

# Effects of an Online Module to Improve Preschool Teachers' Progress Monitoring Implementation and Abilities

**Collin Shepley\***, University of Kentucky, Department of Early Childhood, Special Education, and Counselor Education, 229 Taylor Education Building, Lexington, Kentucky, 40506, US, [collinshepley@uky.edu](mailto:collinshepley@uky.edu)

**Amanda Leigh Duncan**, University of Kentucky, Department of Early Childhood, Special Education, and Counselor Education, 229 Taylor Education Building, Lexington, Kentucky, 40506, US, [aduncan@uky.edu](mailto:aduncan@uky.edu)

**Emily Webb**, University of Kentucky, Department of Early Childhood, Special Education, and Counselor Education, 229 Taylor Education Building, Lexington, Kentucky, 40506, US, [ee webb0@uky.edu](mailto:ee webb0@uky.edu)

\*Corresponding author

## ABSTRACT

For children receiving special education services, the provision of progress monitoring is recommended by professional organizations, required by federal policies, and supported by research. However, research-backed training and professional development focused on progress monitoring for early childhood professionals is largely missing from the research literature. Therefore, we evaluated the impact of an online professional development module called *Progress Monitoring for Preschool Teachers*. Confirmatory outcomes measured teacher ability and implementation quality, and exploratory outcomes measured teacher self-efficacy, classroom quality, and children's academic achievement. Results indicated improvements in teachers' progress monitoring abilities and

**Submitted:** 25 November 2024

**Accepted:** 18 August 2025

**Published:** 06 October 2025

*Research in Special Education* is a peer-reviewed open access journal published by the Aperio. © 2025 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which enables reusers to distribute, remix, adapt, and build upon the material in any medium or format for noncommercial purposes only, and only so long as attribution is given to the creator. See <https://creativecommons.org/licenses/by-nc/4.0/>.



3 OPEN ACCESS

implementation following completion of the module relative to teachers who relied solely on their district-provided professional development opportunities. Noteworthy changes were not observed for exploratory outcomes focused on classroom quality nor children's academic achievement. Limitations of the findings and considerations for districts adopting the module for professional development purposes are provided.

**Keywords:** progress monitoring; professional development; preschool; data collection; module

For children with disabilities receiving special education services, the provision of progress monitoring by educators and related services providers is an essential and legally required practice as specified within a child's Individualized Education Program (IEP). Even for children without a diagnosed disability, but who may be at risk for developmental delay and receiving early educational services within a multi-tiered system of support, progress monitoring is recognized as essential to the implementation of these systems (Division for Early Childhood of the Council for Exceptional Children, 2021). Recommended progress monitoring practices in early education for supporting children with or at risk for disability, emphasize that data be (a) collected through the on-going and objective measurement of operationally defined skills and behaviors (McLean et al., 2004; Ledford et al., 2019) and (b) analyzed using low-inference methods that support practical instructional decisions in a timely manner (Bishop et al., 2019).

For decades, statutes surrounding progress monitoring conveyed that the practice be used to support documentation and reporting efforts of teachers and schools (Etscheidt, 2006). However, the legal definition of progress monitoring has undergone a drastic change following the U.S. Supreme Court's decision in *Endrew F. v. Douglas County Schools* (2017). The Supreme Court found that monitoring child progress solely for documentation and reporting purposes was insufficient for meeting a school's burden of providing a free and appropriate public education (Yell & Bateman, 2017). In their ruling, the Supreme Court stressed that progress monitoring should inform the instruction that a child receives (Yell & Bateman, 2019). Simply put, when progress monitoring data indicate that current educational services are failing to result in meaningful progress for a child, then changes to those educational services should be identified and implemented.

The Supreme Court's framing of progress monitoring aligns with the research literature on the topic. One of the most consistent findings across school-based research is that when teachers use progress monitoring to inform their instruction, it contributes to improvements in child development and student outcomes (Black & William, 1998; Foegen et al., 2007; Fuchs & Fuchs, 1986; Graham et al., 2015; Kingston & Nash, 2011, 1986; Lee et al., 2020). One might assume that the teacher training and professional development literature reflect the legally mandated and research-backed role that progress monitoring plays in supporting children's education. However, in exploring this literature, it appears that research on training and preparing early childhood educators is almost exclusively focused on implementing instructional strategies and behavioral interventions (Artman-Meeker et al., 2015; Egert et al., 2018; Elek & Page, 2019). Missing from this corpus of research are considerations of progress monitoring (Shepley et al., 2023). This is a potentially worrisome omission due to the timely findings from numerous research teams reporting that the latest-and-greatest educational interventions are often not as informative nor effective as generally thought to be (Courtade et al., 2014; Kraft, 2020; Kraft, 2023; Lortie-Forgues & Inglis, 2019; Odom, 2021). Thus, if teachers are trained (or required) to implement a newly developed educational intervention, it is critical that they are also trained (and required) to monitor child progress to determine if such interventions are actually helping their children learn.

Regarding the components of effective professional development commonly utilized in early childhood settings, performance-based feedback is consistently identified as essential to supporting teacher acquisition of new practices (Artman-Meeker et al., 2015; Elek & Page, 2019). Contemporary models of professional development embed the

provision of performance-based feedback within coaching frameworks that require repeated one-on-one contacts between a coach and a teacher (Ledford et al., 2019; Snyder et al., 2015). Though generally effective, this type of coaching can be resource intensive (Brock et al., 2017; Schachter et al., 2024), particularly for rural districts (Lang et al., 2024), necessitating the use of alternative professional development methods that deliver training content through asynchronous mediums (e.g., video recordings of instructional practices; text-based feedback; McLeod et al., 2024). It should be noted that there are minimal rigorous evaluations of asynchronous professional development tools for supporting early childhood educators (cf., Snyder et al., 2018); however, within adjacent fields like special education and behavioral sciences, there is a growing body of literature on using asynchronous methods for training practitioners (Jiminez et al., 2016; Marano et al., 2020).

Effective asynchronous trainings often use interactive online modules with embedded competency checks (Gerenscer et al., 2020). Competency checks function as an equivalent to the provision of performance-based feedback (Pollard et al., 2014), as differentiated information can be presented to a user based on responses to questions presented during a competency check (Cheek et al., 2019). On the effectiveness of asynchronous tools for supporting progress monitoring, Higbee and colleagues (2016) found that the sole use of an online module was sufficient to improve the data collection abilities of pre- and in-service special educators; that is, no coaching or follow-up feedback from a trainer was needed. As it relates to early childhood education, asynchronous mediums for delivering teacher training may offer promise across schools and districts where contemporary models of professional development can be resource prohibitive.

To address the need for research on professional development to support teachers' progress monitoring efforts, we evaluated an asynchronous online training module and its impact across teacher, classroom, and child outcomes. The module focused exclusively on practices and strategies for monitoring the progress of children with individualized needs, such those with or at risk for disability or developmental delay. This module was specific to monitoring child progress on skills related to literacy, mathematics, and cognitive concepts. Example topics addressed within the module included (a) selecting meaningful skills, (b) designing data sheets, (c) collecting data, (d) analyzing data, and (e) making data-based decisions. The module was divided into six sections, with each section presenting a video conveying target content, follow-up questions, and competency checks which presented dynamic feedback (i.e., the specific feedback varied based on a user's responses to the follow-up questions).

The presentation of content within the module was designed to align with the components of a research-supported adult training model, called behavioral skills training (BST; Brock et al., 2019). When using the BST model to train an adult to complete a task or perform a specific skill, a trainer (a) provides instructions about the target skill, (b) models what the skill should like when performed correctly, (c) allows the adult to practice performing the skill through a guided rehearsal, and (d) offers performance-based feedback. Within the module, embedded videos functioned as the instruction and modeling components of BST, the follow up questions functioned as the rehearsal component, and the competency checks functioned as the performance-based feedback. The module can be freely accessed by the public at <https://ProgressMonitoringForPreschoolTeachers.org>. At the time of the study, the module was only available to the research team and the participating teachers.

