2023

Practice Elements Delivered by Teachers to Support Children Who Display Challenging Behaviors in Business-as-Usual Early Childhood Settings

Navneet Kaur
Virginia Commonwealth University

Practice Elements Delivered by Teachers to Support Children Who Display Challenging Behaviors in Business-as-Usual Early Childhood Settings

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

By: Navneet Kaur

Bachelor of Science, Arizona State University, 2018

Director: Bryce D. McLeod, Ph.D.

Professor

Department of Psychology

Virginia Commonwealth University

Richmond, Virginia

April 11, 2023

Table of Contents

Abstract

PRACTICE ELEMENTS DELIVERED BY TEACHERS TO SUPPORT CHILDREN WHO DISPLAY CHALLENGING BEHAVIORS IN BUSINESS-AS-USUAL EARLY CHILDHOOD SETTINGS

By: Navneet Kaur, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2021.

Major Director: Bryce D. McLeod, Ph.D. Professor Department of Psychology

Approximately 15-25% of children ages 3 to 5 years in early childhood settings display chronic challenging behaviors that increase their risk for developing an emotional and behavioral disorder. The early onset of emotional and behavioral disorders has long-term negative implications for these children. Therefore, it is vital to address these problems early on. One approach is through teacher-delivered practices (i.e., specific strategies, such as praise) and programs (i.e., manualized interventions that contain a collection of practices) in the classroom. Some practices and programs have shown positive outcomes for children. However, when practices and programs do not outperform business-as-usual group (BAU), it is hard to determine how to improve outcomes for children. One way to enhance programming offered to children is by describing the practices delivered by teachers in BAU early childhood classrooms. BAU is defined as existing practices teachers use without exposure to new programming at the time of the study. Characterizing the practices delivered by teachers in BAU can generate baseline data that can be used to inform quality improvement process in early childhood settings. The present study used data collected to develop and validate the Treatment Integrity Measure for Early Childhood Settings (TIMECS; McLeod et al., 2021) to examine the frequency of practices used, the average dosage of the practices, and how practices clustered together using Exploratory Factor Analysis (EFA). The study sample included 91 children aged 3 to 5

years who displayed challenging behaviors along with 54 teachers from urban and suburban early childhood settings. The study found that teachers were using all practices found in TIMECS, but some practices were observed in more observations than others, teachers used practices at varying dosages, and were found to group practices into two groups: classroom management and supportive relationships. These findings help characterize practices teachers use in BAU early childhood classrooms with children who display challenging behaviors. Findings from this study provide baseline data that can be used to inform quality improvement process and help researchers implement practices that build on teacher's existing knowledge of the practices delivered in the classroom.

Introduction

Approximately 15 to 25% of children in early childhood settings display challenging behaviors (e.g., disruptive behavior; Barbarin, 2007) that put them at an increased risk for developing emotional and behavior disorders (hereafter referred to as *young children who display* challenging *behaviors*; Michigan's MTSS Technical Assistance Center; Nelson et al., 2013). Challenging behavior in young children can be characterized as aggression, noncompliance, tantrums, defiance, inattention/hyperactivity, and property destruction (Barbarin, 2007; Strain & Timm, 2001). The early onset of challenging behaviors in young children impacts their classroom learning and educational performance (de Lijster et al., 2019; Fleming et al., 2017; Gage et al., 2014). Young children who display challenging behaviors are more likely to disrupt the classroom (Yoder & Williford, 2019) and have negative interactions with their teachers and peers (Hamm et al., 2020; Williford & Vitiello, 2020). They are also at an increased risk for developing more severe challenging behaviors later in childhood and adolescence (Larsson, 2020; Sibley et al., 2017; Sibley et al., 2016), such as repeating a grade (Steinberg & Lacoe, 2017), getting suspended or expelled (Sultan et al., 2021), dropping out of school (Kena, 2016), or getting arrested (Barra et al., 2021). It is evident that if young children who display challenging behaviors do not receive proper help, they are at a greater risk of developing severe problems; therefore, it is critical to support these children.

Early childhood settings, such as Head Start or state-funded programs, represent an ideal setting for supporting young children with challenging behaviors. These programs are intended to promote school readiness and prepare children to transition to kindergarten (Theodore, 2020). Within these settings, teacher-delivered practices (i.e., specific strategies, such as praise; The IRIS Center, 2014; McLeod et al., 2017) and programs (i.e., manualized interventions that contain a collection of practices; The IRIS Center, 2014; McLeod et al., 2017) that target social, emotional, and behavioral (SEB) outcomes can benefit young children who

display challenging behaviors (Conroy et al., 2015; Feil et al., 2014; Fullerton et al., 2009; Smith et al., 2011).

Within early childhood settings, the delivery of practices and programs are often grounded in the public health model of prevention and conceptualized as a three-tier framework that provides a continuum of intensifying supports (Cook et al., 2010; Sugai et al., 2000), including a universal level (Tier-1), an indicated level (Tier-2), and a tertiary level (Tier-3; Rones & Hoagwood, 2000; Walker et al., 1996). Within this framework, practices and programs provided to young children who display challenging behaviors are classified as Tier-2. Teachers can deliver Tier-2 practices and programs in an everyday classroom context for young children who exhibit challenging behaviors, which can help maximize benefits for these young children.

Tier 2 Evidence-Based Practice and Program Literature

Tier-2 evidence-based practices and programs (i.e., practices and programs that have shown beneficial SEB outcomes in empirical research; Forman et al., 2009) exist for young children who display challenging behaviors in early childhood settings. Certain teacher-delivered classroom practices have shown positive outcomes for young children who display challenging behaviors. In particular, teacher-delivered practices like behavior-specific praise (i.e., acknowledging when a child engages in an appropriate behavior; Stormont et al., 2008) have demonstrated significant reductions in challenging behaviors, improvements in on-task behavior, and decreases in physical aggression (Fullerton et al., 2009; Smith et al., 2011). The delivery of precorrective statements (i.e., teacher statements that provide specific appropriate behavior the child should engage in before the problem behavior occurs; De Pry & Sugai, 2002) has also been associated with improvements for children with challenging behavior, including significant reductions in challenging behaviors, improvements in on-task behavior, and decreases in physical aggression (Smith et al., 2011). Thus, the delivery of certain teacher-delivered practices are associated with improved SEB outcomes for young children who display challenging behaviors.

Several Tier-2 programs have also demonstrated positive outcomes for young children who display challenging behavior in early childhood settings. Preschool First Steps to Success (Feil et al., 2014) is a program that provides classroom management training (i.e., establishing rules in the classroom) to preschool teachers designed to help young children who display challenging behaviors in the classroom. The intervention aims to improve children's social skills and reduce problem behavior (Feil et al., 2014). In a randomized controlled trial, Preschool First Steps to Success significantly improved children's prosocial behaviors and decreased problem behaviors compared to a business-as-usual (BAU; defined as existing practices teachers use without exposure to new programming at the time of the study) control group (Feil et al., 2014). Behavioral, Emotional, and Social Training: Competent Learners Achieving School Success (BEST in CLASS) is another program designed to help teachers improve their effective instructional practices (i.e., establishing rules, behavior-specific praise) that help promote positive teacher-child interactions and improve behavioral outcomes for young children who display challenging behaviors (Sutherland et al., 2019; Vo et al., 2012). In a randomized trial, BEST in CLASS demonstrated increased positive teacher-child interactions and engagement, decreased challenging behaviors, and decreased negative teacher-child interactions relative to a BAU control group (Conroy et al., 2015). Similarly, Teacher- Child Interaction Training (TCIT), focused on improving teacher-child interactions and teaching teachers' skills to manage child behavior, has reported improvements in disruptive and prosocial behaviors (Campbell, 2011).

