Journal of Educational Psychology

A Comparative Analysis of Instructional Coaching Approaches: Face-to-Face Versus Remote Coaching in Preschool Classrooms

April Crawford, Cheryl Varghese, Hsien-Yuan Hsu, Tricia Zucker, Susan Landry, Michael Assel, Pauline Monsegue-Bailey, and Vibhutibala Bhavsar

Online First Publication, September 23, 2021. http://dx.doi.org/10.1037/edu0000691

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Educational Psychology

ISSN: 0022-0663

https://doi.org/10.1037/edu0000691

A Comparative Analysis of Instructional Coaching Approaches: Face-to-Face Versus Remote Coaching in Preschool Classrooms

April Crawford[1], Cheryl Varghese[1], Hsien-Yuan Hsu[2], Tricia Zucker[1], Susan Landry[1], Michael Assel[1], Pauline Monsegue-Bailey[1], and Vibhutibala Bhavsar[1]

[1] Children's Learning Institute, Department of Pediatrics, McGovern Medical School, University of Texas Health Science Center at Houston

[2] Department of Curriculum and Instruction, College of Education, University of Massachusetts, Lowell

This study contrasted face-to-face and remote coaching models using an empirically-based professional development model, The Early Education model (TEEM). Child care teachers (n = 174) were randomly assigned to 1 of 3 conditions: remote coaching (n = 59), face-to-face coaching (n = 66), and business-as-usual (n = 49). This study examined intervention effects in addition to coaching modality effects on teacher and child outcomes. Findings suggested that child care teachers who received coaching (regardless of modality) demonstrated improved teaching behaviors compared to the business-as-usual group. Compared to remote coaching, there were also advantages of face-to-face coaching on some teaching behaviors, although these advantages were no longer evident after accounting for teachers' levels of responsiveness to the intervention. Additionally, although there were no direct effects of the intervention on child outcomes, there was evidence of positive indirect associations of the intervention and coaching modality on child-level outcomes through improvements in teachers' instructional practices.

Educational Impact and Implications Statement

Remote and face-to-face coaching models were both found to be effective in improving many child care teachers' and children's outcomes. Findings from this study highlight the promise of professional development with coaching and point to the need for making these supports more accessible for early childhood teachers across settings.

Keywords: early childhood professional development, coaching modality, child care providers

Supplemental materials: https://doi.org/10.1037/edu0000691.supp

One of the most important determinants of high-quality preschool programs is the quality of the teaching workforce (Hamre et al., 2017). Effective teachers interact with young children in cognitively stimulating and emotionally supportive ways, which helps children acquire early learning skills (Yoshikawa et al., 2013). Recent research, however, shows tremendous variability in preschool teachers' provision of cognitively stimulating and emotionally supportive interactions with children (Hamre, 2014; Jamison et al., 2014; Justice et al., 2008). This is attributed to a variety of factors, including an underprepared teacher workforce (Sun et al., 2015) and challenges that are unique to preschool settings operating outside of the more established K-12 system (Hamre et al., 2017). In preschool child care settings, teachers often earn low wages and experience great turnover among staff; instability among child care teachers ranks near turnover rates of the fast food industry (National Survey of Early Care and Education Project Team [NSECE], 2013; Whitebook & Sakai, 2003). High turnover rates in early childhood education settings often leads to unstable environments that can compromise efforts to promote continuity in relational-care settings and adversely affect the quality of child care (Cassidy et al., 2011; McMullen et al., 2020). In efforts to build and maintain a stronger teacher workforce, many schools have invested in professional development opportunities for teachers, including coaching-based models that have the potential to provide sustained and individualized support for teachers (Kane & Rosenquist, 2019; Lemons & Toste, 2019).

Cheryl Varghese https://orcid.org/0000-0002-5537-4902

Hsien-Yuan Hsu https://orcid.org/0000-0003-2155-2093

Tricia Zucker https://orcid.org/0000-0003-3707-4946

We thank and acknowledge all of the Children's Learning Institute coaches and staff who worked on the study. Additionally, the research reported in this article was funded by a Goal 3 Grant from the Institute of Educational Sciences, R305A140378.

Correspondence concerning this article should be addressed to April Crawford, Children's Learning Institute, Department of Pediatrics, McGovern Medical School, University of Texas Health Science Center at Houston, 7000 Fannin Street, Suite 1920, Houston, TX 77030, United States. Email: April.crawford@uth.tmc.edu

Over the past decade, researchers have increasingly investigated how technology could be used to deliver professional development to teachers at larger scales (Downer et al., 2009; 2009; Landry et al., 2006; Powell & Diamond, 2013). Technology has been recently used as a way to provide coaching support for teachers (Crawford et al., 2017; Vernon-Feagans et al., 2018; Rock et al., 2009). Remote coaching models (i.e., coaching delivered through technology) eliminate travel so that rural locales can readily participate, allow for more flexible scheduling than traditional face-to-face coaching, and may be lower cost (Vernon-Feagans et al., 2015; Bates et al., 2019). However, despite the feasibility and potential benefits of remote coaching, there has been very little research comparing coaching delivery modalities (i.e., remote or face-to-face), particularly for preschool teachers in child care settings. To that end, this study examined an empirically-tested preschool teacher professional development program, The Early Education model (TEEM; Landry et al., 2009), which focused on building child care teachers' language and literacy instruction as well as their levels of responsiveness, to compare face-to-face and remote coaching through an experimental research design.

Compared to public-funded programs such as state preschool or Head Start, child care centers typically have fewer financial resources to provide professional development for their teachers and fewer incentives to meet early learning standards (Lin & Magnuson, 2018). Moreover, child care centers have been found to provide less optimal quality of care compared to Head Start and state preschool programs, which often have more stringent requirements for teacher qualifications and training in early childhood education and use more quality improvement monitoring (Bassok et al., 2016). Thus, the current study focused on efforts to provide TEEM to child care teachers as a way to improve the quality of high-need, center-based child care.

Findings from previous efficacy studies of TEEM demonstrated that the combination of online courses, face-to-face coaching, child progress monitoring tools, and curricular resources resulted in the most optimal changes in teachers' instructional practices (Cohen's d = .52 to 1.11; Landry et al., 2006). A follow-up efficacy study of TEEM utilizing a face-to-face coaching model suggested strong positive changes in teachers' responsive interactions and literacy, language, and math practices, with Cohen's d effect sizes ranging from .39 to 1.04 for teachers and .15 to .75 for child outcomes (Landry et al., 2009, 2011). Given the efficacy of TEEM professional development with a face-to-face coaching component, the purpose of the study was to understand if and how a remote coaching component would complement the TEEM professional development model (as compared to face-to-face coaching). The current study examined the effects of face-to-face versus remote coaching for teachers randomly assigned to one of the three conditions: face-to-face, remote, and business-as-usual. Additionally, face-to-face and remote coaching models were compared across a range of teacher- and child-level outcomes.

Coaching + Professional Development and Teacher and Child Outcomes

Professional development models that consist of standalone workshops typically do not provide the level of intensity or individualized support that is necessary for improving and sustaining teachers' instruction (Powell et al., 2009; Snyder et al., 2018). As a result, individualized coaching has emerged as a promising professional development method for improving the quality of early childhood classrooms (Yoon et al., 2007). For example, one experimental study showed that preschool teachers who received ongoing, job-embedded coaching focused on reading instruction were more likely to demonstrate higher scores on a measure of language and literacy-related teaching behaviors (Sibley & Sewell, 2011). Children are also likely to experience the benefits of their teachers working with coaches, with some evidence suggesting greater gains in language and literacy skills (Cabell & Downer, 2011; Landry et al., 2011) and improvements in their behavioral and social competencies (e.g., externalizing, prosocial, disruptive behaviors; Hemmeter et al., 2016; Reinke et al., 2014). Effects of coaching on child outcomes have been small to modest in comparison to larger positive effects of coaching on teacher outcomes (Kraft et al., 2018). This is likely attributed to the foci of many models that embed coaching within professional development models, which is to specifically improve teacher practices (Kennedy, 2016; Kraft et al., 2018; Varghese et al., 2016). In the current study, coaches specifically helped teachers to learn how to respond to student signals and use appropriate scaffolding strategies within the context of language and literacy instruction. Thus, it was hypothesized in the current study that associations of professional development with coaching would have proximal effects on teacher outcomes (language and literacy instruction) and indirect effects on child outcomes (language and literacy outcomes; behavioral and social competencies such as externalizing, internalizing, and prosocial behaviors) through teacher outcomes. The indirect effects of coaching suggest that without improving teacher practices, child outcomes are unlikely to improve. In previous studies, researchers have found early childhood teachers' language and literacy instructional practices to be a mediator of student outcomes (e.g., Piasta et al., 2020).

Coaching Modality and Teacher and Child Outcomes

Although face-to-face and remote coaching models have been separately linked to a variety of positive teacher outcomes, only one experimental study has compared face-to-face versus remote coaching models that supported preschool teachers' language and literacy instruction (Powell et al., 2010). Remote coaching has been characterized as being delivered asynchronously (e.g., delayed feedback on videos of teachers' lessons) or synchronously (e.g., immediate support during instruction). In asynchronous remote coaching models (used in the current study), teachers participate in individualized coaching sessions to review and reflect on a recorded video of instructional practice. For example, in a study conducted by Pianta and colleagues (2008), teachers submitted a video recording of a targeted instructional practice and a coach provided written feedback on segments of the video for the teacher to review. Coaches generated reflective questions for teachers that were specifically focused on interactions between teachers and students across a variety of domains. Coaches later followed up with teachers to debrief about the lesson and identify goals and modifications for instruction.

In the experimental study conducted by Powell and colleagues (2010), Head Start teachers were randomly assigned to one of two coaching conditions (on-site or online) or a business-as-usual condition. Although there were positive effects of both coaching models on

teachers' general classroom environment compared to the business-as-usual condition, there were few significant differences between the coaching models on the general classroom environment and teachers' instructional practices. Of the six teacher-level outcomes (general classroom environment; language, literacy, and curriculum; word definition instruction; code-focused instruction; language-eliciting book reading; and language-eliciting free play), teachers' code-focused instruction (e.g., points out a sound of a letter) was the only outcome positively influenced by on-site coaching. The researchers also found positive effects for remote coaching on two of the seven child outcomes (receptive vocabulary and initial sound-matching skills), but no effects for on-site coaching.

Despite the limited differences between the remote and face-to-face coaching groups, it is difficult to discern the unique contribution of coaching modality on teacher and child outcomes given the differences in coaching dosages and activities for the face-to-face and remote coaching groups. Additionally, the reported coaching protocol for both face-to-face and remote conditions primarily consisted of observations of specific practices and written feedback based on the observations. In the current study, face-to-face and remote coaches used a variety of strategies, in addition to observation and feedback (e.g., demonstration, side-by-side coaching, coteaching). Furthermore, the effects of both face-to-face and remote coaching models is likely to be contingent upon teachers' levels of responsiveness to and engagement with the professional development content and coach (Berkel et al., 2011; LoCasale-Crouch et al., 2016). Yet, there has been limited investigation as to whether teachers' levels of responsiveness accounts for differences between the two coaching modalities.

Similarities and Tradeoffs Between Face-to-Face and Remote Coaching Modalities

Both face-to-face and remote coaching models are aligned with tenets of situated learning theory, which emphasize that learning most often occurs in authentic contexts and that interactions between adults help further knowledge. Practical components of situated learning theory include cognitive apprenticeship and coaching, opportunities for practice, collaboration, and reflection. In both coaching modalities, coaches help teachers to apply content learned from professional development into their classroom settings. Coaches help teachers to make sense of their experiences in applying new content to classroom practices through ongoing conversations and help teachers to reflect on their experiences. In both coaching modalities, coaches are also able to model concepts and help teachers engage in observational learning (social learning theory, Bandura & Walters, 1977). For example, in face-to-face and remote coaching models, coaches may choose to either show demonstrations of effective implementation through video exemplars or demonstrate effective implementation themselves in order to improve teacher practice.