Research questions guiding this study were as follows:

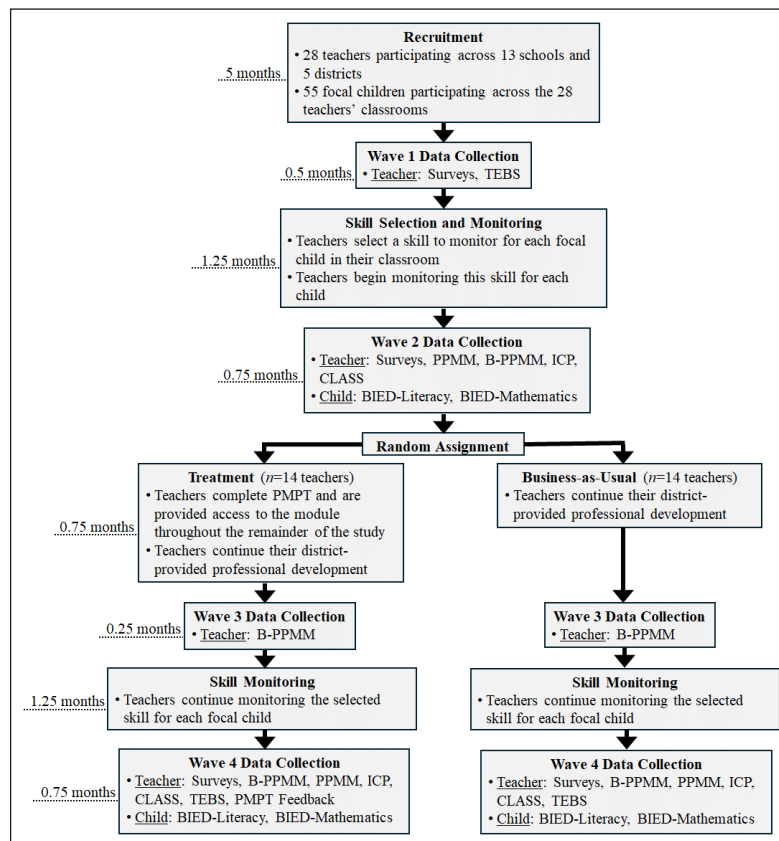
1. Does completion of an online professional development module result in improved progress monitoring implementation quality and ability by preschool teachers? (Confirmatory)
2. Is completion of the module associated with improvements to teacher self-efficacy, classroom quality, and child outcomes? (Exploratory)

## METHOD

This study was preregistered through Open Science Framework prior to the start of data collection (<https://doi.org/10.17605/osf.io/vkb6u>). Deviations from the preregistration protocol were made and include analyzing outcome variables that were not specified in the protocol. Throughout the narrative and tables of this manuscript, we have specified when these deviations are present. Per the Institutional Review Board approval for this study, all reported data and syntax are available for public review, as well as for secondary research purposes at <https://osf.io/mwbdul/>.

## Design

This study employed a randomized group design across 28 classrooms, in which preschool teachers were randomly assigned to a business-as-usual (BaU) condition or a treatment condition. Randomization was conducted by assigning each teacher a random number using the RANDOM function in Microsoft Excel, and then ordering the numbers from lowest to highest. The 14 teachers with the lowest numbers were assigned to the treatment condition and the teachers with the 14 highest numbers were assigned to the BaU condition. Assignment occurred after the second wave of data collection (discussed further in the Data Collection section). The study also included 55 child participants, referred to as focal children within each classroom. Given that randomization was completed at the teacher level, the participating children were assigned to the BaU or treatment condition based on the assignment of their classroom teacher. Refer to Figure 1 for an overview of the study's timeline and activities. The total duration of the study was 5.5 months.



**Figure 1:** Study timeline and activities.

*Note.* TEBS = Teacher Efficacy Beliefs Scale-Self; PPMM = Preschool Progress Monitoring Measure, B-PPMM = Brief Preschool Progress Monitoring Measure, ICP = Inclusive Classroom Profile, CLASS = Classroom Assessment Scoring System, BIED = Brigrance Inventory of Early Development, PMPT = Progress Monitoring for Preschool Teachers.

The independent variable was access to and completion of the online training module, *Progress Monitoring for Preschool Teachers* (PMPT). Teachers assigned to the treatment condition completed the PMPT after the second wave of data collection. No additional coaching or follow-up support was provided as part of the independent variable. Teachers retained access to the module throughout the remainder of the study. Teachers assigned to the BaU condition did not receive access to the module until after the study was completed.

## Setting, Sampling, and Participants

Seven school districts in the southeast United States were initially recruited for participation. These districts were selected given their (a) between-district variation and (b) proximity to the research team. Five of the seven districts ultimately participated in the study. Representatives for one district declined to participate given prior professional development commitments for their teachers, and representatives for one district declined to participate due to perceptions of the study being misaligned with the needs of their teachers at the time. Of the participating districts, one encompassed the second most populous city in the state and was classified as a midsize city, two districts comprised town-sized communities, and two districts were classified as rural (National Center for Education Statistics, 2021).

The research team presented information about the study to each district, and then teachers volunteered to participate. Eligible teachers needed to be employed by a school district as a certified preschool teacher and serve children 3 to 5 years old in a classroom setting. Thirty teachers were initially recruited, but two teachers decommitted prior to submitting their consent forms, and there was insufficient time for the research team to recruit replacement teachers before the start of data collection. The final sample of 28 participating teachers was comprised of 50% from the midsize city and 50% from the town and rural districts. These percentages reflect relative oversampling from the town and rural communities given that the midsize city comprised two-thirds of the preschool teachers eligible for study participation across all participating districts.

After teachers consented to participate, consent forms were sent home with two children in each classroom. These children were initially selected based on teacher nomination. After a child's consent form was returned, the research team confirmed through a teacher-completed questionnaire that the child met the study's inclusion criteria. Participating children needed to (a) demonstrate educational needs focused on pre-academic literacy or math skills, (b) be able to follow simple one-step directions (e.g., stand up, clap hands), and (c) be able to imitate speech and motor movements. Children who did not meet the inclusion criteria were dropped from the study, and a teacher nominated a new child for participation. Given stipulations of the Institutional Review Board approving the study, the research team was unable to determine child eligibility prior to obtaining consent to discuss the child's educational needs with the teacher. A total of 56 children meeting the eligibility criteria were initially recruited for participation (i.e., two children in each classroom); however, one child moved schools prior to the first wave of data collection, resulting in a final sample size of 55 focal children. Based on information reported in the teacher-completed questionnaires, the research team assigned one focal child in their classroom as a *math* child and one focal child as a *literacy* child. The research team then provided the teachers with a list of literacy and math skills to select for each child, for which the teacher would monitor the child's progress throughout the course of the study. Refer to Table 1 for a list of the literacy and math skills from which teachers could select. The skills were identified by the research team to ensure a level of standardization was present across the teachers when engaging in progress monitoring.

TYPE	SKILL	SAMPLE QUESTIONS A TEACHER MAY ASK TO ASSESS CHILD PERFORMANCE
Literacy	Receptive identification of letters	“Where is letter B?”
	Expressive identification of letters	“What letter is this?”
	Receptive identification of letter sounds	“Which letter makes the ‘buh’ sound?”
	Expressive identification of letter sounds	“What sound does B make?”
	Naming rhyming words	“What word rhymes with ‘house’?”
Mathematics	Counting a set of up to 10 objects	“How many blocks do you have?”
	Counting a requested number of up to 10 objects from a larger set	“Hand me 8 blocks, please.”

**Table 1:** Focal Children’s Skills for Progress Monitoring.

## Experimental Conditions

### Business-as-Usual

Teachers assigned to the BaU condition were not provided with any training or professional development opportunities by the research team, nor did these teachers receive access to the online module. Teachers continued to participate in their district-provided professional development opportunities, as well as any extra-curricular training experiences that they were independently pursuing (e.g., graduate courses). Teachers in the BaU condition reported receiving a mean of 10.57 hr of professional development from their school district throughout the study, for which 0.63 hr were specific to progress monitoring. Throughout the study, teachers were instructed by the research team to monitor the progress of their focal children on their selected math or literacy skill. Other than submitting data required for the study, the teachers were not required by the research team to make any other changes to their provision of educational services.

### Treatment

Teachers assigned to the treatment condition received access to and were required to complete the PMPT. The module was presented in a self-paced asynchronous format and could be completed using an individual’s personal (or work) tablet device, laptop, or desktop computer running any mainstream internet browser (e.g., Chrome, Firefox, Safari). The content of the module addressed research-supported practices for monitoring the progress of learners with individualized needs (refer to Shepley et al. [2024] for additional information on the development and content of the module). Downloadable resources were presented in the module (e.g., example data sheets). Throughout the study, teachers were instructed by the research team to monitor the progress of their focal children on their selected math or literacy skill, and upon completion of the module the research team instructed the teachers to use the strategies they learned from the module when monitoring the progress of their focal children. No feedback or coaching was provided by the research team to support the teachers’ integration of the learned strategies.

As with the teachers in the BaU condition, teachers assigned to the treatment condition continued to participate in their district-provided professional development opportunities, as well as any extra-curricular training experiences that they were independently pursuing. As such, the completion of the module should be viewed as an additive training for these teachers, rather than a replacement to what they were already receiving through their districts. Other than monitoring child progress using the learned strategies and submitting data required for the study, the teachers were not required by the research team to make any other changes to their provision of educational services. Teachers in the treatment condition reported receiving a mean of 8.71 hr of professional development from their school district throughout the study, for which 0.81 hr were specific to progress monitoring.