Although some Tier-2 practices and programs demonstrate positive effects, other Tier-2 programs produced mixed results. For example, Social-Emotional Learning Foundation is an intervention designed to allow teachers to extend the language and promote emotional and behavioral regulation while teaching early literary skills (Daunic et al., 2013). The intervention aims to improve children's social skills, behavioral regulation, competence, and reduce internalizing and externalizing problems. In a randomized controlled trial, it was found that the intervention significantly improved behavioral regulation, competence, and internalizing

problems compared to a BAU control group. However, no group differences were observed for social skills and externalizing problems (Daunic et al., 2013). Similarly, Banking Time is another early childhood intervention designed to improve children's externalizing behaviors and increase student-teacher relationship closeness through teachers providing one-on-one time with the child. Teachers observe a child's behaviors and expressed emotions, narrate the child's actions, and label the child's emotions and relational themes during one-on-one interactions (Williford et al., 2017). In a randomized trial, teachers were randomly assigned to one of the three conditions, Banking Time, a child time group (teacher spent individualized time with the child but were not instructed how to use that time), and a BAU control group. The study found reduced challenging behaviors in the Banking Time and child time groups. Teachers in the Banking Time condition reported significant improvements in children's externalizing behaviors, but there were no significant differences in children's observed positive engagement and behavioral control (Williford et al., 2017). When interventions do not produce the expected effects on child outcomes, questions are raised about improving Tier-2 supports for children in early childhood settings.

Identifying ways to improve the effectiveness of teacher-delivered Tier-2 programming offered to young children who display challenging behaviors in early childhood settings is vital for the families, teachers, and researchers who are invested in the SEB outcomes of these children. However, where can we support research efforts to help improve the SEB outcomes for these young children in early childhood settings? It is challenging to enhance the programming offered to children in early childhood settings because little is known about the practices teachers use in BAU to support children with challenging behaviors. Without the ability to describe BAU, the conclusions and action implications of effectiveness trials that produce mixed findings are limited. We do not know if new programs bring novel practices into the classroom or reiterate something teachers already do (Garland et al., 2010; Smith et al., 2017).

Many efforts designed to improve programming in early childhood settings do not consider the practices already being delivered. This approach can be costly and unproductive because when interventions do not produce optimal outcomes for children, it is hard to determine why they do not work. One way to address this problem is first to characterize what practices teachers deliver in BAU. Characterizing BAU is an important first step because it allows for a description of the practices used by teachers in BAU that can be used to inform efforts to improve children's SEB outcomes through a quality improvement process. Without the description of BAU, it is hard to determine how to improve BAU.

By characterizing BAU, schools can engage in a quality improvement process (i.e., continuous efforts to evaluate the performance of a system to identify ways to improve service delivery and outcomes; Riley et al., 2010). The first step of this process is to collect data on SEB problems, indicating a problem that needs to be addressed. The second step is to collect data on what practices teachers deliver in early childhood settings. The second part is missing as we do not know what practices teachers deliver in BAU early childhood settings for children who display challenging behavior. Without this information, engaging in a quality improvement process is challenging. Characterizing BAU can fill in the critical missing information. If data about children's SEB problems and teacher practices are collected, schools can create better services for young children with challenging behaviors by evaluating the current practices and examining how they can be improved.

Business-as-Usual Literature

To the best of my knowledge, no study has characterized the practices delivered by teachers in BAU with children who display challenging behaviors in early childhood settings. However, studies have examined practices delivered by BAU teachers for the whole class. Research has evaluated what practices found in the Pyramid Model (e.g., a tiered framework used to support social-emotional competence; Hemmeter et al., 2007) teachers delivered in the classroom using the Teaching Pyramid Observation Tool (TPOT; Branson & Demchak, 2011;

Snyder et al., 2013). The TPOT is an observational measure that assesses the presence and absence of practices or "red flags" for 21 items and rates 15 specific practices on a scale from 0 to 5, with 5 indicating all specific behaviors associated with the practice were observed.

To date, two studies have used the TPOT to examine practices delivered by teachers (Branson & Demchak, 2011; Snyder et al., 2013). Overall, these studies found teachers used 26 to 80% of the specific behaviors associated with practices, with an average of 50%. Only 3 out of the 50 teachers were implementing more than 70% of the specific behaviors, and only one teacher was implementing more than 81% of the specific behaviors (Snyder et al., 2013). In particular, results show teachers were appropriately responding to challenging behaviors, teaching social-emotional strategies and emotion labeling, and addressing challenging behaviors by discussing social-emotional development and informally planning to address challenging behaviors at home and school. Teachers were not posting rules or expectations, not consistently teaching social skills and emotional competencies, nor supporting children with challenging behaviors in assessing and developing a plan to address those problems. Overall, this study found that teachers used practices to build positive relationships with children and families but did not explicitly teach behavior expectations, social skills, or problem-solving skills (Branson & Demchak, 2011).

Although these studies examined practices delivered by teachers before they were trained on specific EBPs, they only examined practices found in the Pyramid Model, which gives a limited understanding of overall practices used in BAU classrooms. A more generalized tool capturing practices found in various EBPs and other practices used in classrooms can provide a better characterization of BAU. These studies also did not identify children at risk of developing challenging behaviors. Instead, the studies focused on identifying Tier-2 practices observed for all children in the classroom. When examining the practices delivered within the Tier-2 model, it is important to identify children at risk for developing behavioral problems and practices teachers use with those children to understand BAU classrooms better. Although it is helpful to

know if teachers used practices found in the research literature; it is also valuable to know how much they used the practice or at what dosage. The TPOT only provides an occurrence rating; it does not provide the dosage of the practices delivered. Therefore, a scale measuring the dosage of the practices delivered may be more informative about BAU classrooms. Knowing the dosage of practices used in BAU tells us how strong of dosage teachers are using for the practices they already use in the classroom for children with challenging behaviors.

In addition to the school literature, several studies in the mental health literature illustrate that there is utility in characterizing BAU. I summarized the usual care studies here because the methods used in these studies to examine practices delivered by usual clinical care (hereafter referred to as *usual care*) clinicians are similar to the ones I used in this study. These studies examined the frequency and dosage of the practices used for the target population. Garland et al. (2010) published one of the first studies describing usual care provided to children with disruptive behavior. The study aimed to identify similarities and discrepancies between usual care and evidence-based models (i.e., manualized interventions that include multiple practices; McLeod et al., 2017) for children with disruptive behavior (Garland et al., 2010). The study recorded usual care treatment sessions from six community mental health clinics and coded them using an adapted version of an observer-rated measure called the Practice and Research: Advancing Collaboration, Therapy Process Observational Coding System for Child Psychotherapy- Strategies scale (PRAC-TPOCS-S; Garland et al., 2006; McLeod & Weisz, 2010). The PRAC-TPOCS was designed to assess practices found in evidence-based programs for disruptive behaviors.