Despite these similarities, each coaching modality (i.e., face-to-face vs. remote coaching) may provide unique affordances that contribute to variations in teacher outcomes (Fishman et al., 2013). In the sections below, theoretical and empirical evidence are used to describe the advantages and potential obstacles of both face-to-face and remote coaching models, with face-to-face coaching models showing more advantages compared to remote coaching models.

Face-to-Face Coaching

Teachers who receive face-to-face coaching may experience a greater sense of relatedness and competence that increase teachers' intrinsic motivation to try new instructional practices and help teachers internalize feedback on instructional practices (self-determination theory). Face-to-face coaches may be more readily able to form trusting relationships with teachers (relatedness), which then influence a teacher's openness to change. Face-to-face coaches may also be better positioned to use a gradual release of responsibility approach, which can help improve teachers' sense of competence. For example, in face-to-face coaching models, coaches may coteach or model lessons/activities so that teachers can transition toward implementing the lessons or activities independently. Relatedly, theoretical tenets of cognitive apprenticeship suggest that learning occurs through guided experiences and that the initial stages of learning require demonstrations and guidance. Face-to-face coaching provides teachers with an opportunity to have someone demonstrate how a new strategy or lesson can work in their own classroom context (Poglinco & Bach, 2004). Approaches that encourage teachers to make personal connections with what they are learning may improve the chances that teachers will "buy in" to the intervention (Poglinco & Bach, 2004) and increase their receptivity to change. When in the classroom, coaches are able to provide 'in the moment' feedback and demonstrate new strategies or lessons that also allow teachers to attend to the subtleties (i.e., tacit knowledge) of implementing new practices that cannot always be easily articulated verbally. Relatedly, an application of cognitive theory suggests that face-to-face coaches may be better positioned to help teachers learn how to manage the competing demands that arise during classroom instruction (e.g., managing student cues, delivering lessons, adapting instruction based on student signals) by using strategies such as side-by-side coaching. Finally, face-to-face coaching makes it easier to observe and reinforce the use of teaching strategies across multiple classroom contexts (e.g., coach can provide in-the-moment examples of using open-ended questions during transitions, read alouds, or snack time). Collectively, the advantages of face-to-face coaching models may be important for child care teachers, who may have lower levels of education and fewer years of experience compared to teachers working in public preschool or Head Start programs. Immediacy of feedback, modeling activities in teachers' classroom contexts, and gradual release of responsibility, in particular, may be important attributes of face-to-face coaching models for teachers with lower levels of education and experience.

The primary potential disadvantages to face-to-face coaching include the practical need to schedule classroom modeling/observation, travel costs (particularly to rural areas or to programs that lack internal coaching staff/specialists), and the difficulty of timing feedback and reflection sessions when children are present. This may not allow teachers the time/space to reflect as in-depth as a later coaching call.

Remote Coaching

Remote coaching provides an opportunity for teachers to review and reflect upon a video of classroom instruction. Video recordings can be a powerful tool for supporting reflection as it allows teachers an opportunity to see what children are experiencing. This may be especially helpful in situations in which a teacher struggles to recognize or misses opportunities to respond to student cues. Therefore, the video

recording can provide an opportunity for the teacher to gain a different perspective. Relatedly, teachers may also improve in their abilities to critically evaluate and improve upon their own instructional practices and engage in productive inquiry (Matsumura et al., 2019; Pianta et al., 2008; Powell et al., 2010). Second, remote coaches may be able to consistently provide teachers with more rationale and reasoning in the context of instruction because they have focused forums in which they are able to review recordings of teachers' instruction and help the teacher attend to specific instructional practices (Vernon-Feagans et al., 2015). In face-to-face coaching models, coaches may not be able to consistently provide extensive rationale and reasoning for specific instructional practices because of competing demands in early childhood classrooms (e.g., teacher has to attend to other students, has to transition to other classroom activities). Additionally, teachers may not be able to fully process coach feedback in-the-moment and may simply follow coach prompts because the teacher is focused on returning to the classroom. Teachers may be able to internalize and process more of the coach's feedback when they have an opportunity to review their instructional practices alongside coach feedback/commentary. Third, child care teachers often have limited time away from students during their working hours because, unlike district-based preschools, they are often responsible for food service, classroom cleaning, and sometimes student transportation. Therefore, remote delivery can allow for scheduling coaching calls when the teacher is off-duty and has sufficient time to focus on reflection and feedback.

There are also potential disadvantages to remote coaching, which can include technology-related difficulties, limited opportunities to build strong coach-teacher relationships, missed opportunities to intervene in-the-moment, and inabilities to model lessons with the teacher's own students. Technology-related difficulties may include slow Internet speed (particularly in rural areas) and a learning process for using technology to record and upload videos of instructional practices and connect with coaches remotely. Remote coaching models may also limit social cues that build trust and rapport between the coach and teacher, which may compromise coach-teacher relationships. Additionally, in asynchronous remote coaching models, coaches are not able to intervene in-the-moment, which often results in delayed teacher feedback. Moreover, remote coaches are not typically able to model lessons or activities using the teacher's own students (rather than video exemplars). This may compromise teachers' abilities to transfer coach feedback to their own practices.

Goals of Present Study

The present study examined the benefits of coaching delivery approaches for improving teachers' instructional practices. This study built upon and extended two past lines of research (Landry et al., 2006; 2011; Powell et al., 2010) in five key ways: (1) the present sample focused exclusively on child care settings; (2) the content of professional development was based on an online language and literacy course; (3) coaching support was provided throughout the school year; (4) coaching activities and professional development resources between the face-to-face and remote coaching groups were designed to be similar (to the extent possible), but more diverse than what was reported in previous studies; (5) the theory of change for TEEM (with coaching) assumed indirect child effects through improvements in teacher practices.

Classrooms were randomized to one of three conditions: remote coaching, face-to-face coaching, or business-as-usual. Both intervention groups received equivalent additional supports including online courses for teachers, online child progress monitoring tools, and curricular resources. The primary goal of the study was to examine the extent to which professional development training with coaching support led to greater gains in child care teachers' instructional quality that, in turn, improved child outcomes. The following research questions were addressed:

1. What were the intervention effects (regardless of coaching modality) on child care teachers' language and literacy instruction compared to business-as-usual?

2. What were the effects of face-to-face coaching delivery compared to remote delivery with regard to child care teachers' language and literacy instruction?

3. How did effects of the intervention and coaching modality on child care teachers' language and literacy instruction further influence children's language and literacy outcomes and behavior and social competencies?

The hypothesis for the first research question was that intervention teachers would show greater gains in the quality of language and literacy instruction than the business-as-usual group teachers. The hypothesis for the second research question was that there would be stronger effects for the face-to-face coaching condition compared to the remote coaching condition, such that teachers receiving face-to-face coaching would show greater gains in the quality of language and literacy instruction than teachers receiving remote coaching. The hypothesis for the third research question was that the effects of the intervention and face-to-face coaching in teachers' language and literacy instruction would be indirectly associated with children's language and literacy outcomes and would improve children's behavioral and social competencies.

Method

The study took place in center-based child care classrooms with three cohorts. Each cohort participated for one school year beginning in 2014, 2015, or 2016. Written permission from the university IRB, the director of each participating school, each teacher, and parent of each student were obtained. The study was a cluster randomized controlled trial, with students nested within classes, and one class per building. Each classroom was randomly assigned to one of the three conditions: face-to-face coaching, remote coaching, or business-as-usual.

Participants

Teacher demographics are shown by condition in Table 1. The vast majority of child care teachers (72.42%) in this sample had only a high school or Associate's degree and about 16% of teachers had a bachelor's degree or higher. Almost all teachers were female. Most teachers were African American (52.87%), Hispanic (27.59%), or White (10.92%). Teachers reported an average of

Table 1
Descriptive Information for Teacher Sample

Teacher-level variables	Full sample (n = 174)% or M (SD)	Face-to-face coaching (n = 66)% or M (SD)	Remote coaching (n = 59)% or M (SD)	Business-as-usual (n = 49)% or M (SD)
Ethnicity				
White	10.92	10.61	10.17	12.24
African American	52.87	50.00	52.54	57.14
Hispanic/Latinx	27.59	34.85	23.73	22.45
Other	8.62	4.55	13.56	8.16
Education level				
No high school diploma	1.72	1.52	0.00	4.08
High school diploma/GED	64.37	66.67	57.63	69.39
Associate's degree	8.05	9.09	8.47	6.12
Bachelor's degree	8.62	1.52	16.95	8.16
Some graduate work	4.60	7.58	1.69	4.08
Master's degree	2.87	4.55	1.69	2.04
Other	5.75	7.58	5.08	4.08
CDA credential	4.02	1.52	8.47	2.04
Years of teaching experience	9.67 (8.50)	9.35 (7.76)	9.58 (8.29)	10.22 (9.39)

9.67 years of teaching experience, although there was considerable variability (SD = 8.50; range 0–40 years).

Child participants were mostly four-year-old children from low-income backgrounds. Average child age was 4.0 at the beginning of the school-year. Participating children were 46.50% African American, 30.64% White non-Hispanic, 11.43% Hispanic/Latinx, and 11.43% other, as shown in Table 2.

Setting

Classrooms were located in two large urban Southwestern cities. To be eligible to participate, classrooms had to serve toddlers, provide a full-day program (not half-day), and deliver the majority of instruction in English. Classrooms were not eligible if they had participated in past professional development projects offered by the University researchers within the last 5 years. Based on past child care studies with large turnover, steps were taken to reduce attrition before randomization. Program directors and teachers signed commitment letters stating their intentions to make efforts for focal teachers and children to remain within the same classroom throughout the study period.

Randomization

The random assignment intervention was carried out by the lead methodologist who had no involvement in intervention activities. Classrooms were the unit of assignment (with one classroom per building). A total of 217 classrooms/teachers were initially randomized throughout three years. The initial allocation was: (a) face-to-face = 74, (b) remote = 75, and (c) business-as-usual = 68.

Teacher Attrition. As expected within this child care setting, there was considerable attrition in this study. Table 3 shows the numbers of stayers, leavers, and joiners, and the attrition rate for the full teacher sample and by experimental group throughout the study. Of the original 217 teachers randomized, 60 teachers (remote = 20; face-to-face = 16; business-as-usual = 24) were dropped from the study (i.e., leavers). The reasons for attrition were presented by experimental groups in Appendix A.[1] Seventeen substitute teachers (i.e., joiners; face-to-face = 8; remote = 4; business-as-usual = 5) replaced the dropped teachers before they were aware of the randomly assigned conditions of the classrooms and remained from the pretest through posttest. The final analytic sample of classrooms/teachers was 174 (face-to-face = 66; remote = 59; business-as-usual = 49).

After taking the number of joiners into account, the overall attrition rate was 19.82%, as presented in Table 3. The attrition rates for three experimental groups were 10.81% (face-to-face), 21.33% (remote), and 27.94% (business-as-usual). Results of chi-square tests suggest that the business-as-usual group had a statistically higher attrition rate than the face-to-face group ($p < .05$), but no other statistically significant results were found. Moreover, the absolute differential attrition rates for face-to-face v. remote (10.52%) and for remote v. business-as-usual (6.61%) were in tolerable levels of potential bias, whereas the absolute differential rates for face-to-face v. business-as-usual (17.13%) were not in tolerable levels of potential bias (What Works Clearinghouse, 2020). The high differential attrition rate was due to the high attrition rate of the business-as-usual group (27.94%). Although there was a higher differential attrition rate between the face-to-face and business-as-usual group, the reasons for attrition in the business-as-usual group were unrelated to the study (see Appendix A).

Child Attrition. Table 4 shows the numbers of children who participated and children who dropped out of the study at posttest, as well as the attrition rate for the full sample (by group). A total of 1,297 children participated in the study (482 children in the face-to-face group; 443 children in the remote group; and 372 children in the business-as-usual group). There were 282 children (face-to-face = 96; remote = 109; business-as-usual = 77) who did not complete both pretest and posttest for the study (i.e., children with nonresponses). As a result, the final analytic sample of children was 1,015 (face-to-face = 386; remote = 334; business-as-usual = 295). The overall attrition rate was 21.74%. The attrition rates for three experimental groups were 19.92% (face-to-face), 24.60% (remote), and 20.70% (business-as-usual). The absolute differential attrition rates between three experimental groups were 4.69% (face-to-face v. remote), .78% (face-to-face v. business-as-usual), and 3.91%

[1] Appendices A–H can be found in the online supplemental materials.