Backend data and teacher reports provided information about how teachers interacted with the module. Teachers completed the module in an average of 128 min ( $SD = 34$  min), with 11 teachers completing the module in one sitting and three teachers completing it over multiple sittings. All teachers used a laptop or desktop computer to complete the module, with operating systems including Windows, MacOS, and ChromeOS. Five teachers completed the module using a personal device and nine teachers used a work device. All teachers downloaded content from the module and printed an average of six pages ( $SD = 6$  pages) of resources. Following completion of the module, teachers spent an average of 92 min ( $SD = 71$  min) revisiting the content from the module throughout the remaining two months of the study.

## Dependent Measures

Two types of dependent measures were utilized in this study, confirmatory and exploratory. Confirmatory measures were those for which the research team was seeking to identify causal relations, which informed the study's a priori power analysis and resultant sample size. Exploratory measures were included to inform future confirmatory research studies.

### Confirmatory

This study employed two confirmatory measures, the (a) *Preschool Progress Monitoring Measure* (PPMM) and (b) *Brief Preschool Progress Monitoring Measure* (B-PPMM). Refer to Shepley et al. (2022) and Shepley et al. (2024) for information on the development and initial validation of the measures.

The PPMM is an observation-based assessment that provides a measure of a teacher's progress monitoring implementation quality in accordance with the practices detailed in the online training module. The assessment is completed by having a trained observer (a) watch a teacher via videorecording interact with a child to collect progress monitoring data on a target skill, (b) review the teacher's data sheet and scored data, and (c) appraise the appropriateness of the teacher's analysis and decision-making in response to the collected data via a questionnaire. In total, the PPMM consisted of 25 indicators of implementation quality, for which the trained observer scores each indicator as being present or not present (see Supplemental Table S1). Initial pilot studies to develop the PPMM indicators suggest adequate inter-rater reliability, with scorers reporting a mean agreement of 93.79% when using the measure across classrooms, teachers, and child skills (Shepley et al., 2022). Scores on the PPMM are reported as the percentage of indicators present; thus, if a teacher demonstrated the presence of 14 indicators, the teacher's score on the PPMM would be 56.00 percentage points. Higher scores on the PPMM suggest greater implementation quality.

In this study, teachers used a tablet device provided by the research team to record themselves collecting data with their focal children, take pictures of their data sheets, and respond through typed text to questions about their analysis of the collected data and changes they may make to their teaching as a result of their collected data. The teachers uploaded their videos, pictures, and typed responses using a secure portal which was accessible only by one member of the research team. Uploaded content was reviewed on a regular basis throughout the study by a member of the research team to ensure that the content was accessible (e.g., videos were not damaged or corrupted during the upload process) and appropriate for scoring with the PPMM (e.g., all borders of a data sheet were in frame, a child's responses were viewable and or audible throughout the entirety of the video). Prior to uploading content for the study, all teachers completed a self-paced technology training app that was preloaded on the tablets to teach them how to record videos, take pictures, type responses, and access the upload portal.

The B-PPMM was used to evaluate changes in teachers' progress monitoring abilities in accordance with the practices detailed in the online training module. The measure is formatted as an online, test-based assessment, and is designed to be completed independently by a teacher in less than 20 min (Shepley et al., 2024). Throughout the assessment, videos

of data collection sessions and pictures of data sheets are presented alongside questions about (a) data collection, (b) data analysis, and (c) data-based decision making. A validation study utilizing Rasch modeling found the B-PPMM to have adequate item reliability (.94) and unidimensionality (1.67), with all items demonstrating appropriate fit statistics between  $-2.0$  and  $2.0$  (Shepley et al., in press). Of note, item level difficulties fell between  $-.89$  and  $.92$  with a person reliability estimate of  $.57$ , indicating that the measure may not reliably identify teachers at the extremes of ability level. The measure contained 9 question types, with each question type repeated 3 to 4 times, totaling 29 items to which a user needed to provide a response (e.g., *True or False. The teacher in the video always responded appropriately to each child's answer*; see Supplemental Table S2). Scores from the B-PPMM are reported as the percentage of items answered correctly. For example, if a teacher answered 19 of the 29 items correctly, then the teacher's score on the B-PPMM would be 65.52 percentage points. Teachers required an average of 16 min 25 s ( $SD = 8$  min 14 s) to complete the B-PPMM throughout the study.

## Exploratory

The *Instructional Support* domain of the *Classroom Assessment Scoring System* (2nd Edition; CLASS; Teachstone, 2022a) was used to examine the quality of teacher-child interactions throughout classrooms. We perceived that the online module's partial focus on data-based decision making to inform teacher instruction may result in changes to a teachers' instructional practices. The assessment is scored through direct observation of classroom activities by a trained data collector and yields a rating from 1 to 7 across the domain. Higher scores reflect better classroom quality. Research has demonstrated that ratings of a classroom are generally consistent over time when using the CLASS ( $r = 0.65$ ; NICHD Early Childhood Care Research Network, 2002, 2005). It should be noted that we used the second edition of the CLASS for this study, and the psychometric properties reported throughout the literature generally pertain to the first edition of the assessment.

The *Monitoring Children's Learning* item of the *Inclusive Classroom Profile* (Research Edition; ICP; Soukakou, 2016) was used to examine the quality of systemic progress monitoring practices across the classrooms. We perceived that the ICP may function as a more distal outcome than the PPMM and B-PPMM, given that the *Monitoring Children's Learning* item is meant to capture school or district-wide practices. The ICP demonstrates evidence of internal consistency ( $\alpha = .88$ ), with the *Monitoring Children's Learning* item demonstrating an item-total correlation of  $.63$  and a factor loading of  $.62$  (Soukakou et al., 2014). This item is scored on a scale of 1 to 7 by a trained observer conducting a teacher interview and a review of classroom documents (e.g., data sheets). The final score is determined based on 12 indicators that are evaluated in sequential order. The indicators use a gating system, by which if certain indicators are not met then the remaining indicators are not evaluated. For this study, all indicators were evaluated to calculate the percentage of indicators met.

The *Accommodating Individual Differences* component of the Teachers' Efficacy Beliefs System—Self (TEBS) was used to examine teachers' perceptions about their progress monitoring abilities. This component contains 7 items about a teachers' confidence in their ability to engage in various aspects of progress monitoring. Each item is rated on a scale of 1 to 4 by the teacher and then averaged together for a final rating. Higher ratings suggest greater confidence in a teacher's progress monitoring ability. Adequate and consistent reliability estimates for this component have been observed across studies of differing sample sizes and teacher characteristics ( $\alpha = .85-.87$ ; Dellinger et al., 2008).

To examine teachers' efficiency with implementing progress monitoring practices in the classroom, two different measures were used. The first was a teacher report measure and was administered through a researcher-created questionnaire on which teachers self-reported the duration (in minutes) that they estimated spending on (a) collecting and (b) reviewing progress monitoring data for each focal child over the previous four weeks. The second measure was completed by



observing a video recording of a teacher engaged in a data collection session with a focal child. From each video, the research team calculated the total duration (in seconds) of the data collection session. Neither measure was detailed in the study's preregistration protocol as an outcome variable.

The *Literacy* and *Mathematics* subdomains of the standardized version of the Brigance Inventory of Early Development III (Curriculum Associates, 2013; BIED) were used to examine child outcomes. For children between the ages of 3 to 5 years old, the BIED's reliability estimates of internal consistency range from .94 to .97 for the *Literacy* and *Mathematics* subdomains, and test-retest reliability estimates range from .85 to .93 (French, 2012). The subdomains were assessed through direct assessment of children's skills using standardized procedures and testing materials while adhering to the rules for establishing a basal and ceiling for each item within a subdomain. The assessment took approximately 10 to 15 min to complete. Standard scores were obtained from the assessment.

## Data Collection

Four waves of data collection occurred throughout the study. Wave 1 occurred immediately after eligible teachers and children consented to participate. During this wave, teachers completed a series of surveys about their background, classroom, work experience, and professional development history; and the focal children in their classroom. Surveys were distributed via Qualtrics. In addition, teachers completed the TEBS, which was also distributed via Qualtrics. Completion of the surveys and the TEBS required an approximated average of 30 min for each teacher.

Wave 2 occurred 4 to 6 weeks after teachers completed Wave 1 and prior to the assignment of teachers to conditions. During this wave, the research team scheduled observations with each participating teacher to complete the CLASS, ICP, and BIED. Observations were scheduled during center times to ensure that two cycles of the CLASS could be completed across the same activity for each teacher, as this aligned with recommendations for using the latest edition of the CLASS for program level evaluation purposes (Teachstone, 2022b). The ICP was completed at a time during the observation when a teacher could step away from their teaching responsibilities to answer the interview questions and provide the required documentation. The research team completed the *Mathematics* portion of the BIED with the math focal child and the *Literacy* portion with the literacy focal child. The BIED was completed in an area suggested by the teacher, often at a table that was not being used for center activities. If preferred by the child or recommended by the teacher, the research team administered the BIED on the floor, rather than a table. Teachers also uploaded content for the PPMM and completed the B-PPMM. Access to the B-PPMM was provided to the teachers through a password-protected web address. Videos submitted by the teachers for the PPMM were used by the research team to extract the duration of each teacher's data collection sessions. Lastly, teachers completed a brief questionnaire asking them to report the estimated amount of time that they spent collecting and reviewing data on their focal children's selected skill.