In 1,184 treatment sessions with 191 children and 96 clinicians, the study found that clinicians discussed therapeutic content areas (i.e., information, knowledge, or understanding that the clinician is trying to convey to the child, parents, or the family; Garland et al., 2010), and used therapeutic techniques (i.e., methods to convey therapeutic content; Garland et al., 2008) that aligned with the evidence-based model used for children with disruptive behaviors (Garland

et al., 2010). For example, affect education (i.e., understanding, identifying and labeling emotions, and recognizing environmental and physical cues of emotions; Garland et al., 2008) was observed in 81% of sessions, affect or anger management (i.e., a method to manage anger through perspective-taking, recognizing triggers and practicing relaxation skills; Garland et al., 2008) in 36% of sessions, and anticipating setbacks (i.e., predicting future reversals; Garland et al., 2008) in only 6% of sessions (Garland et al., 2010). This study found that usual care clinicians used various practices to treat children with disruptive behavior, but they used some practices more than others.

Similar findings were observed for usual care clinicians treating children diagnosed with autism spectrum disorder (Brookman-Frazee et al., 2010). This study also used the PRAC-TPOCS-S (Garland et al., 2006; McLeod & Weisz, 2010) to code 99 treatment sessions with 19 children diagnosed with an autism spectrum disorder to determine if clinicians were using practices found in evidence-based models to treat children with an autism spectrum disorder. The study found that usual care clinicians used practices found in evidence-based models. Out of the practices that aligned with evidence-based models, this study found that positive reinforcement (i.e., providing strategic attention; labeled praise, physical, verbal and material rewards; Garland et al., 2008) was used in 87% of coded sessions, modeling (i.e., therapist demonstrating skills through live, imagined or videotaped methods; Garland et al., 2008) was used in 48% of the sessions, role-play (i.e., practicing/rehearsing skills or reenacting hypothetical situations; Garland et al., 2008) in 34% of the sessions, and assigning or reviewing homework (i.e., assigning and reviewing task assigned to complete in the previous sessions; Garland et al., 2008) in only 18% of the sessions (Brookman-Frazee et al., 2010). Overall, this study found that usual care clinicians used practices aligned with evidence-based models; some practices were used more than others.

Herschell et al. (2019) similarly examined evidence-based and broader psychotherapy practices used to treat children at risk for physical abuse. This study used two observer-rated

measures, the Therapy Process Observational Coding System for Child Psychotherapy-Strategies scale (TPOCS-S; McLeod & Weisz, 2010) and the Alternatives for Families: A Cognitive Behavior Therapy Adherence Coding System (AF-CBT Adherence Coding System; Kolko,1996). The TPOCS-S was designed to assess various practices from five theory-based domains. AF-CBT Adherence Coding System was designed to evaluate adherence to AF-CBT, an evidence-based treatment for children at risk for physical abuse (Kolko & Swenson, 2002; Kolko, 1996). Both systems were used to code 278 treatment sessions to determine if usual care clinicians used elements of AF-CBT and practices from five theory-based psychotherapy domains to treat children at risk for physical abuse. The study found that usual care clinicians used practices found in AF-CBT and practices found in five theory-based psychotherapy domains to some extent (Herschell et al., 2019). For example, encourages affect (i.e., discusses and encourages a client to express affect; McLeod & Weisz, 2010) was observed in 65% of the sessions, general cognitive focus (i.e., extent to which the clinician uses cognitive interventions during the therapy session; McLeod & Weisz, 2010) in 11% of the sessions, role-play and modeling in only in 6% of the sessions (Herschell et al., 2019). Overall, this study found some overlap between usual care practices and practices found in AF-CBT, suggesting that usual care clinicians use practices found in the evidence-based model for children at risk of experiencing abuse.

Lastly, Smith et al. (2017) used 954 recorded treatment sessions to examine the extent to which clinicians trained in cognitive-behavior therapy in research settings, clinicians trained on cognitive behavior therapy in community settings, and usual care clinicians delivered cognitive-behavioral therapy to treat children with anxiety. Using the Therapy Process Observational Coding System for Child Psychotherapy-Revised Strategies scale (TPOCS-RS; McLeod et al., 2015), the study found that usual care clinicians used cognitive-behavioral techniques to treat children with anxiety but not as much as the research or community setting clinicians who were trained in cognitive-behavior therapy. Similarly, McLeod and Weisz (2010)

used the TPOCS-S to examine practices delivered to children with depression or anxiety in usual care. They found usual care clinicians used some practices found in cognitive and behavior therapy models (McLeod & Weisz, 2010). These findings are significant because it is often assumed that usual care does not contain practices found in evidence-based models (Garland et al., 2010). Identifying practices used in usual care helps address part of the quality improvement process. However, information about the dosage of the practices used will further assist in establishing the quality improvement process.

The usual care literature also used extensiveness, measured by considering the frequency and thoroughness of delivery throughout a treatment session (Hogue et al., 1996), to describe dosage of practices delivered. Assessing the extensiveness of delivery captures variation in dosage across clinicians, whereas considering delivery with a dichotomous variable (present/absent) does not capture variation (Hogue et al., 1996). These studies assessed extensiveness by rating each practice on a 7-point Likert scale (ranging from 0 to 6 or 1 to 7). Garland et al. (2010) found in almost half of the sessions; no single practice was used at high extensiveness for children with disruptive behaviors. The average extensiveness reported for all practices was 2.3 (*SD* = 0.30, range 0 to 6; Garland et al., 2010). Similarly, Brookman-Frazee et al. (2010) found that for children diagnosed with autism spectrum disorder, an average of low extensiveness was observed (*M* = 2.70, range 0 to 6; Brookman- Frazee et al., 2010). Interestingly, role-play had the highest average extensiveness of 3.20 (*SD* = 1.60) when observed, even though it was only observed in 34% of the 99 sessions. This finding suggests that when clinicians used role-play, they spent more time on the practice and reviewed it more thoroughly than other practices. Likewise, Herschell et al. (2019) observed low extensiveness for practices measured by the TPOCS-S that were part of AF-CBT for children at risk for physical abuse. For example, the average extensiveness for general cognitive focus was 1.20 (*SD* = 0.60), on an extensiveness scale ranging from 1 to 7 (Herschell et al., 2019; McLeod & Weisz, 2010).

Similarly, McLeod and Weisz (2010) found low extensiveness for children with internalizing problems. For example, the average extensiveness of respondent interventions (develop a hierarchy, use mastery ratings, and perform an exposure; McLeod & Weisz, 2010) was 1.10 (*SD* = 0.40). Altogether, an average of low to medium extensiveness was observed for practices delivered in usual care for various presenting problems. Since no study found high extensiveness for any practices used in usual care, it could indicate usual care clinicians are not consistently using these practices across sessions.

Describing practices used in usual care and their dosage is valuable in the quality improvement process because it provides baseline data that can be used to inform quality improvement efforts. Since practices were used at an average of low to medium dosage, the extensiveness of the practices can be improved by providing tools to increase the dosage through the quality improvement process. Understanding what practices teachers use and their extensiveness is essential for the quality improvement process; however, the current mental health usual care literature is missing a critical component about how practices are clustered. Characterizing how practices are clustered can give a comprehensive picture of what happens in BAU early childhood settings.