Table 2
Descriptive Information for Child Sample

Child-level variables	Full sample (N = 1,015)%	Face-to-face coaching (n = 386)%	Remote coaching (n = 334)%	Business-as-usual (n = 295)%
Gender				
Male	51.82	52.85	50.60	51.86
Female	48.18	47.15	49.40	48.14
Ethnicity				
White	30.64	32.38	26.05	33.56
African American	46.50	45.85	52.10	41.02
Hispanic/Latinx	11.43	9.84	11.08	13.90
Other	11.43	11.92	10.78	11.53

(remote v. business-as-usual). According to What Works Clearinghouse (2020) standards, these differential attrition rates were in tolerable levels of potential bias.

TEEM Description

The TEEM intervention components primarily consisted of: 1) nine online courses aimed at increasing knowledge of appropriate language, literacy, and responsiveness strategies; 2) training and resources to conduct student progress monitoring, and 3) training in the use of a supplemental curriculum. Course work and training in the use of student progress monitoring software and the supplemental curriculum were conducted using a web-based platform and laptop or tablet. The supplemental curriculum consisted of the Developing Talkers program (e.g., books with linked language and literacy activities, activity guides) and school readiness kits (e.g., literacy, math and science materials). The TEEM intervention was designed to help teachers understand how to use both assessments and curriculum-linked instruction in language and literacy. All TEEM teachers spent approximately 50 hours interacting with course content over the course of a year. Coaching support (described more fully in the section 'TEEM Professional Development and Coaching Procedures') was also provided to all intervention teachers in order to facilitate implementation of the three intervention components. TEEM teachers received similar coaching support that helped: (a) reinforce concepts introduced in the course, (b) ensure proper implementation of student progress monitoring in areas of letter knowledge, phonological awareness, and vocabulary at three time points during the year, and (c) provide curriculum support. In addition, teachers also had access to a video library that included video exemplars of a broad range of language and literacy lessons (e.g., book reading, phonological awareness, print and letter knowledge, written expression, and oral language) that took place in a variety of settings (e.g., whole group, small group, transitions, and centers).

Teacher Training

In the fall of each study year, intervention teachers (both face-to-face and remote) attended a one-day overview session during which they received information on the intervention materials they would receive, how to access TEEM course work through an online platform, and how to conduct student progress monitoring on the online platform. Teachers in the remote intervention group attended an additional session on how to record and upload videos of their teaching footage. During the year, intervention teachers also attended two one-hour webinars to help support language development and dialogic reading.

Coach Training

Over the course of the study, nine coaches and three coach supervisors were hired through the university prior to the start of the first study year. Of the nine coaches, one was an exclusively remote coach and two were exclusively face-to-face coaches. The remaining six coaches worked with both face-to-face and remote groups. One coach supervisor worked with two teachers in the remote group and two of the coach supervisors worked with seven teachers in the face-to-face group. All coaches received ongoing training on recruitment of study sites, study overview (e.g., key coaching responsibilities, implementation plan), coach data collection measures (e.g., coaching session logs), and content training (e.g., coaching competencies, coaching strategies and feedback); in total, this amounted to 13 hours of training. Coaches also attended 18 one-hour biweekly meetings to discuss coaching content and implementation progress. During the weekly meetings, a coach supervisor conducted professional learning community

Table 3
Teacher Attrition Rates

Teacher participation groups	Face-to-face coaching n or %	Remote coaching n or %	Business-as-usual n or %	Total n or %
Random assignment (A + B)	74	75	68	217
Stayers (A)	58	55	44	157
Leavers (B)	16	20	24	60
Joiners (C)	8	4	5	17
Final analytical sample (A + C)	66	59	49	174
Attrition rate after replacement [=(B − C)/(A + B)]	10.81%	21.33%	27.94%	19.82%

Table 4
Child Attrition Rates

Child participation groups	Face-to-face coaching *n* or %	Remote coaching *n* or %	Business-as-usual *n* or %	Total *N* or %
Participating children (A)	482	443	372	1,297
Children who left the study prior to posttest (B)	96	109	77	282
Final analytical sample (A-B)	386	334	295	1,015
Attrition rate [=B/A]	19.92%	24.60%	20.70%	21.74%

sessions based on coaches' needs, with topics including classroom management, lesson cycle, letter knowledge, phonological awareness, language development, written expressions, reflection, and feedback. Additionally, coaches reviewed coaching videos to identify important coaching competencies and to reflect on ways to improve coaching practices. The intervention director and project manager also attended the weekly meetings to monitor the intervention and provide content expertise.

TEEM Coaching Procedures

Both face-to-face and remote coaches adhered to the same biweekly cycle of: observation, assessment, evaluation, feedback, and planning. Observation and feedback were based on the Classroom Observation Tool (COT; Crawford et al., 2017) and COT Mentoring Guide. The COT consists of short checklists of quality teaching behaviors across a range of instructional areas including: sensitivity and responsiveness, oral language use, book reading, print and letter knowledge, phonological awareness, and written expression (see Appendix B for sample of the COT). Coaches used these observational data to guide their discussions with teachers and collaborate in the drafting of Short Term Goals Reports (STGR). During coach feedback sessions, coaches also selected video exemplars from the video library to help teachers understand high quality implementation before teachers attempted to implement the agreed upon changes. The STGR helped to provide a focus for coaching, track fidelity, and set goals aligned with course content. Teachers received an electronic copy of the STGR, which also included connections to preschool guidelines, activities in a supplemental curriculum, and a record of the action plan for the following week (jointly developed at the end of the coaching session).

Coaches also had access to the coach version of the STGR, which linked COT behaviors to specific coaching strategies from the COT Mentoring Guide (see Appendix C). These reports helped coaches to focus and adhere to the intervention during coaching sessions. Coaches in both conditions spent approximately 90 minutes working with each teacher on a biweekly or monthly basis from October to April of each study year. In the sections below, face-to-face and remote coaching models are described.

Face-to-Face Coaching

The face-to-face coaching model followed a gradual release progression in which the coach withdrew support as the teacher was able to implement lessons or activities independently. Face-to-face coaches drew upon several coaching strategies to individualize support for teachers within the gradual release progression. For example, the progression may have begun with an in-class demonstration, followed by coteaching, and concluded with side-by-side prompting, depending on the teacher's level of expertise in the instructional area. These strategies provide context-specific, in-the-moment support that is difficult to replicate in remote coaching models. Face-to-face coaches also drew upon video reflection strategies when necessary, recording a lesson and using the footage as a discussion prompt during the feedback session. Overall, coaching sessions in the face-to-face coaching condition were structured as follows: (a) implement the action plan for 60 minutes (i.e., support teacher according to the gradual release model in agreed upon content areas), (b) conduct reflective follow-up session for 30 minutes (i.e., discuss observed teaching practices; determine needs based on observation, video reflection, course work discussions, and progress monitoring; and set goals), and (c) provide teacher with short term goals report and action plan for the coming week.

Remote Coaching

The remote coaching model followed the face-to-face coaching model to the extent that it was feasible. Teachers in the remote condition were provided with tablets that had video recording capabilities in order to record weekly teaching footage that corresponded with established coaching goals. For example, if a coach and teacher set goals around conducting effective read alouds, the teacher would record a daily read aloud, upload the footage securely, and view the uploaded video while completing a reflection form that was sent to the coach. When the coach received the video, their protocol was as follows: (a) review the description, reflection form, and video, (b) select clips and embed onscreen captioning and voiceover for repeated viewing and discussion, (c) select video exemplars that support identified goals, (d), and provide the teacher with a video link and discussion points to review before the conference call. During the coaching call, teachers and coaches viewed the video jointly while following the coaching cycle (i.e., discuss observed teaching practices, identify areas of need based on observed teaching, discuss course work and child progress monitoring, set goals, view demonstrations). At the conclusion of the session, the coach provided the teacher with the agreed-upon goals using the STGR and action plan for the coming week. Delays associated with video recording and upload required remote coaching to be delivered in two phases. In the first phase, the teacher spent 30 minutes reviewing coach feedback (i.e., video annotations, video voiceovers, and reflective questions), and in the second phase, the teacher participated in a 60-minute remote coaching session to talk about the annotated lesson, establish new goals, and review or discuss the video demonstration.

Business-as-Usual Condition

To understand the counterfactual condition, teachers in all conditions reported on their core curriculum. Teachers in the control group reported using a variety of core curricula (e.g., Frog Street, Mother Goose). They also completed an adaptation of the Early Childhood Longitudinal study Teacher Questionnaire items about content coverage for language arts and classroom resources available for this instruction. The ECLS survey included 27 items that asked teachers to report how often they taught specific literacy and social-emotional skills by selecting one of four options for each skill or strategy: 1 = not at all, 2 = somewhat, 3 = a fair amount, or 4 = a lot. On average, teachers reported an average of a "fair amount" for both teaching literacy skills (M = 3.03, SD = .36) and social-emotional skills (M = 3.54, SD = .36).

Fidelity of Implementation

Teacher Fidelity

Teachers' exposure and responsiveness to the intervention were the two primary fidelity indicators examined in the current study. Table 5 presents descriptive statistics of teacher-level fidelity variables. Teachers' exposure to the intervention was conceptualized as the extent to which teachers participated in the key components of the TEEM intervention (online courses and coaching cycles). Exposure to the intervention was assessed in two ways: the number of online courses completed and the number of coaching cycles (i.e., goals set, practice, and feedback). On average, teachers completed seven out of nine online courses. Teacher goals were tracked within coaching cycles using the Classroom Observation Tool (COT), which had a selection of over 200 goals across 13 instructional domains for teachers to choose from based on their own professional goals and classroom needs (it should be noted that there was not a maximum number of goals that teachers were expected to set). Face-to-face teachers participated in an average of 11 coaching cycles (out of 12 possible coaching cycles), set approximately 44 goals with their coaches, and met approximately 39 of their goals. Remote coaches participated in an average of 11 coaching cycles, set approximately 59 goals with their coaches, and met approximately 55 of their goals.

Teachers' responsiveness was conceptualized as the quality of their participation in the key TEEM intervention components (online courses and coaching cycles). After each coaching cycle, coaches used a 5-point Likert-type scale (1 = low to 5 = high) to rate teachers' levels of coachability, engagement, mastery of the target activity, mastery of the course content, and reflectiveness; these domains captured the extent to which teachers actively engaged with the TEEM intervention components (see Appendix D for descriptions of each dimension). Prior to coaching teachers, coaches attended three training sessions (led by the Principal Investigator and coach supervisor) and independently rated two coaching sessions after the training. Coach ratings were then reviewed to ensure adequate understanding of the rating form. For face-to-face teachers, coach ratings ranged from 3.28 to 4.11 (SDs ranged from .57 to .68). For remote teachers, coach ratings ranged from 2.58 to 3.97 (SDs ranged from .63 to .77). Coaches also documented the specific coaching strategies they used during the coaching session (e.g., coteaching, video reflection) in addition to the classroom context (e.g., small group, whole group), content area (e.g., letter knowledge), and activity (e.g., supplemental curricula lesson).

Coach Fidelity

Coach fidelity data were collected to inform feedback to coaching staff and improve intervention delivery. The Principal Investigator met with the coach supervisors to discuss each item on the coach fidelity rating form and the coach supervisor provided examples for each of the items to clarify items on the coach fidelity rating form. Coach supervisors independently rated three videos; these ratings were discussed during a later meeting to ensure consistency. Supervisors rated each of the coaches during coaching sessions twice per month using a Supervisor Fidelity Rating form, which consisted of five items scored on a Likert-type scale (e.g., 1 = minimal, 3 = moderate, 5 = high). Each of the five items assessed different coaching practices. On average, supervisors rated intervention coaches as demonstrating moderate levels across the four items: Actionable Feedback (M = 3.31, SD = 1.28); Content focused feedback (M = 3.49, SD = 1.16); Supportive (M = 3.82, SD = .90); and Encouraging Reflective Thinking (M = 2.56, SD = 1.20). Supervisors rated face-to-face and remote coaches as being able to moderately match coaching strategies to teachers' needs (M = 2.76, SD = 1.01; M = 3.56, SD = 1.35, respectively). See Table 6 for more descriptive information on coaches' implementation fidelity.