During Wave 3, teachers completed the B-PPMM. This occurred after teachers were assigned to conditions and the teachers in the treatment condition had completed the PMPT. No additional data were collected during this wave.

Wave 4 occurred four to six weeks after teachers in the treatment condition completed the PMPT. Teachers again uploaded content for the PPMM, completed the B-PPMM, and responded to questions about the amount of time they spent collecting and reviewing data. Teachers also completed a brief survey about any changes to their classroom during the course of the study and the amount of professional development they received throughout the study. The research team scheduled observations and readministered the CLASS, ICP, and relevant portions of the BIED. Teachers in the treatment condition also completed a researcher-created survey containing open-ended questions about the usability of the module. Lastly, teachers completed the TEBS again, but the administration was different. During this administration of the TEBS, teachers were asked to indicate both (a) their perception of how confident they felt in their teaching abilities at the start of the study and (b) at the present time. This change to the administration was done to control for

potential self-enhancement bias (Deffuant et al., 2024) and response-shift bias (Ortega-Gómez et al., 2022). These biases occur when an individual completes a self-evaluation and rates themselves as being better than they actually are, which can cause a problem in evaluation studies. If an individual thinks they know a lot about a topic at the start of a study, then they will rate themselves favorably on the topic; but then as they progress through the study, they may come to realize how much they actually do not know, and so then at the end of the study they rate themselves poorly on the topic despite actually learning some things. By asking participants at the end of a study to reflect on their abilities at the start of a study, this can help control for these biases (Sajobi et al., 2018).

## Interobserver Agreement

All data collectors on the research team were past teachers with classroom experience serving young children with special needs. Each data collector held an active certification as a CLASS Observer for the Pre-K and K–3 versions of the CLASS (2nd Edition). In preparation for the study, the data collectors received a curated training on the ICP provided through the assessment’s publisher. The training was led by one of the original evaluators of the ICP, and the trainer provided consultative support to the research team throughout the study. For the CLASS, ICP, and BIED, interobserver agreement data were gathered. Given that teachers were not assigned to conditions until after Wave 2 data collection, the assignment of teachers to conditions was unknown by the data collectors and the teachers themselves throughout Waves 1 and 2 of data collection. Following treatment, only one data collector was aware of teacher assignments, and this data collector served as the secondary data collector to establish interobserver agreement during post-treatment data collection. Interobserver agreement data are reported in Table 2.

ASSESSMENT	PERCENTAGE OF TEACHERS/ CHILDREN FOR WHOM RELIABILITY DATA COLLECTED		AVERAGE PERCENTAGE OF AGREEMENT BETWEEN DATA COLLECTORS	
	PRE-Tx	POST-Tx	PRE-Tx	POST-Tx
CLASS: Emotional climate <sup>a</sup>	28.57%	–	93.3%	–
CLASS: Organization <sup>a</sup>	28.57%	–	86.65%	–
CLASS: Instructional support <sup>a</sup>	28.57%	10.71%	75.51%	89.00%
ICP: Monitoring children’s learning <sup>b</sup>	17.86%	11.1%	91.67%	91.7%
BIED: Literacy <sup>c</sup>	10.00%	5.00%	97.19%	100%
BIED: Mathematics <sup>c</sup>	12.5%	12.50%	95.81%	98.44%

**Table 2:** Interobserver Agreement Data.

*Note.* <sup>a</sup>percentage of agreement calculated according to the assessment’s protocol manual, <sup>b</sup>percentage of agreement calculated using an item-by-item analysis across all 12 indicators assessed, <sup>c</sup>percentage of agreement calculated using the gross method based on the summed raw score reported by each data collector; CLASS = *Classroom Assessment Scoring System*, ICP = *Inclusive Classroom Profile*, BIED = *Brigance Inventory of Early Development*. Tx = treatment.

PPMM scoring occurred after the study concluded. The two primary data collectors from Waves 3 and 4 (i.e., those who were naïve to condition assignments), each independently scored every teacher’s PPMM content. That is, 100% of the PPMM content was double-scored. To prevent observer drift throughout the scoring of the PPMM, disagreements were reviewed and consensus was established by the data collectors after a score was derived for every ten PPMM assessments. It should be noted that although data collectors were masked to teachers’ assignment to the treatment or BaU conditions,

the data collectors were not masked to the time point pertaining to PPMM content (i.e., Wave 2 or Wave 4), as teachers often included dates on their data sheets. Thus, data collectors were likely aware of whether the content they were scoring pertained to pre-treatment or post-treatment information.

## Analysis

Data on the pre- and post-treatment characteristics of the teachers, classrooms, and focal children were analyzed to identify differences between participants in the treatment and BaU groups. For continuous variables, and categorical variables with only two categories reported (e.g., all teachers reported identifying as either *male* or *female*), differences were assessed by entering each variable as the outcome within an OLS regression model and including a dichotomous assignment variable (0 = BaU, 1 = treatment) as the independent variable. The p-value for the assignment variable was used to gauge the significance of a difference between the groups. For categorical variables with more than two categories reported, only descriptive statistics are provided given that inferential statistics would likely yield unreliable estimates due to the low sample size of teachers and children ( $n \leq 28$ ).

Dependent variables were analyzed to evaluate the impact of the online module on confirmatory measures and to examine associations between the online module and exploratory measures. Both confirmatory and exploratory outcomes were analyzed within OLS regression models with robust standard errors applied and Benjamini-Hochberg corrections assessed for p-values to account for multiple significance tests. Post-treatment scores were entered as the dependent variable ( $Y$ ) with assignment to treatment (1) or the BaU condition (0) as the primary effect of interest ( $\beta_1$ ). In alignment with the preregistered analysis plan, pre-treatment scores on the corresponding dependent variable were entered as a covariate ( $\beta_2$ ). We report outcomes across models that exclude (Model 1) and include (Model 2) the covariate for pre-treatment scores, as output from Model 1 provides context for interpreting the output from our primary analytical model, Model 2. The formula for Model 2, with subscript  $i$  referring to each teacher (or child depending upon the dependent variable) and  $\epsilon$  as the error term, is:

$$PostTx\_Score_i = \beta_0 + \beta_1 Assignment_i + \beta_2 PreTx\_Score_i + \epsilon_i$$

## RESULTS

Pre- and post-treatment information on differences between groups is presented in Tables 3–8. Notable pre-treatment differences include (a) the treatment group having no teachers with experience using the Qualtrics platform and the BaU group having five teachers with experience ( $p = .01$ ), (b) the treatment group scoring 3.39 percentage points higher on the PPMM than the BaU group when collecting progress monitoring data with children on math skills ( $p = .04$ ), (c) the BaU group having a .64 higher average rating on the *Organization* section of the CLASS than the treatment group ( $p = .07$ ), and (d) children in the treatment group having a 4.33 higher average standard score on the *Mathematics* subdomain of the BIED than the BaU group ( $p = .06$ ). Given the number of statistical significance tests that were conducted for identifying differences between the groups, the identification of four variables on which the groups meaningfully differ suggests that the study's random assignment mechanism functioned appropriately. Table 9 provides data from the analytical models used to estimate the impact of, and association between, the module and dependent variables. Results and model information not reported in the manuscript may be accessed at <https://osf.io/mwbdu/>.

VARIABLE CATEGORY/VALUE	BaU (n = 14)	Tx (n = 14)	p
Gender (n)			
Female	13	14	.33
Male	1	0	
Race (n)			
White	11	13	–
Black	2	0	
Other <sup>a</sup>	1	1	
Age (yrs)			
Mean (SD)	36.9 (9.8)	34.9 (7.5)	.55
Education (n)			
Bachelor	4	8	.14
Masters	10	6	
Years in paid position with preschool-aged children			
Mean (SD)	10.8 (6.8)	10.9 (7.0)	.96
Years working in current position			
Mean (SD)	5.8 (5.3)	5.6 (4.2)	.94
Data collection frequency for IEP goals (n)			
Daily	4	2	–
Every few days	4	5	
Weekly	6	7	
Data review frequency for IEP goals (n)			
Every few days	1	1	–
Weekly	8	8	
Bi-weekly	1	3	
Monthly	2	1	
Twice a year	0	1	
Ever received PD on PM (n)			
Yes	0	3	.70
No or I'm not sure	14	11	
Hours of PD on PM in last year (n)			
None	8	9	–
Less than 2	2	3	
2 to 5	3	2	
6 to 10	1	0	
Experience with online training modules (n)			
Yes	13	12	.56
No or I'm not sure	1	2	
Experience with the Qualtrics platform (n)			
Yes	5	0	.01
No or I'm not sure	9	14	

**Table 3:** Teacher Information Prior to Treatment.

*Note.* <sup>a</sup>one teacher in the BaU group identified as Asian and one teacher in the Tx group identified as Multiracial; BaU = business as usual control group, PD = professional development, PM = progress monitoring, Tx = treatment group.