In addition to understanding what practices are used and their dosage, it is essential to know which practices teachers cluster. Teachers likely do not use a single practice but a combination of practices when interacting with young children who display challenging behaviors, as seen in Tier-2 programs (Feil et al., 2014; Vo et al., 2012). Thus far, no literature in education or mental health research has attempted to explore which practices are clustered. Knowing if teachers use specific practices when interacting with young children who display challenging behaviors is critical and will help inform quality improvement efforts by allowing them to consider what practices were clustered. Characterizing BAU early childhood settings by examining which practices teachers use, the dosage of those practices, and how teachers combine the practices will aid with efforts designed to improve outcomes for young children with

challenging behaviors. Knowing these three things can inform efforts to develop tailored programming to enhance care in early childhood settings.

Present Study

This study aimed to characterize the practices delivered by teachers in BAU early childhood settings for young children who display challenging behaviors. Characterizing BAU provided baseline data for change efforts and improved practice delivery as part of quality improvement efforts. To characterize BAU, data collected for the development and validation of the Treatment Integrity Measure for Early Childhood Settings measure (TIMECS; McLeod et al., 2021) is utilized. This study's sample and design were ideal for this research's aims. First, the child participants were 3 to 5 years old and who attended early childhood settings. Second, all the children displayed challenging behaviors that placed them at risk for emotional and behavioral disorders. Third, the TIMECS is ideally suited to characterize BAU in early childhood settings because it is a generic measure not tied to a particular evidence-based intervention. It is designed to capture practices in the research literature (McLeod et al., 2017). The TIMECS is a direct observation rating scale designed to assess the quantity (i.e., items that determine the extensiveness of practice delivery) of practices delivered by teachers to foster positive SEB outcomes for young children who display challenging behaviors in early childhood settings.

Using the TIMECS, the following hypotheses and research questions were evaluated in the current study:

1. Hypothesis 1. It was hypothesized teachers would use all practices when interacting with young children who display challenging behaviors.
2. Hypothesis 2. It was hypothesized that teachers would use practices at an average of above *not at all* (i.e., above a score of 1) to *considerably* (i.e., score of 5) extensiveness across all observations.
3. Research Question 1. Since the proposed study was the first to ask how teachers cluster the practices, a hypothesis was not drawn. The proposed study thus sought to answer

the question: which practices were teachers clustering when delivering practices in BAU settings for children who display challenging behaviors?

Method

Data Source

The data for the current study was drawn from an Institute for Education Sciences Measure Development (R305A140487; PIs McLeod & Sutherland) study designed to develop and validate the TIMECS (McLeod et al., 2021). Demographic and children's challenging behavior data were collected from the teachers before the observations. Observational data were collected using the TIMECS.

Participants and Settings

The participants in this study included teachers and children in early childhood settings in a Southeastern state. Early childhood settings included federal or state-funded programs located in urban and suburban communities. On average, early childhood classrooms consisted of 17.26 (*SD* = 3.54) children and 2.09 (*SD* = 0.29) adults. Teachers reported using various comprehensive curricula in their classrooms to support social-emotional competencies, such as the High Scope Early Childhood Curriculum (Schweinhart & Weikart, 1997). A few teachers also implemented targeted SEB learning curricula.

Teachers were contacted each school year in October to participate in the study. Contacting teachers in October allowed teachers to settle into their new class and identify students who had the most severe and chronic challenging behaviors. Teachers identified up to 5 focal children in their classroom who demonstrated the most severe and chronic problem behaviors using the first stage of the *Early Screening Project* (ESP; Walker et al., 1995). After the children were nominated, caregivers were contacted by the teachers to obtain parental consent. Once consent was received, the teachers completed the second stage of ESP, the Externalizer Teacher Questionnaire, to confirm the challenging behaviors. Lastly, *Battelle Developmental Inventory, Second Edition Screener* (Newborg, 2005), was used for each

nominated child to demonstrate average or above-average cognitive abilities. Children with the highest scores on ESP and who met the average or above-average cognitive abilities criteria were identified as the focal children in this sample.

Teachers Participants. The study included 54 teachers (94.4% 26-55+ years old; 92.6% female, 7.4% male; 61.1% White, 35.2% Black, 3.7% multiracial) who averaged 7.70 (*SD* = 8.00) years of teaching in early childhood settings. The teachers held varying degrees; 38.9% had a Bachelor's degree, 48.1% had a Master's degree, and 13.0% had other degrees. See Table 1 for more information about teacher participants.

Of the 54 teachers, 17 had previously been trained on BEST in CLASS, a Tier-2 program developed to increase positive interactions between children and teachers. Training on BEST in CLASS included a 6-hour professional development workshop in which teachers about BEST in CLASS and how those practices can be implemented with the focal students (Conroy et al., 2019). Teachers also received 14 weeks of coaching to ensure they implemented BEST in CLASS practices with high quality in their classrooms. Training and coaching focused on six modules: (a) Rules, Expectations, and Routines; (b) Behavior Specific Praise; (c) Precorrection and Active Supervision; (d) Opportunities to Respond and Instructional Pacing; and (e) Instructive and Corrective Feedback. The final module on Linking and Mastery assisted teachers in delivering practices efficiently with proficiency (Conroy et al., 2015).

After the initial recruitment, 21 teachers were not included in the study because children in their classroom did not qualify for the study, the caregiver did not return the consent form, or teachers withdrew from the study. Some teachers withdrew without a reason, and others expressed they were no longer interested in the study, were too overwhelmed by other responsibilities to participate, or were on medical leave.

Child Participants. The study included 91 children (25.3% female, 54.9% male, 19.8% not reported); average age 4.31 (*SD* = .67) years old; 45.1% Black, 8.8% White, 1.1% Native American/American Indian, 1.1% Asian/Pacific Islander, 5.5% multiracial or other, 38.4% race

was not reported, and 7.7% Latinx ethnicity. See Table 2 for more information about child participants.

Measurement

Treatment Integrity Measure for Early Childhood Settings (TIMECS; McLeod et al., 2021). TIMECS is a 21-item direct observation rating scale designed to assess the quantity and quality of practice elements (i.e., discrete skills or principle; McLeod et al., 2017) delivered by teachers to foster positive SEB outcomes for children in early childhood settings (McLeod et al., 2021). The items on the TIMECS were developed using the following methods. First, an iterative process was used to establish content validity. A literature search identified 49 studies evaluating models and practices delivered in early childhood settings. Then individual practices were extracted from each study and grouped. Practices that targeted the same domain of child functioning were combined to form a single practice element. For example, practices such as “labeled praise” and “behavior-specific praise” were grouped under “praise” (see McLeod., 2017). This process resulted in a total of 24 practice elements. Then five expert raters were asked to rate each practice element's relevance in addressing SEB problems in young children. They were asked to evaluate if the practice was "not necessary," “useful” but not essential, or “essential” to use in early childhood settings to address SEB problems in young children. The practices were retained if 75% of the raters rated them as useful or essential (McLeod et al., 2017). Out of the 24 items, only 21 that could be observed were kept as part of the TIMCES measure.

This study focuses on quantity items of the TIMECS because quantity items were designed to capture the breadth and depth of the practice elements delivered (McLeod et al., 2021), which were ideal for addressing the aims of this study. The quantity items consist of content items and delivery items. Content items were defined as the principle that guides a practice (McLeod et al., 2017) and include five items: Social Skills, Emotion Regulation, Problem-Solving, Promoting Behavioral Competence, and Teacher-Child Relationship (see

Table 2 for definitions). The delivery items were the way a teacher provides instruction (McLeod et al., 2017) and include 16 items: Narrating, Supportive Listening, Choices, Monitoring, Rules, Modeling, Rehearsal, Precorrection, Opportunities to Respond, Visual Cueing, Premack Principle, Tangible Reward, Time-Out, Praise, Error Correction, and Instructive Feedback (see Table 2 for definitions).