Table 5
Descriptive Information for Teacher Level Fidelity

Teacher fidelity indicators	All						Face-to-face coaching						Remote coaching					
	N	*M*	*SD*	Range		Skewness	*N*	*M*	*SD*	Range		Skewness	*N*	*M*	*SD*	Range		Skewness
Number of online courses	83	7.66	2.24	0.00	9.00	−1.87	44	7.34	2.45	0.00	9.00	−1.57	39	8.03	1.95	1.00	9.00	−2.43
Number of cycles	83	11.46	1.57	4.00	12.00	−3.46	44	11.55	1.25	5.00	12.00	−3.87	39	11.36	1.88	4.00	12.00	−3.09
Coachability	125	3.85	0.72	1.71	5.00	−0.59	66	4.08	0.60	2.90	5.00	−0.13	59	3.58	0.77	1.71	4.83	−0.58
Current course content mastery	125	2.95	0.80	1.00	4.23	−0.34	66	3.28	0.68	1.33	4.23	−0.73	59	2.58	0.76	1.00	4.08	0.07
Current target activity mastery	125	3.30	0.60	1.44	4.42	−0.54	66	3.36	0.57	1.67	4.18	−0.79	59	3.24	0.63	1.44	4.42	−0.30
Reflectiveness	125	3.51	0.70	1.57	5.00	−0.32	66	3.71	0.68	2.17	5.00	−0.29	59	3.30	0.65	1.57	4.42	−0.58
Engagement	125	3.95	0.62	2.29	5.00	−0.46	66	4.11	0.58	2.93	5.00	−0.37	59	3.78	0.63	2.29	4.90	−0.49
Coaching support	125	4.03	0.62	2.09	5.00	−0.40	66	4.07	0.61	3.00	5.00	−0.02	59	3.97	0.63	2.09	5.00	−0.80

Table 6
Descriptive Information for Supervisor Rating of Coaches' Fidelity

	All						Face-to-face coaching						Remote coaching					
Coach fidelity indicators	*N*	*M*	*SD*	Range		Skewness	*N*	*M*	*SD*	Range		Skewness	*N*	*M*	*SD*	Range		Skewness
Content-focused feedback	103	3.49	1.16	1	5	−0.27	51	3.04	1.06	1	5	−0.19	52	3.88	1.13	1	5	−0.53
Actionable feedback	103	3.31	1.28	1	5	−0.09	51	2.82	1.05	1	5	0.26	52	3.69	1.38	1	5	−0.58
Supportive feedback	103	3.82	0.90	2	5	−0.19	51	3.33	0.77	2	5	−0.1	52	4.29	0.78	3	5	−0.56
Reflective thinking	103	2.56	1.20	1	5	0.45	51	2.08	0.89	1	4	0.37	52	2.98	1.32	1	5	0.09
Intensity	103	3.17	1.25	1	5	0.02	51	2.76	1.01	1	4	0.02	52	3.56	1.35	1	5	−0.38

Note. Likert scale ranges from *1*-low to *5*-high.

Data Collection Procedures

Following the training sessions and once deemed reliable (for information on the training procedures of classroom observers and child assessors, please see Appendix E), research assistants conducted the teacher observations and child assessments. In the fall and spring of each study year, teachers completed questionnaires about their professional background, classroom characteristics, and demographic information. Teachers assisted the research staff by helping to ensure consent forms and parent questionnaires were delivered home to children via "backpack communications." These questionnaires were completed by caregivers and included family demographic information. Children were assessed in the fall and spring of each study year in quiet locations within their schools (e.g., empty classroom, unused office space). Classroom observations occurred during the school day and were scheduled during portions of the day when teachers were actively engaged in providing instruction. In other words, efforts were made to schedule observations during small and large group lessons versus when children were engaged in other activities (e.g., lunch, nap, outside time).

Measures

Teacher Outcomes

To assess changes in teacher behaviors, assessors administered the Teacher Behavior Rating Scale (TBRS; Landry et al., 2001); and to assess the quality of teachers' language and literacy environment, assessors administered the Early Language and Literacy Classroom Observation (ELLCO; Smith et al., 2002).

TBRS. The TBRS evaluated the following content areas: (a) classroom community behaviors; (b) teacher sensitivity; (c) learning centers; (d) lesson plans; (e) book reading behaviors; (f) print and alphabet knowledge; (g) written expression; (h) phonics; (i) phonological awareness; and (j) oral language use with students. For TBRS each subscale, items on both quality ratings (1 = low; 4 = high) and quantity ratings (1 = rare; 4 = often) were scored using a 4-point scale. Quality mean score and quantity mean score were averaged to compute the subscale score used in analyses. The total score (ranged from 11.64 to 33.23) was calculated by adding the subscale's scores. The means and standard deviations for each of the TBRS subscales are presented in Table 7. The TBRS has previously shown high reliability and validity (e.g., interrater reliability ranged from .71 to 1.00; Landry et al., 2001). In the current study, ICC (intraclass correlation coefficients) was computed as an indicator of interrater reliability. Cicchetti's (1994) criteria of ICC values helped to inform different levels of interrater reliability: poor (ICC $<$ .40), fair (.40 $\leq$ ICC $<$ .60), good (.60 $\leq$ ICC $<$ .75), and excellent (.75 $\leq$ ICC). In general, subscales had good to excellent ICC values (ranged from .72 to .93) except for Phonics. The Phonics subscale had an ICC of .44, which was still considered as a fair interrater reliability coefficient.

ELLCO. Two subtests from the ELLCO were used in the current study: General Classroom Environment (GCE) and the

Table 7
Descriptive Statistics of Teacher Assessment Scores

	Full sample (n = 174) *M* (*SD*)		Face-to-face coaching (n = 66) *M* (*SD*)		Remote coaching (n = 59) *M* (*SD*)		Business-as-usual (n = 49) *M* (*SD*)	
TBRS dimensions	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
TBRS total score	17.51 (3.01)	20.00 (4.25)	17.93 (3.29)	22.04 (4.36)	17.17 (2.84)	20.30 (3.51)	17.34 (2.82)	16.88 (3.00)
Classroom community	2.10 (0.45)	2.24 (0.52)	2.14 (0.44)	2.33 (0.52)	2.04 (0.46)	2.28 (0.48)	2.12 (0.46)	2.05 (0.53)
Teacher sensitivity	2.77 (0.46)	2.88 (0.46)	2.80 (0.48)	3.03 (0.47)	2.75 (0.50)	2.88 (0.39)	2.76 (0.40)	2.68 (0.47)
Learning centers	1.52 (0.53)	1.87 (0.68)	1.61 (0.65)	2.18 (0.72)	1.40 (0.40)	1.86 (0.58)	1.55 (0.50)	1.45 (0.49)
Lesson plans[a]	1.58 (0.69)	1.64 (0.87)	1.69 (0.74)	1.76 (0.98)	1.47 (0.65)	1.68 (0.89)	1.57 (0.66)	1.42 (0.62)
Book reading	1.72 (0.45)	2.12 (0.58)	1.71 (0.49)	2.36 (0.57)	1.75 (0.43)	2.19 (0.50)	1.68 (0.41)	1.72 (0.48)
Print and letter knowledge	1.67 (0.39)	2.02 (0.55)	1.74 (0.44)	2.28 (0.55)	1.66 (0.38)	2.09 (0.48)	1.58 (0.32)	1.60 (0.38)
Written expression	1.33 (0.33)	1.58 (0.54)	1.35 (0.38)	1.86 (0.6)	1.32 (0.25)	1.55 (0.48)	1.33 (0.34)	1.24 (0.27)
Phonics	1.51 (0.53)	1.77 (0.66)	1.52 (0.55)	1.96 (0.8)	1.48 (0.49)	1.84 (0.54)	1.52 (0.56)	1.42 (0.44)
Phonological awareness	1.25 (0.37)	1.58 (0.59)	1.31 (0.4)	1.82 (0.58)	1.23 (0.36)	1.62 (0.62)	1.21 (0.33)	1.20 (0.34)
Oral language	2.09 (0.46)	2.32 (0.51)	2.10 (0.48)	2.50 (0.47)	2.07 (0.50)	2.30 (0.47)	2.08 (0.40)	2.11 (0.53)
ELLCO total score	46.26 (10.33)	52.36 (12.72)	48.20 (11.10)	57.50 (11.44)	45.00 (9.55)	53.53 (11.74)	45.16 (9.97)	44.02 (11.41)

Note. TBRS = Teacher Behavior Rating Scale; ELLCO = Early Language and Literacy Classroom Observation.
[a] Four teachers (2 in face-to-face and 2 in business-as-usual) did not have complete data in pretest or post scores of lesson plans and those teachers were not included in the analysis of lesson plans.

Language, Literacy, and Curriculum (LLC). The GCE subtotal was calculated based on ratings of classroom materials, organization, and management. The LLC subtotal was calculated based on ratings of oral language facilitation, approaches to book reading, and other curriculum aspects. In the current study, the ICCs for the GCE and LLC were .80 and .62, respectively. Items on ELLCO were scored using a 5-point Likert-type scale, which ranged from 1 (deficient) to 5 (exemplary). The total score was calculated by adding all 19 items on the two subtests. The total score for the ELLCO was used in the current study (see Table 7 for means and standard deviations for the full sample and groups).

Child Outcomes

Three standardized measures of children's literacy and language skills were used in the current study: Preschool Language Scales-Fifth Edition (PLS-5; Zimmerman et al., 2002); Test of Preschool Early Literacy (TOPEL; Lonigan et al., 2007); and Expressive One-Word Picture Vocabulary Test (EOWPVT; Brownell, 2000). Additionally, two teacher-rated measures of child behaviors were used: Child Behavior Questionnaire (CBQ; Rothbart et al., 2001) and Social Competence and Behavior Evaluation—Teacher ratings (SCBE; LaFreniere & Dumas, 1996). The means and standard deviations for each of the child assessments are presented in Appendix G (online supplemental materials).

PLS-5. The PLS-5 is a global oral language measure that consists of the Auditory Comprehension subscale, which assesses receptive language, and the Expressive Communication subscale, which assesses the child's ability to communicate with others. Internal consistency estimates were high for both Auditory Comprehension (α = .94 and .93 at preand posttest) and Expressive Communication (α = .96 at preand posttest).

TOPEL. Two TOPEL subtests were administered: Print Knowledge, which consisted of 36 items that measured skills such as letter recognition and book and print concepts, and Phonological Awareness, which consisted of 27 items that measured skills such as elision and blending. Internal consistency estimates were high for Print Knowledge (α = .97 at pretest and α = .96 at posttest) and Phonological Awareness (α = .88 at pretest and α = .90 at posttest).

EOWPVT. The EOWPVT is a norm-referenced assessment that measures expressive vocabulary. Children are presented with a series of illustrations depicting an object, action, or concept and asked to name each illustration. Internal consistency estimates were high (α = .97 at preand posttest).

CBQ. The CBQ consists of three subscales: Attention/Focus, which consisted of 14 items (e.g., when picking up toys, usually keeps at the task until it is done), Inhibitory Control, which consisted of 13 items (e.g., can wait before entering into new activities if asked), and Impulsivity, which consisted of 13 items (e.g., usually rushes into an activity without thinking about it). Teachers rated students across the three domains, using a seven-point Likert-type scale (1 = extremely false to 7 = extremely true). Internal consistency estimates were medium to high for Attention/Focus (α = .81 at preand posttest), Inhibitory Control (α = .86 at preand posttest), and Impulsivity (α = .71 at preand .68 at posttest).

SCBE. The SCBE consists of three subscales with 10 items each: Anxiety/Withdrawal (e.g., avoids new situations), Social Competence (e.g., cooperates with other children), and Anger/Aggression (e.g., easily frustrated). Teachers rated students across the three domains, using a six-point Likert-type scale (1 = never to 6 = always). Cronbach's alpha was .79 at preand .84 at posttest for Anxiety/Withdrawal, .85 at preand .87 at posttest for Social Competence, and .91 at preand .93 at posttest for Anger/Aggression.