<b>VARIABLE CATEGORY/VALUE</b>	<b>BaU (n = 14)</b>	<b>Tx (n = 14)</b>	<b>p</b>
Day structure (n)			
Two half-day classes <sup>a</sup>	13	13	1.0
One full-day class	1	1	
Number of days per week children attend (n)			
4	6	8	.47
5	8	6	
Number of additional adults in class			
Mean (SD)	2.4 (0.8)	2.0 (0.6)	.16
Number of children served (Mean, SD)			
Total	27.0 (7.3)	29.1 (7.6)	.47
Female	10.9 (3.6)	11.8 (4.1)	.57
Male	16.1 (4.0)	17.3 (4.7)	.47
White	13.5 (9.9)	15.6 (12.6)	.62
Black	7.1 (4.4)	5.4 (3.7)	.27
Latinx	4.6 (3.8)	4.6 (3.2)	1.0
Asian	0.4 (0.8)	1.9 (3.2)	.10
DLL	4.6 (4.4)	5.2 (3.5)	.67
IEP	11.9 (5.2)	13.6 (4.2)	.37
CLASS Scores (Mean, SD)			
Emotional Climate	6.0 (0.6)	5.7 (1.0)	.37
Organization	5.7 (0.7)	5.1 (0.8)	.07

**Table 4:** Classroom Information Prior to Treatment.

*Note.* <sup>a</sup>half-day classes consisted of different children with one group served in the morning and the other group served in the afternoon. BaU = business as usual control group, CLASS = Classroom Assessment Scoring System, DLL = Dual Language Learner, IEP = Individualized Education Plan, Tx = treatment group.

<b>VARIABLE CATEGORY/VALUE</b>	<b>LITERACY</b>			<b>MATH</b>		
	<b>BaU (n = 14)</b>	<b>Tx (n = 13)</b>	<b>p</b>	<b>BaU (n = 14)</b>	<b>Tx (n = 14)</b>	<b>p</b>
Age (months)						
Mean (SD)	57.8 (4.3)	58.7 (6.2)	.66	57.4 (7.6)	56.7 (7.0)	.82
Gender (n)						
Female	4	3	.76	6	6	1.0
Male	10	10		8	8	
Race (n)						
White	6	7	–	6	6	–
Black	5	2		4	5	
Latinx	0	2		2	1	

(Contd.)



VARIABLE CATEGORY/VALUE	LITERACY			MATH		
	BaU (n = 14)	Tx (n = 13)	p	BaU (n = 14)	Tx (n = 14)	p
Asian	0	1		0	0	
Multi	3	1		2	2	
Dual Language Learner (n)						
No	14	10	.06	13	13	1.0
Yes	0	3		1	1	
Eligibility (n)						
N/A	5	4	–	6	5	–
Autism	3	2		4	1	
DD	5	4		3	6	
OHI	1	0		0	0	
SLI	0	3		1	2	
Household (n)						
Two parent	10	9	–	10	10	–
One parent	3	2		0	1	
Other	1	2		4	3	
Development Index <sup>a</sup>						
Mean (SD)	1.8 (8.0)	1.6 (0.5)	.56	1.6 (0.6)	1.6 (0.5)	.94

**Table 5:** Focal Child Information Prior to Treatment.

*Note.* “scores based on teacher-report using The Abilities Index (Simeonsson & Bailey, 1991). BaU = business as usual control group, DD = Developmental Delay, OHI = Other Health Impairment, SLI = Speech-Language Impairment, Tx = treatment group.

VARIABLE CATEGORY/VALUE	LITERACY			MATH		
	BaU (n = 14)	Tx (n = 13)	p	BaU (n = 14)	Tx (n = 14)	p
Selected Literacy Skill (n)						
Rec letter ID	7	7	–	–	–	
Exp letter ID	5	4		–	–	
Exp letter sound ID	2	1		–	–	
Rhyming words	0	1		–	–	
Selected Math Skill (n)						
Counting sets	–	–		8	6	.47
Counting requested	–	–		6	8	
Months working with child						
Mean (SD)	11.4 (9.4)	10.2 (8.3)	.74	9.6 (7.1)	11.0 (8.8)	.65
Data collection frequency for child's goals (n)						
Daily	3	1	–	4	1	–
Every few days	5	2		4	1	

(Contd.)

VARIABLE CATEGORY/VALUE	LITERACY			MATH		
	BaU (n = 14)	Tx (n = 13)	p	BaU (n = 14)	Tx (n = 14)	p
Weekly	5	7		3	10	
Bi-weekly	0	0		1	0	
Monthly	1	0		2	0	
Bi-monthly	0	3		0	2	
Data review frequency for child's goals (n)						
Daily	2	0	–	2	0	–
Every few days	1	2		1	2	
Weekly	9	4		8	7	
Bi-weekly	1	1		1	2	
Monthly	1	4		2	2	
Bi-monthly	0	2		0	1	

**Table 6:** Teacher Information Specific to Focal Children Prior to Treatment.

*Note.* BaU = business as usual control group, Exp = Expressive, ID = Identification, Rec = Receptive, Tx = treatment group.

LEVEL OF ANALYSIS OUTCOME VARIABLE	BaU (n = 14)	Tx (n = 14)	p
Teacher/Classroom			
PPMM (Literacy focal child)	40.15 (3.14)	43.54 (3.49)	.32
PPMM (Math focal child)	35.71 (1.78)	41.16 (2.53)	.04
B-PPMM	39.16 (11.28)	41.63 (11.59)	.57
CLASS: Instructional support	2.75 (1.18)	2.66 (1.16)	.83
ICP: Monitoring children's learning (score)	2.86 (0.95)	3.21 (2.01)	.55
ICP: Monitoring children's learning (percentage of indicators met)	67.86 (10.77)	69.05 (14.41)	.81
TEBS: Accommodating individual differences	2.89 (0.34)	2.82 (0.48)	.65
Minutes reported collecting data on focal children's skills in past month	65.07 (39.38)	73.00 (47.45)	.63
Minutes reported reviewing data on focal children's skills in past month	34.14 (20.51)	36.14 (17.63)	.78
Minutes observed collecting data across focal children's skills	5.93 (2.86)	6.37 (4.25)	.75
<b>Child</b> (Literacy)	n = 11	n = 9	
BIED: Literacy	77.55 (6.62)	78.11 (4.68)	.83
<b>Child</b> (Math)	n = 12	n = 12	
BIED: Mathematics	79.50 (6.04)	83.83 (4.76)	.06

**Table 7:** Outcome Information Prior to Treatment.

*Note.* Values reported as means with standard deviations in parentheses; BaU = business as usual control group, Tx = treatment group, PPMM = Preschool Progress Monitoring Measure, B-PPMM = Brief Preschool Progress Monitoring Measure, CLASS = Classroom Assessment Scoring System, ICP = Inclusive Classroom Profile, TEBS = Teacher Efficacy Beliefs Scale-Self, BIED = Brigance Inventory of Early Development.

VARIABLE CATEGORY/VALUE	BaU (n = 14)	Tx (n = 14)	p
Number of children served (Mean, SD)			
Total	27.64 (6.64)	30.07 (7.02)	.36
Total (change from pre-treatment) <sup>a</sup>	0.64 (1.91)	1.00 (1.47)	.58
IEP	14.36 (4.41)	14.93 (4.57)	.74
IEP (change from pre-treatment) <sup>a</sup>	2.43 (3.41)	1.36 (3.50)	.42
Number of additional adults in class (Mean, SD)			
Total	2.25 (1.01)	1.93 (0.81)	.36
Total (change from pre-treatment) <sup>a</sup>	-0.11 (0.68)	-0.07 (0.81)	.90
Hours of school-district-provided PD since start of study (Mean, SD)			
Total	10.57 (10.01)	8.71 (9.27)	.62
Progress monitoring <sup>b</sup>	0.63 (1.06)	0.81 (1.47)	.71

**Table 8:** Teacher and Classroom Information at Post-Treatment.

*Note.* <sup>a</sup>refers to the difference when subtracting post-treatment values by pre-treatment values for each teacher and then averaging for the group, <sup>b</sup>refers solely to district-provided professional development addressing progress monitoring and does not include time spent engaging with the online training module as part of this study; BaU = business as usual control group, IEP = Individualized Education Plan, PD = Professional Development, Tx = treatment group.