The quantity items were scored on a 7-point Likert extensiveness scale with the anchors of extensiveness: 1 = *not at all*, 3 = *somewhat*, 5 = *considerably*, and 7 = *extensively*. Extensiveness scoring captures the degree to which teachers deliver specific practice elements during the entire observation through thoroughness and frequency (McLeod et al., 2021). Thoroughness was defined by: “(a) the concentration of effort or commitment the teacher puts into the practice; (b) the detail in which the teacher describes the practice; (c) the depth or intensity of the practice; (d) the extent to which the teacher followed through with the practice; or (e) the extent to which the practice was pursued intensively across an observation” (McLeod et al., 2021, pg. 4). Frequency was defined as the number of instances a teacher used a specific instructional practice during the observation regardless of the thoroughness of the practice (McLeod et al., 2021). Both thoroughness and frequency were considered when giving the extensiveness rating. For example, it was possible to score "7" on *Emotion Regulation* based on the teacher encouragement of the child to label their emotions throughout the observation (i.e., frequency); however, it was also likely that a "7" can be given based on the teacher encouragement of the child to use emotion labeling and assisting the child in regulating their emotions thoroughly during a specific interaction (i.e., thoroughness). The extensiveness rating provides quantity or dosage information for each practice.

Coding Procedures

Coders. Seven doctoral students and five data staff members (*M* age = 25.92 years old, *SD*= 3.90; 83.0% female, 17.0% male; 58.0% White, 25.0% Latinx, 8.0 % Asian, 9.0% multiracial) were trained on TIMECS.

Coder Training. Training started with didactic instruction and discussion of the manual, then coding exercises designed to expand understanding of each item led by the Principal Investigators. After the initial training, coders independently coded recorded observations and discussed the results at the weekly meetings. As the final step, coders independently scored 40 recordings, and their codes were compared against the main codes created by the Principal Investigators. To be certified for independent coding, each coder had to meet "good" score reliability on each item, defined as an intraclass correlation coefficient (ICC)[2,2] ≥ .60 (Cicchetti, 1994). Once coders were trained, regular reliability assessments were performed and discussed at the weekly meetings to prevent coder drift (Margolin et al., 1998).

Positionality. Following the guidelines established by Quantitative Critical Race Theory, researchers are encouraged to reflect on the potential impact of the team members' identities on the results (Castillo & Gillborn, 2022). The coding team was trained by experienced faculty members and included graduate students and data staff. The team was diverse in gender representation and racial/ethnic backgrounds. Although team's gender and racial/ ethnic identities may influence the results, we took steps to ensure the trustworthiness of the results in the section below (Zuberi & Bonilla-Silva, 2008).

Trustworthiness of the Data. To ensure the trustworthiness of the data, this study engaged in multiple recommended practices (Patton, 1999; Connelly, 2016). For example, coders received intensive training where they coded recorded observations and discussed the results in weekly meetings. Coders independently scored 40 recordings to be certified. Additionally, two coders were sent to collect data during each observation. The coders were recommended not to talk to each other, teachers, or students as that may influence the data quality. Further, regular reliability assessments were performed and discussed at weekly meetings to prevent coder drift. These procedures helped ensure data quality and reliability; however, individual bias may still play a role during coding of the observations. In particular, stereotypical bias about Black women may play a role in coding of practices, including

perceiving Black women as more aggressive and hostile (Ashley, 2013). These biases may influence how coders code the positive interaction between teachers and children during the observations. These biases may impact the frequency and dosages of practices observed. Further influence of identities is discussed in the discussion.

Classroom Observation Coding. Observations were assigned using a balanced incomplete block design (Fleiss, 1981), stratifying for classroom and time. The aim was to schedule up to 12 observations (M = 7.14, SD = 1.97) for each focal child. Observations were scheduled during the week at a convenient time for the teacher. Two coders were sent to each class and instructed to observe each focal child separately. Observations were assigned randomly if a teacher had multiple focal children in the classroom. The coders were instructed to sit in a discrete location in the classroom and not interact with teachers, students, or each other. Coders took notes throughout the observation and, at the end of the observation, scored the 21 TIMECS items. Observations occurred during a teacher-led instructional activity (e.g., small group, circle time, or story time), child-led activity (e.g., center time), or transitions. The observation could have occurred during multiple instructional contexts. Most observations were during a teacher-led instructional activity, where the teacher directly engaged with the focal child or the focal child's group. The observation had to be at least 30 minutes long, and the data were discarded if an observation was less than 30 minutes long. The average observation was 40.56 (SD = 11.08) minutes.

Results

The purpose of the current study was to use the TIMECS items to characterize BAU in early childhood classrooms, by understanding what practices were used, the extensiveness of scores on the TIMECS items, and how scores on the items co-occurred during observations. Preliminary analyses focused on interrater reliability and descriptive statistics for all TIMECS items, and sample bias analyses were conducted for items that mapped on to BEST in CLASS content. Primary analyses focused on calculating the percentage of items observed across

observations, estimating the average extensiveness of each item score, and conducting an exploratory factor analysis (EFA) to estimate how item scores clustered together.

Preliminary Analyses

Interrater Reliability. Interrater reliability for coder scores on each item was calculated using the intra-class correlations (ICC (2,2), Shrout & Fleiss, 1979). Following the Cicchetti (1994) standard, ICC (2,2) below .40 reflect "poor" agreement, from .40 to .59 reflects "fair" agreement, from .60 to .74 reflects "good" agreement, and above .75 reflects "excellent" agreement. The ICC (2,2) was appropriate for this study because it represents the two-way random-effects models, which provide reliability estimates for all coders and allows for generalizability of findings to other samples.

Interrater reliability was calculated for the 21 quantity TIMECS items. A total of 650 observations independently scored by two coders were used for the reliability analyses. ICC (2,2) values for the items ranged from .68 to .95 (M = .81, SD = .07; see Table 4). The ICC (2,2) for 18 of the 21 items fell within the "excellent" range, and three items fell within the "good" range. These findings suggest interrater reliability for items was in the good-to-excellent range (Cicchetti, 1994).

Descriptive Statistics. As shown in Table 4, descriptive statistics, including mean, standard deviation, range, skewness, and kurtosis, were calculated for the scores on the 21 TIMECS items. Item scores were computed by averaging across the two coder's ratings. Mean scores on the TIMECS items ranged from 1.04 (SD = .25; Time-out) to 4.80 (SD = .92; monitoring). The range of scores for the items went from 2.50 to 6.00. Item scores were considered to be normally distributed if skewness and kurtosis were between -2 to 2 (George & Mallery, 2019). Skewness values for the 21 items ranged from -.49 (SE = .10; Opportunities to Respond) to 7.66 (SE = .10; Time-out), whereas kurtosis values ranged from -.64 (SE= .19; Praise) to 66.55 (SE = .19; Time-out). Seven item scores had skewness above 2 (Emotion Regulation, Problem Solving, Choices, Rehearsal, Premack Principle, Tangible Reward, Time-

out), and 8 item scores had kurtosis above 2 (Emotion Regulation, Problem Solving, Supportive Listening, Choices, Rehearsal, Premack Principle, Tangible Reward, Time-out). Based on these descriptive findings, scores for 13 items were normally distributed and scores for 8 items were not normally distributed.