Control Variables

Teachers' covariates considered in this study included years of teaching and educational attainment. Ethnicity was used as the child-level covariate, given prior research that has documented significant differences in academic achievement based on ethnicity (e.g., Chatterji, 2006; Davis-Kean & Jager, 2014; Reardon & Galindo, 2009).

Analysis Strategy

The analysis of covariance (ANCOVA) was conducted using SAS 9.4 to examine the intervention effects (research question 1) and coaching modality effects (research question 2) on end-of-year teacher outcome measures, statistically controlling for variability in beginning-of-the-year (baseline) scores as well as teacher-level covariates (years of teaching and educational attainment). The intervention effects were estimated by a contrast variable that was coded .5 for teachers in the intervention group (face-to-face coaching and remote coaching) and -1 for teachers in the business-as-usual group, while the coaching modality effects were estimated by the contrast variable that was coded -1 for teachers who received remote coaching, 1 for teachers who received face-to-face coaching, and 0 for teachers in the business-as-usual group. In addition, since data collection spanned three years, dummy indicators representing data collection waves to explore and control for the impact of data collection periods were included. The ANCOVA models were conducted using the SAS PROC MIXED procedure.

First, because teachers were nested in coaches, the dependence in teacher outcome scores was modeled by introducing a random effect in ANCOVA. As a result, the effects of interest were estimated taking into account teachers' dependence. The values of intraclass correlation coefficients for teacher outcome scores indicating teacher-level dependence (ranging from .000 to .055 for pretest scores; ranging from .000 to .177 for posttest scores) are reported in Appendix H. Second, additional analysis of the comparison between face-to-face coaching and remote coaching was conducted, taking into account teachers' responsiveness rated by coaches (i.e., teacher fidelity). Specifically, teacher responsiveness score was an average score of teachers' levels of coachability, mastery of the target activity, mastery of the course content, reflectiveness, engagement, and coaching support, which were domains capturing the extent to which teachers actively engaged with the TEEM intervention components. The descriptive statistics of teacher responsiveness scores are reported in Table 5. A posthoc analysis suggested that face-to-face teachers' responsiveness mean score was statistically significantly higher than remote teachers' ($p < .01$), implying teacher responsiveness varied by modality. We then controlled for teachers' responsiveness when examining coaching modality effects on teacher outcomes. We also examined interaction effects between coaching modality and teachers' responsiveness for the second research question. Third, we also conducted additional analyses to examine two interaction effects, namely modality*teaching experience and modality*educational

Table 8
Estimates of Mean Differences and Corresponding Effect Sizes on Teacher Assessments: Intervention Groups Versus BAU Group

	Pretest			Posttest		
Teacher outcomes	Mean difference (*SE*)	*p*-value	Effect size	Mean difference (*SE*)	*p*-value	Effect size
TBRS total score	0.21 (0.49)	.67	0.07	4.12 (0.47)	<.01	1.09
Classroom community	−0.03 (0.08)	.74	−0.06	0.27 (0.08)	<.01	0.53
Teacher sensitivity	0.02 (0.07)	.76	0.05	0.26 (0.07)	<.01	0.58
Learning centers	−0.05 (0.09)	.59	−0.09	0.60 (0.08)	<.01	0.95
Lesson plans[a]	0.01 (0.11)	.91	0.02	0.28 (0.12)	<.05	0.32
Book reading	0.05 (0.07)	.49	0.11	0.54 (0.08)	<.01	1.02
Print and letter knowledge	0.12 (0.06)	.04	0.31	0.54 (0.07)	<.01	1.10
Written expression	0.00 (0.06)	.97	0.01	0.46 (0.06)	<.01	0.92
Phonics	−0.01 (0.09)	.87	−0.03	0.48 (0.09)	<.01	0.76
Phonological awareness	0.06 (0.06)	.29	0.17	0.52 (0.07)	<.01	0.95
Oral language	0.01 (0.07)	.87	0.02	0.29 (0.08)	<.01	0.59
ELLCO total score	1.44 (1.70)	.40	0.14	10.90 (1.62)	<.01	0.94

Note. TBRS = Teacher Behavior Rating Scale- max score = 35.79; ELLCO = Early Language and Literacy Classroom Observation – Likert scale ranges from *1*-Deficient to *5*-Exemplary. Hedge's *g* was reported as a measure of effect size.
[a] Four teachers (2 in face-to-face and 2 in business-as-usual) did not have complete data in pretest or post scores of lesson plans and those teachers were not included in the analysis of lesson plans.

attainment. Results suggested these interaction effects were not statistically significant and thus were not included in our final analysis. Finally, teachers' educational attainment was examined as a moderator of coaching modality effects and no statistically significant results were found. Therefore, the interaction term between educational attainment and the contrast variable was not included in the analysis.

There were no missing values on teacher outcome variables except for one outcome - lesson plans. Only 2.30% of the teacher sample (4 of 174 teachers; 2 in face-to-face and 2 in business-as-usual) did not have complete data in pretest or post scores of lesson plans, and those teachers were not included in the analysis of lesson plans. Moreover, missing data on teacher covariates (years of teaching, and educational attainment) used in the ANCOVA models was minimal; approximately 2.87% of teacher samples (5 of 174 teachers) did not have full information on teacher covariates. Five teachers had missing values on years of teaching; the average year of teaching computed based on nonmissing cases was used to replace the missing values. Four teachers had missing information on educational attainment and the missing values were replaced with the mode value. To address research question 3, the structural equation modeling (SEM) was applied to investigate the direct and indirect effects of intervention effects/coaching delivery effects on child outcome measures via teacher outcome measures (e.g., contrast variables → teacher outcome scores → child outcome scores) using Mplus 8.4 (Muthén & Muthén, 1998). Both indirect and direct effects were estimated taking into account the impacts of teacher and child covariates as well as teacher's responsiveness scores. The model Indirect command was applied to request estimated indirect effects and corresponding results of statistically significant tests. The command Type = Complex in Mplus was used to take into account clustering of the children considering children being nested in the classroom. The values of intraclass correlation coefficients for child outcome scores indicating child-level dependence (ranging from .082 to .172 for pretest scores; ranging from .071 to .189 for posttest scores) are reported in Appendix H. Full information maximum likelihood (FIML) estimation in Mplus was used as a means of handling the missing data. For simplicity, only the statistically significant effects are reported even though all possible direct and indirect effects were tested.[2] The standardized effects were reported as a form of effect size (Preacher & Kelley, 2011). Although the theory of change posited indirect child effects via teacher changes, direct effects of the intervention/coaching delivery on child outcomes are also reported.

Results

Intervention Effects on Teachers' Language and Literacy Instruction

Table 7 presents descriptive statistics of the pretest and posttest TBRS total scores and subscores and ELLCO total scores for the full analytic teacher sample and for each experimental group. The ranges of TBRS and ELLCO scores were reported in Appendix F. Table 8 includes the results of estimated mean differences (Diff), corresponding standard errors and effect size Hedge's *g* for the pretest and posttest. The results showed at pretest, no statistically significant mean differences emerged between intervention group and business-as-usual group on the TBRS total score ($g = .07$), TBRS subscores (g ranging from $-.09$ to .31), and ELLCO total score ($g = .14$). The findings support the baseline equivalence in the analytical data.

Results showed that years of teaching was a statistically significant predictor of the TBRS learning centers posttest score ($p < .05$) and print and letter knowledge posttest score ($p < .05$). Teacher educational attainment was not a statistically significant predictor of any teacher outcomes. At posttest, results showed teachers who received the intervention (regardless of coaching modality) were rated higher on the TBRS total score (Diff = 4.12, $p < .01$, $g = 1.09$), including all TBRS subscores (Diff ranging from .26 to .60; g ranging from .32 to 1.10), as well as the ELLCO total score (Diff = 10.90, $p < .01$, $g = .94$).

[2] For information about the full set of tested direct and indirect effects, readers should contact the corresponding author.

Table 9
Estimates of Mean Differences and Corresponding Effect Sizes on Teacher Assessments: Face-to-Face Coaching Group Versus Remote Coaching Group

	Pretest			Posttest (without controlling for teachers' responsiveness)			Posttest (controlling for teachers' responsiveness)		
Teacher outcomes	Mean difference (*SE*)	*p*-value	Effect size	Mean difference (*SE*)	*p*-value	Effect size	Mean difference (*SE*)	*p*-value	Effect size
TBRS total score	0.76 (0.55)	.17	0.25	1.32 (0.66)	.05	.33	1.04 (0.69)	.14	0.26
Classroom community	0.10 (0.08)	.20	0.23	0.01 (0.08)	.96	.01	−0.04 (0.09)	.69	−0.07
Teacher sensitivity	0.05 (0.09)	.60	0.09	0.14 (0.08)	.06	.33	0.13 (0.08)	.12	0.29
Learning centers	0.22 (0.09)	.02	0.40	0.23 (0.11)	<.05	.35	0.16 (0.12)	.19	0.24
Lesson plans[a]	0.23 (0.12)	.07	0.33	0.01 (0.17)	.93	.02	−0.15 (0.17)	.39	−0.16
Book reading	−0.04 (0.08)	.60	−0.09	0.19 (0.10)	.06	.36	0.16 (0.11)	.14	0.31
Print and letter knowledge	0.08 (0.07)	.26	0.20	0.15 (0.09)	.09	.29	0.12 (0.09)	.19	0.24
Written expression	0.03 (0.06)	.62	0.09	0.30 (0.10)	<.01	.54	0.19 (0.10)	.06	0.34
Phonics	0.04 (0.09)	.66	0.08	0.10 (0.12)	.40	.14	0.10 (0.12)	.43	0.14
Phonological awareness	0.08 (0.07)	.22	0.22	0.19 (0.12)	.11	.32	0.23 (0.12)	.07	0.38
Oral language	0.03 (0.09)	.72	0.06	0.19 (0.08)	<.05	.40	0.15 (0.09)	.10	0.31
ELLCO total score	3.20 (1.85)	.09	0.31	2.43 (1.90)	.21	.21	0.93 (1.95)	.63	0.08

Note. TBRS = Teacher Behavior Rating Scale; ELLCO = Early Language and Literacy Classroom Observation. Hedge's *g* was reported as a measure of effect size.

Coaching Modality Effects on Teacher Outcomes

Table 9 presents the results of the effects of coaching modality (face-to-face v. remote) on teacher assessment scores at pretest and posttest. In our sensitivity analysis, teachers' responsiveness to the intervention was controlled for and these results are presented in the third column of Table 9. At pretest, there were no statistically significant mean differences between the two coaching modality groups on the TBRS total score (g = .25), TBRS subscores (g ranging from −.09 to .40), or ELLCO total score (g = .31). Years of teaching was a statistically significant predictor of the TBRS learning centers posttest score ($p < .05$), TBRS print and letter knowledge posttest score ($p < .05$), and ELLCO total posttest score ($p < .05$). Teacher educational attainment was not a statistically significant predictor of any teacher outcomes.

At posttest (without controlling for teachers' responsiveness), results showed that teachers who received face-to-face coaching had greater scores on some TBRS subscores including learning centers (Diff = .23, $p < .05$, g = .35), written expression (Diff = .30, $p < .01$, g = .54), and oral language (Diff = .19, $p < .05$, g = .40), compared to teachers who received remote coaching. In addition, the magnitudes of effect sizes suggested that teachers in the face-to-face coaching group outperformed teachers in the remote coaching group in terms of TBRS total scores (Diff = 1.32, p = .05, g = .33), teacher sensitivity (Diff = .14, p = .06, g = .33), book reading (Diff = .19, p = .06, g = .36), print and letter (Diff = .15, p = .09, g = .29), and phonological awareness (Diff = .19, p = .11, g = .32). However, after accounting for teachers' responsiveness, all of the statistically significant coaching modality effects on teachers' TBRS subscores disappeared and the magnitudes of effects also diminished. There were no statistically significant interaction effects of teachers' covariates or coaching modality.