LEVEL OF ANALYSIS OUTCOME VARIABLE	MODELS (B, SE)	
	1	2
Teacher/Classroom		
PPMM (Literacy focal child)	<b>29.55</b> (3.30)***	<b>29.13</b> (3.37)***
PPMM (Math focal child)	<b>31.35</b> (4.58)***	<b>32.53</b> (4.87)***
B-PPMM (1 week after treatment) <sup>a</sup>	<b>19.71</b> *** (3.76)	<b>19.19</b> *** (3.63)
B-PPMM (1 month after treatment) <sup>a</sup>	<b>20.69</b> *** (4.29)	<b>20.38</b> *** (4.45)
CLASS: Instructional support <sup>b</sup>	-0.35 (0.34)	-0.30 (0.26)
ICP: Monitoring children's learning (score) <sup>b</sup>	0.21 (0.58)	0.06 (0.53)
ICP: Monitoring children's learning (percentage of indicators met) <sup>c</sup>	6.00 (3.40)	5.39 (3.22)
TEBS: Accommodating individual differences <sup>b</sup>	<b>0.43</b> ** (0.13)	<b>0.46</b> *** (0.11)
Minutes reported collecting data on focal children's skills in past month <sup>c</sup>	-20.79 (11.73)	-22.42 (11.77)
Minutes reported reviewing data on focal children's skills in past month <sup>c</sup>	-7.93 (10.73)	-9.70 (8.80)
Minutes observed collecting data across focal children's skills <sup>c</sup>	-1.51* (0.72)	-1.60* (0.69)
<b>Child</b> (Literacy)		
BIED: Literacy <sup>b</sup>	1.64 (4.83)	1.04 (4.11)
<b>Child</b> (Math)		
BIED: Mathematics <sup>b</sup>	5.17 (3.91)	-0.07 (3.30)

**Table 9:** Analysis of Outcome Variables at Post-Treatment.

*Note.* Coefficient estimate in each model reflects the average change for a teacher who completed the online module relative to a teacher who did not complete the module; Coefficient estimates are expressed in unstandardized values; <sup>a</sup>indicates an outcome variable for confirmatory analysis per the preregistration protocol, <sup>b</sup>indicates an outcome for exploratory analysis per the preregistration protocol; <sup>c</sup>indicates an outcome variable that was not included in the preregistration protocol; PPMM = Preschool Progress Monitoring Measure; B-PPMM = Brief Preschool Progress Monitoring Measure, CLASS = Classroom Assessment Scoring System, ICP = Inclusive Classroom Profile, TEBS = Teacher Efficacy Beliefs Scale-Self, BIED = Brigance Inventory of Early Development; \*p < .05, \*\*p < .01, \*\*\*p < .001, boldened values indicate statistical significance at the 5% level when applying the Benjamini-Hochberg p-value adjustment.

## Confirmatory

For teachers who completed the PMPT, they scored 29.55 percentage points ( $SE = 3.30$ , 95% CI [22.74, 36.35]) higher on the PPMM, on average, than teachers in the BaU group when working with a literacy focal child. This finding suggests that if a teacher in the BaU group scored 40.00 percentage points on the PPMM when working with a child on a literacy skill, a teacher who completed the online module would score 69.55 percentage points on the PPMM. When working with a math focal child, teachers in the treatment group scored 31.35 percentage points ( $SE = 4.58$ , 95% CI [21.94, 40.75]) higher on the PPMM, on average, than teachers in the BaU group who relied solely on their district provided professional development. When controlling for pre-treatment scores in Model 2, both the magnitude and precision of the effects were relatively unchanged for the PPMM scores pertaining to literacy and math focal children. Treatment estimates of the PMPT's impact on teachers' PPMM scores retained statistical significance at the 5 percent level when applying Benjamini-Hochberg corrections ( $p = .003$ ).

One week after teachers completed the PMPT, they scored an average of 19.71 percentage points ( $SE = 3.76$ , 95% CI [11.99, 27.43]) higher on the B-PPMM than teachers in the BaU group. This finding suggests that if an average teacher in the BaU group scored 45.00 percentage points on the B-PPMM, then an average teacher who completed the PMPT would score 64.71 percentage points on the B-PPMM. When controlling for pre-treatment scores on the B-PPMM in Model 2, both the magnitude of the effect ( $B = 19.19$ ) and its precision ( $SE = 3.63$ ) remained relatively unchanged from Model 1. One month after the initial completion of the PMPT, teachers scored an average of 20.69 percentage points ( $SE = 4.29$ , 95% CI [11.87, 29.51]) higher on the B-PPMM than teachers who relied solely on their district provided professional development. Again, the magnitude and precision of the effect in Model 2 did not indicate meaningful differences from Model 1. All treatment estimates of the PMPT's impact on teachers' B-PPMM scores retained statistical significance at the 5 percent level when applying Benjamini-Hochberg corrections ( $p = .003$ ).

## Exploratory

No meaningful associations were observed across the CLASS or ICP between teachers who did and did not complete the PMPT ( $p > .090$ ). Completion of the PMPT was associated with a 0.43 point ( $SE = .13$ , 95% CI [0.16, 0.70]) increase in self-efficacy as measured by TEBS ratings relative to teachers who did not complete the module. This finding suggests that if an average teacher in the BaU group reported a self-efficacy rating on the TEBS of 3.00, an average teacher who completed the PMPT would report a rating of 3.43. The association between PMPT completion and TEBS ratings retained statistical significance at the 5 percent level when applying Benjamini-Hochberg corrections ( $p = .013$ ). When including a covariate for pre-treatment TEBS scores based on either (a) Wave 1 TEBS data or (b) Wave 4 TEBS data in which teachers reflected on their abilities at the start of the study, coefficient estimates were consistent ( $B = .39$ ,  $B = .46$ ) and retained statistical significance when applying Benjamini-Hochberg corrections ( $p < .008$ ).

Descriptive differences were observed across teachers' reporting of how much time they estimated spending on data collection and review, with teachers who completed the module reporting that they spend less time on both data collection ( $B = -20.79$ ) and review ( $B = -22.42$ ) than teachers who did not complete the module; however, these estimates did not achieve statistical significance ( $p > .088$ ). The observational measure completed by calculating the amount of time teachers spent engaging in data collection based on video recordings, did achieve statistical significance ( $p = .046$ ), with completion of the module being associated with a 90 s ( $SE = 43$  s, 95% CI [-180 s, -2 s]) average decrease in the duration of a teachers' data collection sessions. When controlling for pre-treatment durations, the magnitude of the association further decreased by an average of 6 s, yielding a 25% decrease in data collection duration for the treatment group relative to the BaU group. This finding did not retain statistical significance at the 5 percent level when applying Benjamini-Hochberg corrections ( $p = .099$ ).

When gathering data on child outcomes, assessments could not be completed for three literacy children in the treatment group, three literacy children in the BaU group, two math children in the treatment group, and two math children in the BaU group. The assessments could not be completed due to challenges in securing child attention throughout administration of the assessments or children vocally refusing to participate which the research team considered a form of dissent. Overall attrition for the BIED *Literacy* measure was 25.93% and differential attrition was 9.34%. Overall attrition for the BIED *Mathematics* measure was 14.29% and differential attrition was 0.00%. Anecdotally, it was reported by multiple teachers that these children did not have experience with performance-based standardized assessments, such as the BIED. Authentic assessments (e.g., *Teaching Strategies GOLD*; *Assessment, Evaluation, and Programming System*) were typically used for program planning purposes in the classrooms and to support eligibility decisions along with the use of indirect assessments (e.g., teacher- and parent-rating scales). Exploratory analyses do not indicate a meaningful association between the completion of the PMPT and children's scores on the BIED *Literacy* subdomain nor the BIED *Mathematics* subdomain ( $p > .200$ ).

## DISCUSSION

This study aimed to address the need for research and resources related to professional development for supporting the progress monitoring efforts of teachers serving children with individualized needs. The findings suggest that the online module, PMPT, improved teachers' abilities and implementation quality when collecting progress monitoring data, analyzing those data, and making appropriate data-based decisions. In addition, exploratory data suggest that further research is warranted to evaluate the causal impact of the module on (a) teachers' confidence in their progress monitoring abilities and (b) teachers' efficiency when engaged in progress monitoring activities. This study does not suggest a clear link between completion of the PMPT and changes in broad measures of classroom quality nor child outcomes when measured using standardized measures in the relative short term (i.e., six weeks post treatment); though it should be noted that the a priori determined sample size for this study was not informed by considerations of changes in classroom quality nor child outcomes. Nuanced measures of classroom quality specific to special education or Tier 3 services within a multi-tiered system of support may be more sensitive to the practices conveyed within the PMPT; however, we are unaware of any assessments with evidence of technical adequacy that may be relevant as of this writing (cf., the *Examining Data Informing Teaching* measure [Monahan et al., 2015]). More individualized measures of child outcomes (e.g., progress on objectives depicted within a child's individualized education program) may also be better aligned with the content addressed in the PMPT, given that the module was developed specifically for supporting the individualized needs of children rather than addressing needs about universal or classroom-wide outcomes.

