklist was scored at 100% in the BEST in CLASS condition. Similarly, coach-reported fidelity was slightly higher across each of the self-report fidelity indicators. These data, in sum, suggest that coaching fidelity was slightly higher in the web condition, which may help explain findings that favor the BEST in CLASS–Web condition outcomes. To illustrate, the increased number of follow-up contacts in the web condition might explain some of the findings favoring the web condition in comparison with the in-person condition as teachers who received additional contacts from their coach may have felt more support and alliance with their coach and scored higher in some dimensions of self-efficacy and classroom quality (e.g., emotional support). Although this investigation examined coaching fidelity, including dosage and quality, we did not control for the number of follow-up contacts after coaching and thus are not able to determine the impact. Future research should continue to examine the possible differences between online coaching and in-person coaching, and how delivery of coaching might impact teachers' implementation of practices and self-efficacy.

In regard to teacher outcomes, results were mixed, with improvement on several dimensions of classroom quality favoring the BEST in CLASS–Web condition, whereas the increase in teacher's sense of self-efficacy from pretest to posttest was greater in the in-person condition in comparison with web delivery or the control condition. The BEST in CLASS–Web condition had a large effect on classroom quality, particularly on the dimensions of Emotional Support and Instructional Support (CLASS; Pianta et al., 2008). This effect was present both within condition and in comparison with the BEST in CLASS and control conditions. This is surprising, particularly given that BEST in CLASS condition has shown moderate effects on dimensions of the CLASS in a previous study (Conroy, Sutherland, Algina, et al., 2019). One potential explanation for this finding may be related to coaching fidelity; although teachers in the BEST in CLASS–Web condition did not complete as much training as teachers in the BEST in CLASS condition, coaching fidelity was higher in the BEST in CLASS–Web condition, and teachers in this condition received more follow-up contacts from coaches. As discussed above, teachers who receive coaching have a greater likelihood of using desired practices in their classrooms (Joyce & Showers, 2002); in this case, the quality and amount of contact with coaches may have resulted in larger improvements in classroom quality in the BEST in CLASS–Web condition. While a larger sample size may further inform our ability to determine an effect of the BEST in CLASS

condition on classroom quality, the large effects noted in the BEST in CLASS–Web condition are promising and speak to a robust impact on classroom quality. As noted in Conroy, Sutherland, and Feil (2019), this is an important finding, given the Tier 2 nature of BEST in CLASS and its focus on improving interactions between teachers and focal children. That is, it appears that Tier 2 interventions may have broader effects on the larger classroom ecology.

Relatedly, moderate to large effects were noted for both the BEST in CLASS–Web and BEST in CLASS conditions for all subscales of the TSES, suggesting that teachers who received training and coaching in BEST in CLASS felt more efficacious. There were also large effects found for both of these conditions compared with the control condition. These findings are particularly interesting in light of the direct observation data of teachers' delivery of BEST in CLASS practices, which suggest that teachers in both conditions used the practices at greater rates from pre- to posttest and when compared with the control condition. In sum, some combination of training and coaching in BEST in CLASS practices appears to result in improvements in both teachers' perceptions of their efficacy and the practice delivery, and in this case the BEST in CLASS and BEST in CLASS–Web teachers' use of practices may provide some confirmation of their increased feelings of self-efficacy.

Given the strengths of web-based professional development approaches (e.g., Castle & McGuire, 2010; Dede et al., 2009; Knight, 2012), findings from this study for BEST in CLASS–Web are promising. Furthermore, teachers rated both the web and in-person interventions as useful and acceptable, with no notable differences between conditions. Although costs for the BEST in CLASS–Web condition were higher than the BEST in CLASS condition, it is important to note that these costs are primarily related to start-up and that, over time, BEST in CLASS–Web would likely decline after start-up. Most likely, BEST in CLASS–Web would not only be more cost-effective, but also have the ability to reach more teachers and children.