Sample Bias. Sample bias analyses were conducted to determine whether teachers who were previously trained to deliver BEST in CLASS differed from the remaining teachers in scores on the TIMECS items that map onto the content of the BEST in CLASS program (Rules, Precorrection, Opportunity to Respond, Praise, Error Correction [i.e., Corrective Feedback], and Instructive Feedback). A series of independent t-tests were performed to compare scores on the six TIMECS items from teachers in the BAU and BEST in CLASS groups. Teachers who had received BEST in CLASS training and coaching had significantly higher scores on items, Rules, Precorrection, Opportunity to Respond, Praise, and Error Correction and no group differences were observed for scores on item Instructive Feedback (see Table 3). These results showcase those teachers who had been previously trained on BEST in CLASS had significantly higher scores on five out of the six items that align with BEST in CLASS content. Although there were differences in the groups, the subsequent analysis were reported for the whole sample because that was the main focus of this study. For this study, BAU is defined as existing practices teachers use without exposure to new programming at the time of the study. Teachers that were trained on BEST in CLASS, were trained couple years before this study. Therefore, that training is considered part of the BAU. Additionally, teachers in early childhood settings receive various training on curriculums. Even in this study, teachers reported using other comprehensive curricula in their classrooms to support social-emotional competencies, such as the High Scope Early Childhood Curriculum (Schweinhart & Weikart, 1997) and few teachers also implemented targeted SEB learning curricula. Therefore, teachers receiving training on other programs is part of the BAU for early childhood settings.

Primary Analyses

Practice Use. It was hypothesized teachers would deliver all TIMECS items when interacting with young children who display challenging behaviors. The hypothesis was evaluated by examining the percentage of item scores as present (>1) across all observations (*N* = 650). For each observation, a TIMECS item was counted as observed if scored above 1. To estimate a percentage of items observed across observations, the number of times an item occurred was divided by the total number of observations (*N* = 650). For each item, the hypothesis was considered supported if the percentage was above zero percent.

As seen in Table 5, the average percentage of items observed across the 650 observations on the TIMECS items was 49.2% (*SD* = 32.2), with a range of 2.9% to 99.7%. Six of the 21 TIMECS items were observed in fewer than 25% of the 650 observations. Six TIMECS items were observed between 24.9% to 50% of the 650 observations. Three items were observed between 49.9% to 75%, and five items were observed in 75% or more of the 650 observations. These results indicate all TIMECS items were observed, thus supporting the hypothesis.

Extensiveness of Practices. It was hypothesized that item scores would range from 1 to 5 on a 7-point extensiveness scale across observations. To evaluate the hypothesis 2, I calculated descriptive statistics, including means and standard deviations for scores on the 21 TIMECS Quantity items across all observations (*N* = 650). If, on average, item scores across the 650 observations were above a “1” and below “5” on the extensiveness scale, then the hypothesis was considered supported. As seen in Table 5, the mean score for the 21 TIMECS items was 2.03 (*SD* = .38) across the 650 observations. The mean scores on the items ranged from 1.04 (*SD* = .25; Time-out) to 4.80 (*SD* = .92; Monitoring). These results indicate that mean scores on items were between 1 to 5 on the 7-point extensiveness scale, thus supporting the hypothesis.

Since each item was not observed in all observations, including scores of 1s (i.e., not observed) when calculating the item means may underestimate extensiveness. Therefore, item means and standard deviations were calculated with the scores of 1 removed. As seen in Table 5, with the 1s removed, the mean score on items ranged from 1.89 (*SD* = .39; Choices) to 4.80 (*SD* =.92; Monitoring), with the mean scores increasing for items used at a lower percentage. For example, Time-Out, which was observed in 2.9% of the 650 observations, increased from *M* = 1.04 (*SD* = .25) to *M* = 2.34 (*SD* = .67). Similarly, Tangible Reward, which was observed in 8.8% of the observations increased from *M* =1.11 (*SD* = .39) to *M* = 2.25 (*SD* = .57). However, even with the 1s removed, the mean scores on all items remained below 5. These results showed that when not observed scores were removed, the mean use of items did not increase above 5. Therefore, removing the not observed scores still supports the hypothesis that items were observed between 1 to 5 on the 7-point extensiveness scale.

Clusters of Practices. Since the current study was the first to ask how teachers clustered items together when observed in a BAU setting, a hypothesis was not drawn. This study thus sought to answer the following research question: which items clustered together when teachers interacted with children who display challenging behaviors? To characterize which items clustered together in BAU, an exploratory factor analysis (EFA) was used to describe the factor structure of the scores on the 21 TIMECS items.

Before conducting the EFA, interrater reliability, item distribution, and inter-item correlations were examined to ensure the items were appropriate for an EFA. Following Hogue et al. (2019), items were considered for exclusion if the ICC (2,2) value was < .30. Item distribution was examined by checking item skewness and kurtosis. Following the guidelines by Tabachnick and Fidell (2013), items were considered for exclusion if the absolute value of skewness was greater than 3 or the absolute value of kurtosis was greater than 10. The correlation matrix was examined to assess for multicollinearity which is indicated by highly correlated items (i.e., >.90; Hair Jr. et al., 2021).

Interrater reliability was examined with ICC (2,2). As reported earlier, the ICC (2,2) for all TIMECS items ranged from good-to-excellent. Therefore, all items met the ICC criteria. Six items had out-of-range skewness and kurtosis and thus were removed before analyzing the correlation matrix (see Table 6). Of the 15 items that met the inclusion criteria for normality, the correlation between the TIMECS items ranged from r = -.07 to .69, which is acceptable. These findings suggest that 15 items were appropriate for EFA because they met the ICC criteria, were normally distributed, and did not show multicollinearity.

MPLUS (Version 8; Muthen & Muthen, 2017) was used to perform the EFA. Factor analysis extraction maximum likelihood method was used with oblique (direct oblimin) rotation, and the sandwich variance estimator was used to account for the nested data structure. Factor analysis was used as this approach allows for a wide range of goodness of fit indexes and statistical significance testing of factor loadings and correlation among factors (Fabrigar et al., 1999). Maximum likelihood is the recommended method when data is normally distributed, which is the case for this dataset (Fabrigar et al., 1999). Oblique rotation (direct oblimin) was used because it allowed the factors to be correlated (Costello & Osborne, 2005). Using oblique rotation provides a more accurate representation of data examining behavior that has not been studied before and of behavioral data that is likely to occur in relation with one another (Hogue et al., 2019). Since multiple observations were collected per teacher and children, the sandwich variance estimator was used to account for nesting of observations within children. The sandwich estimator provided consistent variance estimation for dependent data (Kauermann & Carrol, 2001).