On the relevance of the teacher training method employed in this study and its application to the field of early education, we perceive that asynchronous online modules may offer utility for certain types of outcomes targeted through professional development. This study demonstrated that when professional development targeted teachers' progress monitoring practices, module-based learning may be effective. We are hesitant to suggest that online modules may be effective for improving *all* teacher practices, as research from the behavioral sciences indicates that the implementation of complex procedures and interventions may still require observation and individualized feedback from a trained professional; though, the amount of observation and feedback will likely be mitigated if asynchronous training methods are utilized first (Gerencser et al., 2020). Further application and evaluation of asynchronous online modules and training methods for early childhood professionals are warranted, in particular for school districts located in rural and geographically isolated communities where resource intensive professional development methods may be impractical.

For administrators, principals, center directors, and early childhood educators who may be considering using the PMPT for professional development purposes, we think it critical to delineate some of the aspects of our study that may limit (or enhance) the generalization of our findings. First, all children in this study were receiving special education or tiered



support services within publicly funded, district-based, preschool classrooms, from teachers who held at least a bachelor's degree. Some school districts may provide special education services to young children using a different service-provision model. For example, children with special needs may be placed in community-based childcare and education centers in which the teachers have varying levels of education, are not employed directly by the school-district, and are not responsible for managing and providing special education services. Instead, a special education teacher employed by the district travels between centers to provide consultative and direct support in alignment with the children's Individualized Education Programs. Thus, we are unable to gauge the impact of the PMPT on teachers working across different models for delivering special education services to young children.

Second, all teachers in our study reported receiving very little professional development focused on progress monitoring both prior to (see Table 3) and throughout the study (see Table 8). In districts that provide more intensive professional development aimed at supporting teachers' progress monitoring abilities, the impact of the PMPT may differ from the reported results. Third, the population of the students served in the classrooms participating in this study may differ from those in other districts, given that many districts provide preschool special education services to only those children with a special education eligibility. The school districts participating in this study provided services to children with a special education eligibility, as well as children who were considered at risk for delay based on the results of a developmental screening assessment or identified risk factors (e.g., income eligibility). We urge individuals seeking to adopt the PMPT to identify relevant differences between the participants in this study and those for whom the PMPT will be provided in order to consider whether outcomes may vary.

## LIMITATIONS

Readers should be aware of the limitations to this study when interpreting the findings. The pre-treatment interobserver agreement percentage for the *Instructional Support* domain of the CLASS was below 80%, which is the standard commonly adopted across research studies employing the measure (Teachstone, 2022b). Attrition of children participating in the BIED *Literacy* assessment exceeded the conservative attrition standard established by the What Works Clearinghouse (n.d.). Although not pertinent to the internal validity of the study, we think it relevant to recommend that generalizations of our study's findings to early childhood professionals lacking a teaching certification and working in settings other than publicly funded school-district-based settings, should be avoided (e.g., tuition-funded preschools and childcare centers, Head Start classrooms). It should also be stressed that despite observing an improvement in teacher implementation and abilities following completion of the online module, we did not observe a change in child outcomes. Notably, child outcomes in our study functioned as exploratory outcomes and thus did not inform our sample size. Future research should prioritize child outcomes so that study resources can be appropriately allocated to ensure that a sufficient number of children can be recruited and retained to more precisely estimate the impact of the online module on child outcomes. It is highly likely that there are limitations we did not discuss nor recognize, therefore we have posted all our data and analytical syntax for public review.

## CONCLUSION

This study provides evidence to support that teacher's completion of the PMPT, an online professional development module, improves progress monitoring implementation quality and abilities. Findings also suggest that future confirmatory research studies should evaluate the impact of the PMPT across teachers' (a) self-efficacy and (b) efficiency with performing progress monitoring activities. Data do not support that the module was associated with changes to classroom ratings on the CLASS or ICP, nor children's standard scores on subsections of the BIED. Administrators

seeking to adopt the PMPT for use by their teachers should carefully consider the extent to which their teachers and the children they serve are similar (and different) to the individuals who participated in this study, as this should help in determining if the PMPT is appropriate for their populations.

## COMPETING INTERESTS

Collin Shepley receives monetary compensation for the licensing of the *Brief Preschool Progress Monitoring Measure*, which was used as an outcome measure in this study. All authors contributed to the development and validation of the online module evaluated in this study, and as such, their interpretations of the findings may be subject to unconscious biases.

## FUNDING

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant No. R324B210002 to the University of Kentucky. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

## DATA AVAILABILITY AND STUDY RESOURCES

All data, syntax, and supplemental information referenced throughout manuscript may be obtained for public review and secondary research purposes at <https://osf.io/mwbdu/>. The module utilized in this study may be accessed at <https://ProgressMonitoringForPreschoolTeachers.org/>.

## PREREGISTRATION

The study was preregistered through Open Science Framework for which the protocol may be reviewed at <https://doi.org/10.17605/osf.io/vkb6u>.

## AUTHOR CONTRIBUTIONS

**Collin Shepley:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Visualization, Writing Original Draft, Writing Review & Editing. **Amanda Duncan:** Data Curation, Investigation, Project Administration, Resources, Supervision, Writing Review & Editing. **Emily Webb:** Data Curation, Investigation, Project Administration, Writing Review & Editing.

## AUTHOR AFFILIATIONS

**Collin Shepley, PhD, BCBA-D**  [orcid.org/0000-0003-1967-9714](https://orcid.org/0000-0003-1967-9714)

University of Kentucky, Department of Early Childhood, Special Education, and Counselor Education, 229 Taylor Education Building, Lexington, Kentucky, 40506, US

**Amanda Leigh Duncan, PhD, BCBA**  [orcid.org/0000-0003-3342-7556](https://orcid.org/0000-0003-3342-7556)

University of Kentucky, Department of Early Childhood, Special Education, and Counselor Education, 229 Taylor Education Building, Lexington, Kentucky, 40506, US

**Emily Webb, MEd**

University of Kentucky, Department of Early Childhood, Special Education, and Counselor Education, 229 Taylor Education Building, Lexington, Kentucky, 40506, US

## REFERENCES

- Artman-Meeker, K., Fettig, A., Barton, E. E., Penney, A., & Zeng, S. (2015). Applying an evidence-based framework to the early childhood coaching literature. *Topics in Early Childhood Special Education, 35*(3), 183–196. <https://doi.org/10.1177/0271121415595550>
- Bishop, C., Shannon, D., & Harrington, J. (2019). Progress monitoring within the embedded instruction approach: Collecting, sharing, and interpreting data to inform instruction. In M. McLean, R. Banerjee, J. Squires & K. Hebbeler (Eds.), *Assessment: Recommended practices for young children and families: DEC recommended practices monograph series* (No. 7, pp. 135–148). Division for Early Childhood.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74.
- Brock, M. E., Cannella-Malone, H. I., Seaman, R. L., Andzik, N. R., Schaefer, J. M., Page, E. J., Barczak, M. A., & Dueker, S. A. (2017). Findings across practitioner training studies in special education: A comprehensive review and meta-analysis. *Exceptional Children, 84*(1), 7–26. <https://doi.org/10.1177/0014402917698008>
- Cheek, A. E., Rock, M. L., & Jimenez, B. A. (2019). Online module plus eCoaching: The effects on special education teachers' comprehension instruction for students with significant intellectual disability. *Education and Training in Autism and Developmental Disabilities, 54*(4), 343–357. <https://www.jstor.org/stable/26822513>
- Courtade, G. R., Test, D. W., & Cook, B. G. (2014). Evidence-based practices for learners with severe intellectual disability. *Research and Practice for Persons with Severe Disabilities, 39*(4), 305–318. <https://doi.org/10.1177/1540796914566711>
- Deffuant, G., Roubin, T., Nugier, A., & Guimond, S. (2024). A newly detected bias in self-evaluation. *PLoS ONE, 19*(2). <https://doi.org/10.1371/journal.pone.0296383>
- Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*(3), 751–766. <https://doi.org/10.1016/j.tate.2007.02.010>
- Division for Early Childhood of the Council for Exceptional Children. (2021). *Multitiered systems of support framework in early childhood: Description and implications*. <https://www.decdocs.org/position-statement-mtss>
- Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research, 88*(3), 401–433. <https://doi.org/10.3102/0034654317751918>
- Elek, C., & Page, J. (2019). Critical features of effective coaching for early childhood educators: A review of empirical research literature. *Professional Development in Education, 45*(4), 567–585. <https://doi.org/10.1080/19415257.2018.1452781>
- Endrew, F. v. Douglas County School District, 580 U.S. (2017).
- Etscheidt, S. K. (2006). Progress monitoring: Legal issues and recommendations for IEP teams. *Teaching Exceptional Children, 38*(3), 56–60. <https://doi.org/10.1177/004005990603800308>
- Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *The Journal of Special Education, 41*(2), 121–139. <https://doi.org/10.1177/00224669070410020101>
- French, B. F. (2012). *Brigance Inventory of Early Development III: Standardized and validation manual*. Curriculum Associates.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53*(3), 199–208. <https://doi.org/10.1177/001440298605300301>
- Gerencser, K. R., Akers, J. S., Becerra, L. A., Higbee, T. S., & Sellers, T. P. (2020). A review of asynchronous trainings for the implementation of behavior analytic assessments and interventions. *Journal of Behavioral Education, 29*, 122–152. <https://doi.org/10.1007/s10864-019-09332-x>

- Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, 115(4), 523–547. <https://doi.org/10.1086/681947>
- Higbee, T. S., Aporta, A. P., Resende, A., Nogueira, M., Goyos, C., & Pollard, J. S. (2016). Interactive computer training to teach discrete-trial instruction to undergraduates and special educators in Brazil: A replication and extension. *Journal of Applied Behavior Analysis*, 49(4), 780–793. <https://doi.org/10.1002/jaba.329>
- Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37. <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
- Kraft, M. A. (2023). The effect-size benchmark that matters most: Education interventions often fail. *Educational Researcher*, 52(3), 183–187. <https://doi.org/10.3102/0013189X231155154>
- Lang, S. N., Tebben, E., & Odean, R. (2024). Inequities in coaching interventions: A systematic review of who receives and provides coaching within early care and education. *Child Youth Care Forum*, 53, 141–171. <https://doi.org/10.1007/s10566-023-09748-7>
- Ledford, J. R., Lane, J. D., & Barton, E. E. (2019). *Methods for teaching in early education: Contexts for inclusive classrooms*. Routledge.
- Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The effectiveness and features of formative assessment in US K-12 education: A systematic review. *Applied Measurement in Education*, 33(2), 124–140. <https://doi.org/10.1080/08957347.2020.1732383>
- Lortie-Forgues, H., & Inglis, M. (2019). Rigorous largescale educational RCTs are often uninformative: Should we be concerned? *Educational Researcher*, 48(3), 158–166. <https://doi.org/10.3102/0013189X19832850>
- Marano, K. E., Vladescu, J. C., Reeve, K. F., Sidener, T. M., & Cox, D. J. (2020). A review of the literature on staff training strategies that minimize trainer involvement. *Behavioral Interventions*, 35(4), 604–641. <https://doi.org/10.1002/bin.1727>
- McLean, M. E., Bailey, D. B., & Wolery, M. (2004). *Assessing infants and preschoolers with special needs*. Merrill.
- McLeod, R. H., Hardy, J. K., & Carden, K. C. (2024). A review of the literature: Distance coaching in early childhood settings. *Journal of Early Intervention*, 46(1), 3–18. <https://doi.org/10.1177/10538151231159639>
- Monahan, S., Atkins-Burnett, S., Wasik, B. A., Akers, L., Hurwitz, F., & Carta, J. (2015). *Developing a tool to examine teachers' use of ongoing child assessment to individualize instruction*. U.S. Department of Health and Human Services. [https://acf.gov/sites/default/files/documents/opre/40158\\_cpm\\_clin\\_3\\_report\\_111416final\\_updated\\_covers\\_b508.pdf](https://acf.gov/sites/default/files/documents/opre/40158_cpm_clin_3_report_111416final_updated_covers_b508.pdf)
- National Center for Education Statistics. (2021). *Education demographic and geographic estimates*. <https://nces.ed.gov/programs/edge/Geographic/LocaleBoundaries>. Accessed on September 1<sup>st</sup>, 2023.
- NICHD Early Child Care Research Network. (2002). The relation of global first grade classroom environment to structural classroom features, and teacher and student behaviors. *The Elementary School Journal*, 102(5), 367–387. <https://doi.org/10.1086/499709>
- NICHD Early Child Care Research Network. (2005). A day in third grade: A large-scale study of classroom quality and teacher and student behavior. *The Elementary School Journal*, 105(3), 305–323. <https://doi.org/10.1086/428746>
- Odom, S. L. (2021). Education of students with disabilities, science, and randomized controlled trials. *Research and Practice for Persons with Severe Disabilities*, 46(3), 132–145. <https://doi.org/10.1177/15407969211032341>
- Ortega-Gómez, E., Vicente-Galindo, P., Martín-Rodero, H., & Galindo-Villardón, P. (2022). Detection of response shift in health-related quality of life studies: a systematic review. *Health and Quality of Life Outcomes*, 20(1), 1–10. <https://doi.org/10.1186/s12955-022-01926-w>

- Pollard, J. S., Higbee, T. S., Akers, J. S., & Brodhead, M. T. (2014). An evaluation of interactive computer training to teach instructors to implement discrete trials with children with autism. *Journal of Applied Behavior Analysis*, 47(4), 765–776. <https://doi.org/10.1002/jaba.152>
- Sajobi, T. T., Brahmabatt, R., Lix, L. M., Zumbo, B. D., & Sawatzky, R. (2018). Scoping review of response shift methods: current reporting practices and recommendations. *Quality of Life Research*, 27, 1133–1146. <https://doi.org/10.1007/s11136-017-1751-x>
- Schachter, R. E., Knoche, L. L., Goldberg, M. J., & Lu, J. (2024). What is the empirical research base of early childhood coaching? A mapping review. *Review of Educational Research*, 94, 627–659. <https://doi.org/10.3102/00346543231195836>
- Shepley, C., Duncan, A. L., & Setari, A. P. (2024). Toward developing and validating a measure to appraise progress monitoring ability. *Journal of Early Intervention*. Advanced online publication. <https://doi.org/10.1177/10538151241235557>
- Shepley, C., Graley, D., & Lane, J. D. (2023). Preparing preschool educators to monitor child progress: A best-evidence synthesis and call to action. *Infants & Young Children*, 37(1), 20–35. <https://doi.org/10.1097/IYC.0000000000000255>
- Shepley, C., Grisham-Brown, J., Lane, J. D., & Ault, M. J. (2022). Training teachers in inclusive classrooms to collect data on individualized child goals. *Topics in Early Childhood Special Education*, 41(4), 253–266. <https://doi.org/10.1177/0271121420915770>
- Shepley, C., Setari, A., Duncan, A. L., & Webb, E. (in press). Validity of an online assessment to appraise teacher progress monitoring ability. *Assessment for Effective Intervention*.
- Simeonsson, R. J., & Bailey, D. B. (1991). *The Abilities Index*. Frank Porter Graham Child Development Center University of North Carolina at Chapel Hill. [https://fpg.unc.edu/sites/fpg.unc.edu/files/resource-files/FPG\\_AbilitiesIndex.pdf](https://fpg.unc.edu/sites/fpg.unc.edu/files/resource-files/FPG_AbilitiesIndex.pdf)
- 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>
- Snyder, P. A., Hemmeter, M. L., & Fox, L. (2015). Supporting implementation of evidence-based practices through practice-based coaching. *Topics in Early Childhood Special Education*, 35(3), 133–143. <https://doi.org/10.1177/0271121415594925>
- Soukakou, E. (2016). *The Inclusive Classroom Profile*, research edition. Brookes Publishing.
- Soukakou, E. P., Winton, P. J., West, T. A., Sideris, J. H., & Rucker, L. M. (2014). Measuring the quality of inclusive practices: Findings from the inclusive classroom profile pilot. *Journal of Early Intervention*, 36(3), 223–240. <https://doi.org/10.1177/1053815115569732>
- Teachstone. (2022a). Classroom Assessment Scoring System 2nd edition: Pre-K–3rd. Teachstone, Inc.
- Teachstone. (2022b). Reference manual: Classroom Assessment Scoring System. Teachstone, Inc.
- What Works Clearinghouse. (n.d.). *WWC standards brief for attrition*. [https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc\\_brief\\_attrition\\_080715.pdf](https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_attrition_080715.pdf)
- Yell, M. L., & Bateman, D. F. (2017). Endrew F. v. Douglas county school district (2017). FAPE and the US supreme court. *Teaching Exceptional Children*, 50(1), 7–15. <https://doi.org/10.1177/0040059917721116>
- Yell, M. L., & Bateman, D. F. (2019). Free appropriate public education and Endrew F. v. Douglas County School System (2017): Implications for personnel preparation. *Teacher Education and Special Education*, 42(1), 6–17. <https://doi.org/10.1177/0888406417754239>