Intervention Effects and Coaching Modality Effects on Child Outcomes

As mentioned in the Methods section, a SEM approach was used to investigate the indirect effects of intervention effects/coaching delivery effects on child outcome measures via the teacher outcome measures. Table 10 presents the statistically significant indirect

Table 10
Results of Indirect Effects: Intervention Effects on Child Outcomes Through Teachers Outcomes

Indirect effect paths (group contrast variable → teacher outcomes → child outcomes)									
Group contrast variable	Teacher outcomes	Child outcomes	Standardized indirect effect	*SE*	*p*-value	RMSEA	CLI	TLI	SRMR
Intervention vs. BAU	TBRS total score	Print Knowledge (TOPEL)	0.04	0.02	<.05	0.03	0.99	0.97	0.02
	Learning centers	EOWPVT	0.02	0.01	<.05	0.00	1.00	1.00	0.02
	Book reading behaviors	Print Knowledge (TOPEL)	0.03	0.02	<.05	0.00	1.00	1.00	0.02
	Print and letter knowledge	Auditory Comprehension (PLS)	0.03	0.01	<.05	0.02	1.00	0.98	0.02
		Print Knowledge (TOPEL)	0.04	0.02	<.05	0.04	0.99	0.95	0.03
	Written expression	Print Knowledge (TOPEL)	0.03	0.01	<.01	0.00	1.00	1.00	0.02
		Phonological Awareness (TOPEL)	0.03	0.01	<.05	0.02	0.99	0.98	0.02
	ELLCO total score	Print Knowledge (TOPEL)	0.03	0.01	<.05	0.03	0.99	0.97	0.03

Note. TOPEL = test of Preschool Early Literacy; EOWPVT = Expressive One-Word Picture Vocabulary Test; PLS = Preschool Language Scales-Fourth Edition.

effects of the intervention effects (intervention v. business-as-usual) on child outcome measures via teacher outcome measures. Model fit indices (e.g., root mean square error of approximation [RMSEA], comparative fit index [CFI], Tucker–Lewis Index [TLI]), and standardized root-mean-square residual (SRMR)] for each analysis are reported in Table 10. Based on the traditional cutoff values (RMSEA $< .06$; CFI and TLI fit indices $> .95$; SRMR $< .08$; Hu & Bentler, 1999), these fit indices suggested adequate model fit for our analysis. Child ethnicity was a statistically significant predictor of one child outcome. Hispanic/Latinx children had statistically lower EOWPVT scores compared to White children ($SE = -.06$; $p < .05$). Significant intervention effects on (a) teachers' TBRS total score, (b) TBRS subscores (learning centers, book reading behaviors, print and letter knowledge, and written expression), and (c) the ELLCO total score were found, which further related to changes in several child outcomes. For example, the first estimated indirect effect presented in the column of standardized indirect effect was .04 ($SE = .02$; $p < .01$), suggesting that the intervention effects on teachers' TBRS total score was positively associated with children's print knowledge scores. That is, teachers who received the intervention (regardless of coaching modality) outperformed teachers in the business-as-usual group on the TBRS total score, and the difference in TBRS total score was further associated with children's performance in print knowledge. The last estimated indirect effect was .03 ($SE = .01$; $p < .05$). This indirect effect suggested that teachers who received the intervention outperformed teachers in the business-as-usual group on the ELLCO total score, and the difference in ELLCO total score was further related to children's performance in print knowledge. Overall, intervention effects on teachers' outcomes were more likely to relate to children's print knowledge, phonological awareness, and auditory comprehension scores. Results suggested there were no statistically significant direct effects of the intervention on child outcomes, implying the intervention was only associated with child outcomes through changes in the teacher outcomes. It should be noted that the indirect effects explored in these analyses did not suggest causal relationships between intervention effects and child outcomes. In addition, no statistically significant indirect or direct effects of coaching modality on child outcomes were found after accounting for teachers' responsiveness scores. These findings were consistent with the previous nonsignificant findings of coaching modality effects on teacher-level outcomes.

Discussion

After Powell and colleagues' (2010) study, the current study is the second known experimental study to investigate the benefit of two different coaching modalities to teachers' instructional practices and child outcomes. Face-to-face and remote coaching models paired with intensive professional development for child care teachers were contrasted. The professional development was unique in that it consisted of three core features (a) course work around classroom management, language, literacy, and responsiveness, (b) training on conducting student progress monitoring, and (c) training on a supplemental language and literacy curriculum. Coaching was also provided in addition to the core intervention features. This approach to professional development is in contrast to other experimental studies of professional development models that have focused on those aspects in isolation rather than collectively (e.g., Pianta et al., 2008; Powell et al., 2010). Additionally, the study was among the first to examine professional development for child care teachers, who have been found to have lower levels of knowledge and skills compared to other early childhood contexts such as public schools and Head Start programs. Teachers were randomly assigned to one of three conditions (face-to-face, remote, and business-as-usual) to contrast coaching delivery modalities on teacher-level outcomes. Teachers in the study demonstrated medium levels of fidelity to the intervention across the intervention conditions.

Three key findings from the study emerged. First, both the face-to-face and remote coaching intervention groups demonstrated improved behaviors on the TBRS and ELLCO measures compared to the business-as-usual group. Second, there were some significant differences between teachers who received face-to-face and remote coaching on observational measures of instructional practices; however, when accounting for teachers' levels of responsiveness to the intervention, the differences between teachers who received face-to-face and remote coaching were no longer evident. Third, there were some positive indirect associations of the intervention on child-level outcomes through teachers' instructional practices.

TEEM and Teacher Outcomes

Findings from this study showed that teachers who participated in the intervention were observed as having stronger instructional practices compared to teachers in the business-as-usual group. All teachers were included in the analyses, regardless of their levels of engagement in the intervention. Generally, years of teaching experience, but not educational attainment, was a significant predictor of teacher-level outcomes. This is consistent with previous research that has documented on-the-job experience as an effective source of professional improvement (Kraft et al., 2018), but degree status as having limited associations with teaching quality (Early et al., 2017; NCTQ, 2005). There were no significant moderation effects for years of teaching experience or educational attainment on any teacher-level outcomes.

Coaching Modality and Teacher Outcomes

When comparing face-to-face and remote groups, there were several significant differences in teachers' observed instructional practices, favoring the face-to-face coaching group. This is somewhat in line with findings from other studies that have similarly compared face-to-face and remote coaching models. For example, Powell and colleagues (2010) found that teachers who received face-to-face coaching demonstrated larger gains in their code-focused instruction compared to teachers who received remote coaching. Similar to the Powell et al. (2010) study, the current study found that teachers who received face-to-face coaching outperformed teachers who received remote coaching on some, but not all, of the observed instructional practices that were measured (effect sizes ranging from .35 to .54). However, these differences disappeared when accounting for teachers' responsiveness to the intervention (e.g., engaged with their coach, demonstrated mastery of activities, reflected on their instructional practice), discussed further below.

A key goal of the current study was to hold constant coaching activities and professional development resources (to the extent

possible) across both coaching delivery modalities. That is, face-to-face and remote coaches were trained to use similar coaching strategies (e.g., helping teachers reflect, modeling) when working with teachers; however, there were coaching strategies that were not always transferable to remote coaching contexts (e.g., demonstration with teacher's own students, coteaching lessons). Coaches, regardless of coaching delivery modality, spent comparable amounts of time working with teachers (approximately 90 minute coaching sessions). Teachers in both coaching groups had access to the same sets of resources (e.g., video exemplars, course work). Thus, this study differed from and thereby extended the work of Powell and colleagues' (2010) study, in which researchers noted several key differences between the face-to-face and remote coaching groups. For example, in Powell and colleagues (2010) study, face-to-face coaches observed classroom activities for approximately 90 minutes each visit and debriefed with teachers for approximately 30 minutes. In the remote coaching condition, teachers submitted 15 minutes of a video-recorded instructional lesson and the coach provided annotated feedback, including direct links to video exemplars. Additionally, the researchers noted that only teachers who received remote coaching had ongoing access to hypermedia resources that showed video exemplars. In their study, it is difficult to determine whether the larger effects in face-to-face coaching on teachers' code-based instruction were attributed to the intensity of face-to-face coaching compared to remote coaching. By aligning coaching strategies when possible, the current study sought to examine the impacts of remote versus face-to-face coaching in a more controlled research design. Moreover, Powell and colleagues (2010) reported using a coaching protocol that was primarily based on observation and feedback (e.g., coaches observed specific practices and then provided written feedback). In the current study, face-to-face and remote coaches used a variety of strategies, in addition to observation and feedback (e.g., demonstration; side-by-side coaching).

It is worth noting that despite a controlled research design, there seemed to be inherent differences between face-to-face and remote coaching models. For example, face-to-face coaching models may allow coaches to view the classroom holistically; in contrast, remote coaching models, by nature, provide a more controlled or isolated aspect of classroom instruction. Face-to-face and remote coaching models may simply provide different stimuli that coaches attend to, thereby contributing to differences in what coaches respond to. Given that teachers in the intervention groups (regardless of coaching modality) showed improvements in instructional practices compared to the business-as-usual group, this suggests some form of professional development with coaching may be helpful for teachers. This finding is consistent with prior research that has compared coaching modalities and found that remote delivery may be a similarly effective way to provide coaching (e.g., Pianta et al., 2008; Powell et al., 2010; Vernon-Feagans et al., 2018). For example, Vernon-Feagans et al. (2018) has similarly found that remote coaching delivery of a literacy-based professional development model was effective in both improving teacher and child outcomes in samples of kindergarten and first grade classrooms. Remote delivery paired with online professional development, particularly when it is linked to specific quality metrics, has potential to be a scalable way to improve the quality of child care programs (Pianta et al., 2008). Taken together, these findings suggest that the quality and content of coaching may be more important contributors to improvements in teachers' practices (Matsumura et al., 2019). In the current study, sensitivity analyses also showed that when accounting for teachers' responsiveness to the intervention (e.g., participation, mastery of content/activities, engagement), face-to-face and remote coaching models were comparable in improving outcomes. Other researchers have similarly recognized teachers' responsiveness in coaching-based professional development as a way to optimize resources and obtain desired outcomes (e.g., Downer et al., 2009; Hindman et al., 2015; LoCasale-Crouch et al., 2016). Researchers have hypothesized that coach-teacher dynamics can boost (or hinder) the degree to which teachers are perceived as responsive (Cox, 2015; Downer et al., 2009; LoCasale-Crouch et al., 2016).

It was originally hypothesized that face-to-face coaching models had many theoretical benefits compared to remote coaching models. For example, technology was hypothesized to be a barrier to strong coach-teacher relationships (i.e., concept of relatedness from self-determination theory) and opportunities that allowed coaches to break down complex instructional practices into discrete steps (cognitive theory). It was also hypothesized that face-to-face coaches would be better able to help teachers manage competing classroom demands because they could readily use strategies such as side-by-side coaching or in-the-moment feedback. Findings showed that although teachers in the face-to-face coaching group were rated as more responsiveness compared to teachers in the remote coaching group, there were no significant interaction effects of responsiveness and modality on teacher outcomes. This suggests that the impact of responsiveness on teacher outcomes was consistent in both face-to-face and remote coaching groups. That is, teachers' responsiveness, regardless of coaching modality, appeared to minimize the theoretical advantages that face-to-face coaching models had over remote coaching models. This finding highlights the importance of the practical tenets of situated learning theory that were embedded within both coaching models. Teachers in the face-to-face and remote coaching groups both had opportunities to socially construct knowledge with guidance from a coach, reflect on their teaching, and practice new instructional strategies multiple times.

The minimal differences between the two coaching modalities are promising, especially when considering factors such as access to coaching support (e.g., especially in rural geographic locations) and costs of face-to-face versus remote coaching models (e.g., mileage reimbursement, between school travel time). Interventions and implementation plans that explicitly design support for increasing teachers' responsiveness may be important, especially in remote coaching models, where it may be more difficult for teachers to engage with their coaches.