Limitations and Future Research

Although findings are promising for BEST in CLASS–Web, results should be viewed with several limitations in mind. First, this was a small, underpowered randomized controlled trial that was part of an Institute of Education Sciences–funded development project; therefore, our ability to test for statistical differences that may have existed between groups is hindered. In addition, although the design (randomly selecting control classrooms from a previous BEST in CLASS investigation) is creative, these remain a small number of classrooms and may not be representative of the larger number of business-as-usual classrooms in early childhood settings. This is highlighted by the competence data for the control condition in this study, which increases over time. While this increase is contrary to the BAU condition (Conroy, Sutherland, Algina, et al., 2019) from which this sample was selected, it should also be noted that, in this sample, teacher adherence in the control condition decreased from pretest to posttest; although teachers may have been delivering practices with more competence, they were delivering practices over time with less extensiveness, thereby decreasing the overall dosage received by focal children.

Another important limitation to note are several pretest differences in the delivery of the practices and teacher self-efficacy across intervention groups in this sample of teachers. To illustrate, across the three groups, teachers differed in their use and delivery of behavior-specific praise and in their delivery of precorrection. When interpreting study findings for each of these practices, these baseline differences should be considered. One final and important limitation is that observers were not blind to condition. Precautions were taken to minimize observer bias, including varying observers' observations across teachers in both conditions, training observers on observer bias, conducting a recalibration training in the middle of data collection, and conducting interobserver and reliability checks across all phases of the study.

Although the findings are promising, this is a preliminary investigation; therefore, interpretation of these differences is limited and we caution readers against overgeneralizing these

findings. Future work should look to replicate study findings across equivalent BEST in CLASS, BEST in CLASS–Web, and control groups. Findings from this study do provide several avenues for future research. First, this study should be replicated to better examine the effects of web-based professional development models on a variety of teacher outcomes. In addition, a replication study should evaluate the effects of delivering the training in the web condition synchronously through an online platform or providing structure so that all participants in the web condition are fully trained in the practices. Second, while the coaching fidelity (both coach report and observed) was high across both conditions in this study, future work should examine how different aspects of coaching (e.g., alliance) and coach characteristics (e.g., age, years of coaching experience) may be related with treatment outcomes for both web-based and in-person delivery options. For example, coaches in this investigation varied across a number of characteristics, including previous experience as a teacher, coach, and age. Although the data indicated these varying characteristics did not impact coaching fidelity, further research is warranted to examine coaches' attributes, coaching implementation, and teacher satisfaction. Finally, although the acceptability data suggest that teachers found the interventions useful and acceptable, future work is needed to examine the sustainment of both the intervention model (e.g., Do early childhood programs continue to implement web-based models in subsequent years?) and teacher delivery of practices (e.g., Do teachers continue to use practices with new focal children in subsequent years without coaching support?).

Conclusion

Many early childhood teachers struggle to support children who exhibit chronic challenging behavior in their classrooms. Given that many teachers and early childhood programs lack access to high-quality professional development and supports, web-based models are a promising approach to help bridge this gap. This study examined one such model, BEST in CLASS–Web, and results suggest that this is a promising approach for providing high-quality professional development to teachers of young children with chronic challenging behavior. Those delivering professional development to early childhood practitioners need to continue to leverage the opportunities that technological advances provide to best meet the needs of the teachers for children in need of socio-emotional learning and supports.

Authors' Note

Katerina M. Marcoulides is now in the Quantitative/Psychometric Methods Program, Department of Psychology at the University of Minnesota.

Declaration of Conflicting Interests

Funding

ORCID iDs

Maureen A. Conroy https://orcid.org/0000-0002-0170-3299

Katerina M. Marcoulides https://orcid.org/0000-0001-8829-870X

Supplemental Material

Supplemental material for this article is available online.