Two decision rules were used to extract the number of factors: an examination of factors with eigenvalues >1.0 (the Kaiser- Guttman rule) and a scree test (Kaiser, 1991). Literature suggests multiple approaches should guide factor extraction (Hair et al., 1995; Thompson & Daniel, 1996). Using the scree test, eigenvalues were plotted on the y-axis, and factor numbers were listed on the x-axis. Factors were listed in decreasing order of their eigenvalues. All factors

above the inflection point (where the curve starts to level off) were retained, identifying the number of factors for this analysis. After the number of the extracted factors was determined, items were trimmed if they did not load at least |0.32| onto one of the factors (Tabachnick & Fidell, 2001). Items cross-loadings were examined. The items that loaded onto multiple factors were examined carefully; items were assigned to a factor with the highest loading; if the difference between the factor loading was below .30 or EFA was run with fewer number of factors to examine if the cross-loading still appeared, if they did then the item was dropped. (Costello & Osbourne, 2005; Tabachnick & Fidell, 2013). The final factor solution was based on the steps described above; three to 10 items were expected to load onto a factor to make a meaningful interpretation (Mvududu & Sink, 2013). The factors were interpreted and labeled based on the conceptual ideas they represented. Eigenvalues and percent of variance for each factor were reported. Fifteen items that met the criteria for normality and multicollinearity were included in the EFA (See Table 6 for included items).

The initial examination of factors with eigenvalues >1.0 and the scree plot test suggested a three-factor solution. The eigenvalues for the three-factor solution were 4.82 for factor one, 1.92 for factor two, and 1.27 for factor three, accounting for 32.1%, 12.8% and 8.5% of the variance. With the three-factor solution, the Social Skills and Emotional Regulation items were dropped because they did not load at least .32 on any factors (Tabachnick & Fidell, 2001). This solution did not show any cross-loadings.

EFA was rerun with 13 items and suggested a three-factor solution. The eigenvalues for the 13 items model were 4.66 for factor one, 1.83 for factor two, and 1.09 for factor three, accounting for 35.8%, 14.1% and 8.4% of the variance. All items loaded at least .32 on to one of the three factors. However, the three-factor solution showcased cross-loading for items Rules and Narrating. Therefore, the two-factor model was run, and this solution did not showcase any cross-loadings; thus, Rules and Narrating were retained in the final two- factor solution.

The final solution resulted in a two-factor solution that exhibited no significant cross-loadings. The two-factor solution had eigenvalues of 4.25 for factor one and 1.78 for factor two and accounted for 70.48% and 29.52% of the variance, respectively. The final two-factor solution included 13 items (see Table 7). Factor one and Factor two correlated r = .17.

Table 7 displays item loadings for each factor. Factor one was named classroom management and included 9 items: Promoting Behavioral Competence, Rules, Narrating, Monitoring, Modeling, Precorrection, Opportunities to Respond, Visual Cueing, Praise, Error Correction, and Instructive Feedback. Factor one loadings ranged from .460 to .806. Factor two was named supportive relationship and included 2 items: Teacher-Child Relationship and Supportive Listening. Factor two loadings were .709 and .899. Although it is recommended that factors comprise of at least 3 or more items, the two-item factor can still be considered reliable if two items are strongly correlated (r = .66) and not highly correlated with other items, which is the case for the two items in this dataset (Tabachnick and Fidell, 2013). Therefore, the two-factor solution provides insight into which items were clustered in observations.

Discussion

Developing a better understanding of the practice elements teachers use in BAU to support young children who display challenging behaviors may identify ways to improve the quality of care in early childhood classrooms. The present study aimed to characterize BAU by describing what practice elements teachers were using, the dosage of those practice elements, and how they grouped practice elements. The results showcase that teachers used various practice elements to support children who display challenging behaviors in early childhood settings. The findings also suggest that the dosage of practice elements varied across observations. Lastly, the results indicate that teachers grouped practice elements into two groups: those that help promote classroom management and those that support teacher-child relationships. These findings suggest that BAU teachers used practice elements in the literature to support children with challenging behaviors.

The study's first aim was to identify the practice elements teachers delivered in early childhood settings with children who display challenging behaviors. The results showed that teachers used all of the practice elements measured by the TIMECS to support children who display challenging behaviors. There was variation in how frequently practice elements were observed, with some practice elements used in most observations and others in very few. Trends appeared in how often practice elements were delivered. In this sample, the BAU teachers frequently used practice elements (e.g., Opportunity to Respond, Praise) shown to reduce challenging behaviors and improve on-task behavior and can be used across different learning contexts within the classroom (e.g., large groups, individual interactions; Floress et al., 2017) to engage the student so they can be used in academic lessons and other activities (Fullerton et al., 2009; Smith et al., 2011; Sutherland et al, 2002). Since these practices were observed in the majority of the observations, it suggests most teachers were using these practice elements in their classrooms with children who display challenging behaviors. These practice elements characterize BAU for children with challenging behaviors.

In contrast, other practice elements found in social-emotional learning (e.g., Problem Solving, Emotion Regulation) were not characteristics of practices used by BAU teachers because these were observed in very few observations. Additionally, practice elements such as Time out which are often used in response to a child's behavior, may not be characteristics of practices delivered in BAU classrooms because they were rarely used during the observations (O'Handley et al., 2019). These other practice elements do not characterize BAU for children with challenging behaviors.

Some of the study findings align with previous research in school settings and the mental health literature. These findings were consistent with previous research findings that show teachers used practices that build positive relationships (e.g., Teacher-Child Relationships). They were infrequently using social-emotional learning practices (e.g., Problem Solving and Social Skills) and expectations (e.g., Rules, Precorrection) during the observations

(Branson & Demchak, 2011). In contrast to previous research, teachers in this sample were observed to use emotional regulation less (Branson & Demchak, 2011). The discrepancy in findings may be due to the difference in the samples of the study; Branson and Demchak's (2011) study examined the delivery of the practices with the whole class, while this study only examined the delivery of practices with children who displayed challenging behaviors. The consistencies in findings show that certain practices are used with the whole class and with children who display challenging behaviors in early childhood classrooms. Similarly, some practices are not frequently used with the whole class or for children with challenging behaviors. In contrast, the discrepancy in findings may show that certain practices may be used with the whole class and not for children with challenging behaviors. Other reasons for discrepancy are discussed in the limitations below.

Similar to this study, usual care therapists were observed using practice such as praise for children with disruptive behavior (Chorpita & Daleiden, 2009; Garland et al., 2010). Similarly, parents trained on Parent-Child Interaction Therapy (PCIT) were also observed using praise for children with disruptive behavior (Abrahamse et al., 2016). This suggests practice element like praise may be applicable across contexts (i.e., home, classrooms and therapy) for children who display challenging behaviors and is characteristic of practice delivered across BAU settings. However, contrary to this study, usual care therapists were more likely to use problem solving, punishment, and tangible rewards than BAU teachers. The usual care therapists were also less likely to use modeling and building relationships than BAU teachers. The difference usage in practices between the two settings suggest some practices may be more likely to be delivered in one setting versus another. Suggest BAU classrooms look different from usual care therapy sessions. The discrepancies between the two studies may exist because the focus and training of teachers versus clinicians are different (i.e., academic achievement versus reduction in challenging behaviors). The study sample also differed in that the usual care literature included children who meet the diagnostic criteria for disruptive behavior disorders. In contrast, this

sample only included children who displayed challenging behaviors. Overall, this suggests that practices delivered in usual care therapy session look different from practices delivered in BAU classrooms.