Influence on Child Outcomes

The results of the current study also highlight small indirect effects of the intervention on children's literacy and language and no impacts on children's social-behavioral outcomes. The lack of indirect effects on children's social-behavioral outcomes may have been attributed to the moderate effects that the intervention had on related teacher instructional practices (e.g., TBRS subscales such as teachers' sensitivity, classroom community). However, children's phonological awareness, print knowledge, and auditory comprehension skills seemed to be most consistently,

albeit indirectly, associated with intervention effects through changes in key teacher practices (e.g., book reading behaviors, print and letter knowledge, written expression, and total ELLCO score). This finding is consistent with previous research on early childhood teachers' language and literacy instructional practices as a mediator that predicts children's outcomes (Piasta et al., 2020) as well as research that has posited professional development with coaching to have more proximal impacts on teacher practices, which thereby influence child outcomes (Kraft et al., 2018). In the current study, coaches directly worked with teachers to improve these types of skills as evidenced by the frequency of focal topic areas such as shared book reading, phonological awareness and phonics, letter knowledge, and written expressions, which were noted in coaching logs. This focus likely helped teachers to integrate knowledge from the professional development content (e.g., online courses, student progress monitoring) within their practices. Responsive interactions within the context of language and literacy were also a central focus of the TEEM intervention and coaches. The online courses and resource library emphasized, for example, the use of scaffolding in order to appropriately match teachers' responses to children's verbal and nonverbal signals. Improvements in those key teacher practices likely supported children's early literacy skills in areas such as phonological awareness and print knowledge. Teachers who were rated higher in those areas may have been more likely to provide opportunities that supported the development of early literacy skills. For example, teachers who were rated as demonstrating stronger book reading behaviors may have provided more developmentally appropriate opportunities for children to attend to the sounds of words or specific reading conventions (e.g., page orientation, connecting print to sounds).

The limited child effects differ from past studies of the TEEM model that found moderate to large effect sizes across children's early language and literacy skills (e.g., language comprehension, phonological awareness, expressive vocabulary, print and letter knowledge). It is hypothesized that the differences in child effects were attributed to the quality of instruction (as measured by TBRS) in child care programs, where the quality of instruction was descriptively lower compared to other contexts (e.g., Head Start, school districts) included in previous TEEM studies (e.g., Landry et al., 2011). Previous research also suggests that implementation fidelity (dosage/exposure, adherence, and quality) has been associated with preschool students' gains in learning outcomes during the year (Hamre et al., 2017; Marti et al., 2018). In the current study, teachers in both intervention groups demonstrated relatively moderate fidelity scores across categories (e.g., mastery of target activity, mastery of course content), which may have also contributed to the inconsistent findings on child outcomes. On average, TEEM teachers showed relatively high levels of participation in the intervention (i.e., course completion) and with coaching. Despite this dosage, coaches reported, for example, that both face-to-face and remote teachers only demonstrated a medium level of course content and target activity mastery. Several reasons for medium levels of mastery over course content and target activity are possible. First, many child care teachers in the current sample had limited training in early childhood education and lower levels of education (e.g., 25% had an Associates degree or higher, 4% had CDA certification). Thus, one year of PD support and online courses may not be sufficient in improving teachers' abilities to fully internalize content from PD courses and implement activities with high levels of mastery. Other studies have found promising effects for two-year professional development programs (e.g., Landry et al., 2006; Vernon-Feagans et al., 2018; Taylor et al., 2005). Findings of a previous study on TEEM suggested that two-year participation in TEEM led to greater gains for children (Landry et al., 2006). Changes in teachers' practices typically occur in small increments and likely require sustained efforts over longer periods of time (Taylor et al., 2005). Second, the moderate quality ratings on other fidelity indicators such as reflectiveness and coachability, suggest that more robust coach and teacher trainings are needed. There are limited training and PD opportunities for coaches to learn effective coaching skills that support teachers' abilities to reflect on instructional practices. Coaches may have needed more support on how to help teachers to be more reflective in their practice.

Overall, this study, like others, found stronger effects for professional development with coaching on teacher outcomes compared to child outcomes (e.g., Kraft et al., 2018; Pianta et al., 2017). Additional analyses suggested that although there were no direct effects of the intervention on child-level outcomes, there were indirect effects of the intervention for child outcomes. This is consistent with hypothesized theories of change for professional development with coaching. That is, researchers have hypothesized that coaches typically work with teachers to improve their instructional practices and knowledge (Kraft et al., 2018; Sheridan et al., 2008). The subsequent improvements in teacher practices and knowledge typically lead to gains in child outcomes. In the current intervention, both face-to-face and remote coaches worked directly with teachers as teachers implemented goals aligned with the course content. It is possible that in the context of child-focused interventions (e.g., teacher targets specific children throughout the year to implement an intervention), coaching may have direct effects on child outcomes (e.g., Vernon-Feagans et al., 2018). On the other hand, less child-specific and more generalized coaching feedback (as used in the current study) may not always translate immediately or directly to child outcomes. It may take longer periods of time to change teacher behaviors and practices, which may take longer to detect effects on child-level outcomes.

Limitations

This study was bound by several limitations. First, despite efforts to minimize levels of attrition (e.g., meeting and explaining study protocol prior to randomization, commitment letters from principals and teachers), there was substantial attrition across the study groups. Although the differential attrition ranged from 6–17%, which still meets the What Works Clearinghouse (2020) standards, it was challenging to work in child care centers where there were higher levels of teacher turnover. Approximately 47% left for reasons unrelated to participation in the intervention, such as job termination or other job opportunities. The high rates of teacher turnover in the current study are consistent with turnover rates in child care settings across the nation. Although providing professional development with coaching for child care teachers continues to be important, it is not without challenges. In the current sample of child care teachers, many did not have high levels of education or training in early childhood education (e.g., only 25% had an Associates degree or higher and only 4% reported having CDA certification). This underscores an important need for providing professional development to this population of the early childhood workforce. Second, the coaches who worked with child

care teachers were employed by the university. It is unclear how these findings generalize to other models of professional development with coaching that use center- or school-employed coaches. Third, utilizing remote coaching models was somewhat challenging across child care programs. For some teachers, it was more challenging to learn how to upload videos and learn how to use the online platform; for these teachers, more time was spent early on helping them learn and use technology. Occasional Internet connectivity and bandwidth issues posed challenges for some teachers, particularly those in more rural locales.

Additionally, there was limited coach fidelity data that was collected in the current study. Expanding the current coach fidelity rating tool to include other ways to measure coaches' adherence to the TEEM coaching model can help schools can sustain and monitor TEEM implementation. There were limitations in the teacher fidelity data that was collected. Data collected on teachers' responsiveness to TEEM suggested relatively moderate levels of responsiveness (ranged from 2.58 to 4.11). Some teachers were not considered fully responsive to the intervention, but the reasons for teachers' lack of responsiveness were not collected. Collecting this type of information can be especially important for studies that specifically target teachers for professional development. It is also important to note that although there were trainings and ongoing monitoring to help coach supervisors and coaches to use the coach and teacher fidelity rating forms (respectively), there was no interrater reliability data collected for these fidelity rating forms.

In addition, although data was collected on teachers' instructional practices across conditions, no data was collected on the professional development and coaching opportunities available to teachers in the business-as-usual group. This would have allowed for a more comprehensive understanding of comparable practices between teachers in the intervention and business-as-usual groups and may have provided more information about why and/or how the TEEM professional development model yielded positive teacher- and child-level effects. Future studies should account for the types of professional development and coaching opportunities available to teachers in the business-as-usual condition. Moreover, although previous studies of TEEM have shown that TEEM + coaching yielded the strongest benefits to teachers and children compared to other contrasts (e.g., teachers who received TEEM without coaching; Landry et al., 2009), we were unable to isolate coaching support from the TEEM intervention components in the three groups included in the current study. In future studies, researchers may consider adding another contrast (e.g., TEEM without any coaching) in addition to a business-as-usual group.

Future Research

Three key areas for future research are warranted. First, as professional development with remote coaching is increasingly used, more research about the efficacy of remote coaching models compared to face-to-face coaching models are needed. Remote coaching models are thought to be more cost-effective compared to face-to-face coaching models, but there is limited research using cost effectiveness analyses that support these hypotheses. In future studies of TEEM, the cost of each intervention delivery modality will be analyzed and compared using cost effectiveness to understand whether remote coaching models reduce costs of coaching. Second, researchers may consider research designs in which they are able to work with teachers for longer periods of time (e.g., two years) and measure child outcomes after teachers have more time to internalize and apply knowledge from working with their coaches. Teachers may also require more time to fully acclimate to the professional development model, as they likely have stronger uptake toward the middle and end of the school year. That is, it may take teachers more time to fully "buy into" a new professional development model. Building in longer periods of time that allow teachers more time to learn and apply content of professional development may translate to stronger child outcomes. Third, this study examined controlled delivery of intervention by a research team. In routine practice, child care quality improvement efforts are more often delivered by Quality Rating and Improvement Systems (which typically provide information about program quality and incentives and resources to improve or maintain quality) or technical assistance specialists. Research is needed to determine the extent to which the coaching approaches (e.g., dose, coaching strategies) tested in this study are feasible and can be implemented with fidelity under more routine conditions (e.g., state professional development systems).

Conclusion

Professional development with coaching continues to be widely adopted across early childhood settings. Findings from the current study add to the existing knowledge base by focusing on the use of professional development with coaching for child care teachers, who are an important, but often overlooked, subset of the early childhood workforce that is especially vulnerable to job stressors and burnout. Additionally, the current study also contrasted two specific types of coaching modalities, face-to-face and remote coaching, in relation to teacher and child outcomes. Findings of the current study suggest that there were some key differences between the two coaching modalities on teacher outcomes (favoring the face-to-face coaching group), but that after accounting for teachers' levels of responsiveness, there were minimal differences between coaching modalities. Results also add empirical evidence to a growing body of research that supports the use of professional development with coaching, particularly for improving teachers' instructional practices. Child care teachers who participated in the professional development with coaching showed improvements in their instructional practices, suggesting that these kinds of opportunities may be beneficial to improving the quality of child care classrooms. This study also extends current research in this area by providing empirical support for what researchers have theorized (Kraft et al., 2018): professional development with coaching influences child outcomes through improvements in teachers' instructional practices. Collectively, the study findings highlight the promise of professional development with coaching and point to the need for making these supports more accessible for early childhood teachers across settings.

References

Brownell, R. (Ed.). (2000). *Expressive one-word picture vocabulary test: Manual*. Academic Therapy Publications.

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Prentice Hall.