References

Artman-Meeker, K. M., & Hemmeter, M. L. (2012). Effects of training and feedback on teachers' use of classroom preventive practices. *Topics in Early Childhood Special Education*, *33*(2), 112–123. https://doi.org/10.1177/0271121412447115

Artman-Meeker, K. M., Hemmeter, M. L., & Snyder, P. (2014). Effects of distance coaching on teachers' use of Pyramid model practices: A pilot study. *Infants & Young Children*, *27*, 325–344. https://doi.org/10.1097/IYC.0000000000000016

Becker, K. D., & Domitrovich, C. E. (2011). The conceptualization, integration, and support of evidence-based interventions in the schools. *School Psychology Review*, *40*(4), 582–589. https://doi.org/10.1080/02796015.2011.12087531

Blair, C., & Raver, C. C. (2012). Child development in the context of adversity: Experiential canalization of brain and behavior. *American Psychologist*, *67*, 309–318. https://doi.org/10.1037/a0027493

Bruder, M. B., Mongro-Wilson, C., Stayton, V. D., & Dietrich, S. L. (2009). The national status of in-service professional development systems for early intervention and early childhood special education practitioners. *Infants and Young Children*, *22*, 13–20. https://doi.org/10.1097/01.IYC.0000343333.49775.f8

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. https://doi.org/10.1016/j.jsp.2012.08.001

Castle, S. R., & McGuire, C. J. (2010). An analysis of student self-assessment of online, blended, and face-to-face learning environments: Implications for sustainable education delivery. *International Education Studies*, *3*(3), 36–40.

Cloney, D., Cleveland, G., Hattie, J., & Tayler, C. (2016). Variations in the availability and quality of early childhood education and care by socioeconomic status of neighborhoods. *Early Education and Development*, *27*(3), 384–401. https://doi.org/10.1080/10409289.2015.1076674

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Conroy, M. A., & Sutherland, K. S. (2011). *BEST in CLASS—PK intervention: Research version 2.0* [Unpublished intervention]. Anita Zucker Center of Excellence in Early Childhood Studies, University of Florida.

Conroy, M. A., Sutherland, K. S., Algina, J., Ladwig, C., Werch, B., Martinez, J., Jessee, G., & Gyure, M. (2019). Outcomes of the BEST in CLASS intervention on teachers' use of effective practices, self-efficacy, and classroom quality. *School Psychology Review*, *48*(1), 31–45. https://dx.doi.org/10.17105/SPR-2018-0003.V48-1

Conroy, M. A., Sutherland, K. S., & Feil, E. (2019). *BEST in CLASS—Web: Research version 1.0* [Unpublished web-based training modules]. Anita Zucker Center of Excellence in Early Childhood Studies, University of Florida.

Dede, C., Ketelhut, D. J., Whitehouse, P., Breit, L., & McCloskey, E. M. (2009). A research agenda for online teacher professional development. *Journal of Teacher Education*, *60*(1), 8–19. https://doi.org/10.1177/0022487108327554

Downer, J. T., Pianta, R. C., Fan, X., Hamre, B. K., Mashburn, A., & Justice, L. (2011). Effects of web-mediated teacher professional development on the language and literacy skills of children enrolled in prekindergarten BEST in CLASS. *NHSA Dialog*, *14*(4), 189–212. https://doi.org/10.1080/15240754.2011.613129

Feil, E. G., Baggett, K., Davis, B., Landry, S. H., Sheeber, L. B., Leve, C., & Johnson, U. (2020). Randomized control trial of an internet-based parenting intervention for mothers of infants. *Early Childhood Research Quarterly*, *50*, 36–44. https://doi.org/10.1016/j.ecresq.2018.11.003

Feil, E. G., Walker, H. M., & Severson, H. H. (1995). The Early Screening Project for young children with behavior problems. *Journal of Emotional and Behavioral Disorders*, *3*(4), 194–202. https://doi.org/10.1177/106342669500300401

Fox, L., Dunlap, G., Hemmeter, M. L., Joseph, G. E., & Strain, P. S. (2003). The Teaching Pyramid: A model for supporting social competence and preventing challenging behavior in young children. *Young Children*, *58*(40), 48–52.

Fox, L., Hemmeter, M. L., Snyder, P., Binder, D. P., & Clarke, S. (2011). Coaching early childhood special educators to implement a comprehensive model for promoting young children's social competence. *Topics in Early Childhood Special Education*, *31*, 178–192. https://doi.org/10.1177/0271121411404440

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution. *Annals of Mathematical Statistics*, *29*, 885–891.

Hartley, H. O. (1940). Testing the homogeneity of a set of variances. *Biometrika*, *31*, 249–255.