Bassok, D., Fitzpatrick, M., Greenberg, E., & Loeb, S. (2016). Within-and between-sector quality differences in early childhood education and care. *Child Development*, *87*(5), 1627–1645. https://doi.org/10.1111/cdev.12551

Bates, C. C., McClure, E. L., Ross, R. L., & Womack, P. (2019). Web-Mediated Professional Development. *Journal of Digital Learning in Teacher Education*, *35*(1), 40–53. https://doi.org/10.1080/21532974.2018.1537817

Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science*, *12*(1), 23–33. https://doi.org/10.1007/s11121-010-0186-1

Cabell, S. Q., & Downer, J. T. (2011). Improving preschoolers' language and literacy skills through web-mediated professional development. *NHSA Dialog*, *14*(4), 316–322. https://doi.org/10.1080/15240754.2011.613130

Cassidy, D. J., Lower, J. K., Kintner-Duffy, V. L., Hegde, A. V., & Shim, J. (2011). The day-to-day reality of teacher turnover in preschool classrooms: An analysis of classroom context and teacher, director, and parent perspectives. *Journal of Research in Childhood Education*, *25*(1), 1–23. https://doi.org/10.1080/02568543.2011.533118

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology*, *98*(3), 489–507. https://doi.org/10.1037/0022-0663.98.3.489

Cox, E. (2015). Coaching and adult learning: Theory and practice. *New Directions for Adult and Continuing Education*, *2015*(148), 27–38. https://doi.org/10.1002/ace.20149

Crawford, A., Zucker, T., Van Horne, B., & Landry, S. (2017). Integrating professional development content and formative assessment with the coaching process: The Texas school ready model. *Theory into Practice*, *56*(1), 56–65. https://doi.org/10.1080/00405841.2016.1241945

Davis-Kean, P. E., & Jager, J. (2014). Trajectories of achievement within race/ethnicity:"Catching up" in achievement across time. *The Journal of Educational Research*, *107*(3), 197–208. https://doi.org/10.1080/00220671.2013.807493

Downer, J. T., Kraft-Sayre, M. E., & Pianta, R. C. (2009). Ongoing, web-mediated professional development focused on teacher–child interactions: Early childhood educators' usage rates and self-reported satisfaction. *Early Education and Development*, *20*(2), 321–345. https://doi.org/10.1080/10409280802595425

Early, D. M., Maxwell, K. L., Ponder, B. B., & Pan, Y. (2017). Improving teacher-child interactions: A randomized control trial of Making the Most of Classroom Interactions and My Teaching Partner professional development models. *Early Childhood Research Quarterly*, *38*, 57–70. https://doi.org/10.1016/j.ecresq.2016.08.005

Fishman, B., Konstantopoulos, S., Kubitskey, B. W., Vath, R., Park, G., Johnson, H., & Edelson, D. C. (2013). Comparing the impact of online and face-to-face professional development in the context of curriculum implementation. *Journal of Teacher Education*, *64*(5), 426–438. https://doi.org/10.1177/0022487113494413

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

Hamre, B. K., Partee, A., & Mulcahy, C. (2017). Enhancing the impact of professional development in the context of preschool expansion. *AERA Open*, *3*(4), 2332858417733686. https://doi.org/10.1177/2332858417733686

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*(3), 133–146. https://doi.org/10.1177/0271121416653386

Hindman, A. H., Snell, E. K., Wasik, B. A., Lewis, K. N., Hammer, C. S., & Iannone-Campbell, C. (2015). Research and practice partnerships for professional development in early childhood: Lessons from ExCELL-e. *Journal of Education for Students Placed at Risk*, *20*, 12–28. https://doi.org/10.1080/10824669.2014.984036

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jamison, K. R., Cabell, S. Q., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. C. (2014). CLASS–Infant: An observational measure for assessing teacher–infant interactions in center-based child care. *Early Education and Development*, *25*(4), 553–572. https://doi.org/10.1080/10409289.2013.822239

Justice, L. M., Mashburn, A., Hamre, B., & Pianta, R. (2008). Quality of Language and Literacy Instruction in Preschool Classrooms Serving At-Risk Pupils. *Early Childhood Research Quarterly*, *23*(1), 51–68. https://doi.org/10.1016/j.ecresq.2007.09.004

Kane, B. D., & Rosenquist, B. (2019). Relationships between instructional coaches' time use and district-and school-level policies and expectations. *American Educational Research Journal*, *56*(5), 1718–1768. https://doi.org/10.3102/0002831219826580

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

LaFreniere, P. J., & Dumas, J. E. (1996). Social competence and behavior evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment*, *8*(4), 369–377. https://doi.org/10.1037/1040-3590.8.4.369

Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, *101*(2), 448–465. https://doi.org/10.1037/a0013842

Landry, S. H., Crawford, A., Gunnewig, S., & Swank, P. R. (2001). *Teacher Behavior Rating Scale* [Unpublished manuscript]. Center for Improving the Readiness of Children for Learning and Education, University of Texas Health Science Center at Houston.

Landry, S. H., Smith, K. E., Swank, P. R., Assel, M. A., & Vellet, S. (2001). Does early responsive parenting have a special importance for children's development or is consistency across early childhood necessary? *Developmental Psychology*, *37*(3), 387–403. https://doi.org/10.1037/0012-1649.37.3.387

Landry, S. H., Swank, P. R., Anthony, J. L., & Assel, M. A. (2011). An experimental study evaluating professional development activities within a state funded pre-kindergarten program. *Reading and Writing*, *24*(8), 971–1010. https://doi.org/10.1007/s11145-010-9243-1

Landry, S. H., Swank, P. R., Smith, K. E., Assel, M. A., & Gunnewig, S. B. (2006). Enhancing early literacy skills for preschool children: Bringing a professional development model to scale. *Journal of Learning Disabilities*, *39*(4), 306–324. https://doi.org/10.1177/00222194060390040501

Lemons, C. J., & Toste, J. R. (2019). Professional Development and Coaching: Addressing the "Last Mile" Problem in Educational Research. *Assessment for Effective Intervention*, *44*(4), 300–304. https://doi.org/10.1177/1534508419862859

Lin, Y. C., & Magnuson, K. A. (2018). Classroom quality and children's academic skills in child care centers: Understanding the role of teacher qualifications. *Early Childhood Research Quarterly*, *42*, 215–227. https://doi.org/10.1016/j.ecresq.2017.10.003

LoCasale-Crouch, J., DeCoster, J., Cabell, S. Q., Pianta, R. C., Hamre, B. K., Downer, J. T., Hatfield, B. E., Larsen, R., Burchinal, M., Howes, C., LaParo, K., Scott-Little, C., & Roberts, A. (2016). Unpacking intervention effects: Teacher responsiveness as a mediator of perceived intervention quality and change in teaching practice. *Early Childhood Research Quarterly*, *36*, 201–209. https://doi.org/10.1016/j.ecresq.2015.12.022

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *TOPEL: Test of preschool early literacy.* Pro-Ed.

Marti, M., Melvin, S., Noble, K. G., & Duch, H. (2018). Intervention fidelity of getting ready for school: Associations with classroom and teacher characteristics and preschooler's school readiness skills. *Early Childhood Research Quarterly*, *44*, 55–71. https://doi.org/10.1016/j.ecresq.2018.02.010

Matsumura, L. C., Correnti, R., Walsh, M., Bickel, D. D., & Zook-Howell, D. (2019). Online content-focused coaching to improve classroom discussion quality. *Technology, Pedagogy and Education*, *28*(2), 191–215. https://doi.org/10.1080/1475939X.2019.1577748

McMullen, M. B., Lee, M. S., McCormick, K. I., & Choi, J. (2020). Early Childhood Professional Well-Being as a Predictor of the Risk of Turnover in Child Care: A Matter of Quality. *Journal of Research in Childhood Education*, *34*, 1–15. https://doi.org/10.1080/02568543.2019.1705446

Muthén, L. K., & Muthén, B. O. (1998). Mplus [Computer software].

National Survey of Early Care and Education Project Team. (2013). *Number and characteristics of early care and education (ECE) teachers and caregivers: Initial findings from the National Survey of Early Care and Education* (Report #2013–38). U.S. Department of Health and Human Services Office of Planning, Research and Evaluation, Administration for Children and Families.

Pianta, R., Hamre, B., Downer, J., Burchinal, M., Williford, A., Locasale-Crouch, J., Howes, C., LaParo, K., & Scott-Little, C. (2017). Early childhood professional development: Coaching and coursework effects on indicators of children's school readiness. *Early Education and Development*, *28*(8), 956–975. https://doi.org/10.1080/10409289.2017.1319783

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediatedprofessional development resources on teacher–child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, *23*(4), 431–451. https://doi.org/10.1016/j.ecresq.2008.02.001

Piasta, S. B., Farley, K. S., Mauck, S. A., Soto Ramirez, P., Schachter, R. E., O'Connell, A. A., Justice, L. M., Spear, C. F., & Weber-Mayrer, M. (2020). At-scale, state-sponsored language and literacy professional development: Impacts on early childhood classroom practices and children's outcomes. *Journal of Educational Psychology*, *112*(2), 329–343. https://doi.org/10.1037/edu0000380

Poglinco, S. M., & Bach, A. J. (2004). The heart of the matter: Coaching as a vehicle for professional development. *Phi Delta Kappan*, *85*(5), 398–400. https://doi.org/10.1177/003172170408500514

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacyprofessional development intervention on head start teachers and children. *Journal of Educational Psychology*, *102*(2), 299–312. https://doi.org/10.1037/a0017763

Powell, D. R., Steed, E. A., & Diamond, K. E. (2009). Dimensions of Literacy Coaching With Head Start Teachers. *Topics in Early Childhood Special Education*, *30*(3), 148–161. https://doi.org/10.1177/0271121409341197

Powell, D. R., & Diamond, K. E. (2013). Implementation fidelity of a coaching-based professional development program for improving Head Start teachers' literacy and language instruction. *Journal of Early Intervention*, *35*(2), 102–128. https://doi.org/10.1177/1053815113516678

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*(2), 93–115. https://doi.org/10.1037/a0022658

Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, *46*(3), 853–891. https://doi.org/10.3102/0002831209333184

Reinke, W. M., Stormont, M., Herman, K. C., Wang, Z., Newcomer, L., & King, K. (2014). Use of coaching and behavior support planning for students with disruptive behavior within a universal classroom management program. *Journal of Emotional and Behavioral Disorders*, *22*(2), 74–82. https://doi.org/10.1177/1063426613519820

Rock, M. L., Gregg, M., Thead, B. K., Acker, S. E., Gable, R. A., & Zigmond, N. P. (2009). Can you hear me now? Evaluation of an online wireless technology to provide real-time feedback to special education teachers-in-training. *Teacher Education and Special Education*, *32*(1), 64–82. https://doi.org/10.1177/0888406408330872

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development*, *72*(5), 1394–1408. https://doi.org/10.1111/1467-8624.00355

Sheridan, S. M., Marvin, C. A., Knoche, L. L., & Edwards, C. P. (2008). Getting ready: Promoting school readiness through a relationship-based partnership model. *An Interdisciplinary Journal of Effectiveness*, *2*, 149–172.

Sibley, A., & Sewell, K. (2011). Can multidimensional professional development improve language andliteracy instruction for young children? *NHSA Dialog*, *14*(4), 263–274. https://doi.org/10.1080/15240754.2011.609948

Smith, M., Dickinson, D., Sangeorge, A., & Anastasopoulos, L. (2002). *Early literacy and language classroom observation scale (ELLCO).* Paul Brookes.

Snyder, P., Hemmeter, M. L., McLean, M., Sandall, S., McLaughlin, T., & Algina, J. (2018). Effects of professional development on preschool teachers' use of embedded instruction practices. *Exceptional Children*, *84*(2), 213–232. https://doi.org/10.1177/0014402917735512

Sun, J., Rao, N., & Pearson, E. (2015). *Policies and strategies to enhance the quality of early childhood educators*. http://education2030-africa.org/images/talent/Atelier_melqo/Policies-and-Strategies-to-Enhance-the-Quality-of-Early-Childhood-Educators.pdf

Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2005). The CIERA school change framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly*, *40*(1), 40–69. https://doi.org/10.1598/RRQ.40.1.3

Varghese, C., Garwood, J. D., Bratsch-Hines, M., & Vernon-Feagans, L. (2016). Exploring magnitude of change in teacher efficacy and implications for students' literacy growth. *Teaching and Teacher Education*, *55*, 228–239. https://doi.org/10.1016/j.tate.2016.01.011

Vernon-Feagans, L., Bratsch-Hines, M., Varghese, C., Bean, A., & Hedrick, A. (2015). The Targeted Reading Intervention: Face-to-face vs. webcam literacy coaching of classroom teachers. *Learning Disabilities Research & Practice*, *30*(3), 135–147. https://doi.org/10.1111/ldrp.12062

Vernon-Feagans, L., Bratsch-Hines, M., Varghese, C., Cutrer, E. A., & Garwood, J. D. (2018). Improving struggling readers' early literacy skills through a Tier 2 professional development program for rural classroom teachers: The targeted reading intervention. *The Elementary School Journal*, *118*(4), 525–548. https://doi.org/10.1086/697491

What Works Clearinghouse. (2020). *What Works Clearinghouse Standards Handbook* (Version 4.1). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Whitebook, M., & Sakai, L. (2003). Turnover begets turnover: An examination of job and occupational instability among child care center staff. *Early Childhood Research Quarterly*, *18*(3), 273–293. https://doi.org/10.1016/S0885-2006(03)00040-1

Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. https://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/rel_2007033.pdf

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., & Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. Society for Research in Child Development and Foundation for Child Development.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale* (4th ed.). The Psychological Corporation.

Received March 30, 2020
Revision received April 14, 2021
Accepted April 23, 2021 ■