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. https://doi.org/10.1177/0271121416653386

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

Ke, F., & Xie, K. (2009). Toward deep learning for adult students in online courses. *The Internet and Higher Education*, *12*(3–4), 136–145. https://doi.org/10.1016/j.iheduc.2009.08.001

Knight, D. S. (2012). Assessing the cost of instructional coaching. *Journal of Education Finance*, *38*(1), 52–80.

Knowles, M. S. (1984). *Andragogy in action: Applying modem principles of adult education*. Jossey-Bass.

Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, *62*, 399–402.

Lloyd, C. M., & Modlin, E. L. (2012). *Coaching as a key component in teachers' professional development improving classroom practices in Head Start settings*. U.S. Department of Health and Human Services, Office of Planning, Research and Evaluation (OPRE). http://www.mdrc.org/publications/620/full.pdf

Mauchley, J. W. (1940). Significance test for sphericity of a normal n-variable distribution. *Annals of Mathematical Statistics*, *11*, 204–209.

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. https://doi.org/10.1002/pits.20570

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews*, *82*, 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Newberg, J. (2005). *Battelle developmental inventory* (2nd ed.). Riverside.

Otte, S., Bangerter, A., Britsch, M., & Wuthrich, U. (2014). Attitudes of coaches towards the use of computer-based technology in coaching. *Consulting Psychology Journal: Practice and Research*, *66*(1), 38–52. https://doi.org/10.1037/a0035592

Patriota, A. G. (2017). On some assumptions of the null hypothesis significance testing. *Educational and Psychological Measurement*, *77*, 507–528. https://doi.org/10.1177/0013164416667979

Pianta, R. C., La Paro, K. M., & Hamre, B. (2008). *Classroom assessment scoring system (CLASS): Pre-K version*. Paul H. Brookes.

Reeves, T. D., & Pedulla, J. J. (2011). Predictors of teacher satisfaction with online professional development: Evidence from the USA's e-learning for educators initiative. *Professional Development in Education*, *37*(4), 591–611. https://doi.org/10.1080/19415257.2011.553824

Schindler, H. S., Kholoptseva, J., Oh, S. S., Yoshikawa, H., Duncan, G. J., Magnuson, D. 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. https://doi.org/10.1016/j.jsp.2015.04.001

Snyder, P. A., Hemmeter, M. L., & Fox, L. K. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Topics in Early Childhood Special Education*, *35*(3), 133–143. https://doi.org/10.1177/0271121415594925

Snyder, P. A., Hemmeter, M. L., & Fox, L. K. (2022). *Essentials of practice-based coaching*. Paul H. Brookes.

Snyder, P. A., Hemmeter, M. L., Meeker, K. A., Kinder, K., Pasia, C., & McLaughlin, T. (2012). Characterizing key features of the early childhood professional development literature. *Infants and Young Children*, *25*(3), 188–212. https://doi.10.1097/IYC.0b013e31825a1ebf

Spybrook, J., & Raudenbush, S. (2009). An examination of the precision and technical accuracy of the first wave of group-randomized trials funded by the Institute of Education Sciences. *Educational Evaluation and Policy Analysis*, *31*(3), 298–318. https://doi.org/10.3102/0162373709339524

Sutherland, K. S., McLeod, B. D., Conroy, M. A., Abrams, L. M., & 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. https://doi.org/10.1177/1063426613497258

Sutherland, K. S., McLeod, B. D., Conroy, M. A., & Cox, J. R. (2013). Measuring implementation of evidence-based programs targeting young children at risk for emotional/behavioral disorders: Conceptual issues and recommendations. *Journal of Early Intervention*, *35*(2), 129–149. https://doi.org/10.1177/1053815113515025

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

Thompson, B. (2002). Statistical, practical, and clinical: How many kinds of significance do counselors need to consider? *Journal of Counseling and Development*, *80*, 64–71. https://doi.org/10.1002/j.1556-6678.2002.tb00167.x

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Yoshikawa, H., Aber, J. L., & Beardslee, W. R. (2012). The effects of poverty on the mental, emotional, and behavioral health of children and youth: Implications for prevention. *American Psychologist*, *67*(4), 272–284. https://doi.org/10.1037/a0028015