The study's second aim was to examine the dosage of practice elements delivered by teachers in BAU classrooms. The results show that teachers delivered practice elements at varying dosages. To receive a high "dosage," teachers had to score 6 or 7 on the TIMECS item (McLeod et al., 2021), which none of the practice elements achieved. Practice elements with the highest dosage (e.g., Opportunities to Respond, Monitoring) were also observed most frequently across observations, even when taking out observations when a practice did not occur. When observed, teachers delivered these practices at higher frequency and thoroughness than other practices. While dosage remained low for other practice elements (e.g., Choices, Premack Principle) when observed. Other practice elements, such as Emotional Regulation and Social Skills, were observed at a lower dosage, which may be because these practice elements are typically delivered as a part of a structured curriculum (e.g., practices delivered at specific time during the week). Therefore, depending on the time of the observation, some of these practices may not be observed if they were delivered. Practice elements like Emotional Regulation and Social Skills were shown to be beneficial in decreasing disruptive behaviors and increasing pro-social and on-task behaviors (Ocasio et al., 2015; Upshur et al., 2017). However, teachers may need help generalizing these practices to the classroom because these practices are typically taught with a structured curriculum (Duncan et al., 2013). Lastly, teachers did not use certain practices at a higher dosage which may be appropriate for the setting and the population. Practice elements such as Time Out are recommended to be used selectively in the classroom management curriculums (Reinke et al., 2012). Overall, these findings indicate that BAU teachers use practices at varying dosages for children with challenging behaviors.

The findings from the current study regarding dosage of practice elements delivered aligns with some literature examining practices delivered by school teachers. Previous research has found that teachers use Opportunity to Respond at a high rate, similar to this study (Scott et al., 2011). Some of the findings from this study do not align with previous literature that found that teachers were posting class rules at high rates and using praise at low rates (Reinke et al., 2013). These inconsistencies may exist because previous studies only looked at rates per minute rather than the dosage, which includes frequency and thoroughness within a given observation. The studies in school literature were also completed with a whole class rather than children who display challenging behaviors. School literature suggests that increasing the dosage for some practice elements, such as Opportunity to Respond and Praise, is beneficial for students to decrease disruptive behavior and improve positive behaviors (Sutherland et al., 2003; Sutherland & Wehby, 2002). However, other research suggests that higher usage and dosage of other practice elements, such as Error Correction, are associated with less child academic engagement (Collier-Meek et al., 2019); therefore, not increasing the dosage may be appropriate. Overall, the findings from this study suggest that increased dosage may be beneficial for some practices while sufficient as it is for others.

Similarly, some of the findings align with practices used by usual care clinicians treating children with disruptive behavior problems in community mental health settings. Except for praise and relationship building, other practices found in usual care clinics, such as modeling, reinforcements (tangible rewards), punishment, and problem-solving, were delivered at relatively low dosages, similar to mental health studies (Garland et al., 2008). Along with therapist, parents receiving PCIT also increased their rates of praise (Abrahamse et al., 2016). These findings suggest that certain practices may be applied to children who display challenging behaviors in various settings, while others may be more appropriate for classroom settings. Compared to teachers, clinicians may receive more training in practices directly focused on addressing challenging behaviors. These findings suggest that teachers may be

trained differently in certain practices than clinicians, which may be appropriate for a given setting.

The final aim of this study was to examine how teachers combine practice elements when interacting with young children who displayed challenging behaviors. As evidenced by Tier 2 programs (Feil et al., 2014; Vo et al., 2012), practice elements are often implemented in a package; therefore, examining which practices teachers cluster is crucial. The EFA was focused on 13 items that yielded two factors. The first factor included 11 practice elements and was labeled classroom management. The second factor was defined by two items and was labeled supportive relationships.

The 11 items included in the first factor labeled as classroom management focused on using clear expectations (i.e., Rules), effective instructional management (i.e., Opportunities to Respond, Monitoring), reinforcing positive behavior (Promoting Behavioral Competence, Praise), and addressing inappropriate academics responses or behaviors (i.e., Error Correction, Instructive Feedback) and engaging children (i.e., Visual Cueing, Narrating, and Modeling). These practice elements can be used to address academic needs and behavior disruption in the classroom. Previous research indicates that interventions that contain some or all of these practice elements help reduce challenging behaviors (Feil et al., 2014; Simonson et al., 2008).

The composition of factor one indicates that teachers use practices aligned with classroom management practices found in the literature, with a few caveats. The teachers used clear expectations, effective instructional management, reinforcing positive behavior, and addressing inappropriate behavior together, consistent with previous findings on classroom management (Feil et al., 2009; Simonson et al., 2008; Sprague & Golly, 2013). In addition, teachers use other practices such as narrating, monitoring, modeling, and visual cues with other practices found in classroom management. Interestingly, this study did not find teachers using minimizing attention for minor inappropriate behaviors and enforcing clear consequences for unacceptable behaviors typically accomplished through Time-out and other practices (Feil et al.,

2009; Sprague & Golly, 2013). That could be because the practice is only observed in a few observations and, therefore, not included in the factor analysis because it is not normally distributed. Additionally, the items on the TIMECS measure did not cover ignoring minor inappropriate behaviors, which may be more difficult to observe than other practices included in the TIMECS.

The second factor was labeled supportive relationships and includes two practices focused on building Teacher-Child Relationships and Supportive Listening. Previous research found that improving the teacher-child relationship has positive outcomes for children who display challenging behaviors (Hamre & Pianta, 2006). By building a relationship with children, teachers may show warmth and closeness when interacting with young children who display challenging behaviors. Previous literature indicates teacher-child supportive relationships, especially closeness, warmth, and low- level of conflict, have increased learning engagement, academic performance, and decreased challenging behaviors in the classroom (Quin, 2017; Roorda et al., 2011; 2017).

The findings from factor two align with practices in programs focused on building supportive relationships between teachers and children. Similar to previous research, this study found teachers were using Teacher-Child Relationship and Supportive Listening to engage with children who display challenging behaviors similar to practices found in evidence-based programs like Banking Time and Teacher-Child Interaction Training (Driscoll et al., 2011; McIntosh e al., 2000; Williford et al., 2017). Contrary to previous research on supportive relationships, such as through programs like Banking Time, teachers in this study were not observed labeling emotions nor narrating (Driscoll et al., 2011; Williford et al., 2017). Instead, in this study, practices such as monitoring and narrating were more likely used along with other classroom management practices. The discrepancy may exist because programs like Banking Time focus on one-on-one interaction between teachers and children. For this study, teachers may have been using these practices with individual children or in group settings.

Since factor two, supportive relationships, only included two items, the results should be interpreted cautiously. The literature suggests that a factor should consist of three or more items to be considered reliable (Tabachnick & Fidell, 2013). However, two-item factors can be regarded as reliable if the two items are highly correlated but not with other items, which was the case for the two items that mapped onto the supportive relationships factor (r= .66, see Table 6).

The two-factor solution indicates that teachers use those practice elements cluster together or that coders group them when observing the teachers. This can occur because certain practices may follow one another. For example, Opportunities to Respond can often be followed by Praise (Sutherland & Wehby, 2002). Therefore, coders may have been observing them together. Another reason why certain practices were observed clustering may be because the similarity in item definitions. For example, Supportive Listening is defined as actively understanding the topic the child is discussing, and that can also be conveyed with warmth, closeness, and interest when listening and interacting, which is part of the Teacher-Child Relationship item. Therefore, item definition may make it difficult to decipher between the two items. Overall, these findings suggest that teachers in BAU classrooms were clustering practice elements aligned with practices found in Tier 2 interventions to support children who display challenging behaviors.

The results show teachers use several practice elements in BAU classrooms already, indicating teachers may already know how to use those practices. Therefore, when new programs are implemented in schools, they may not need to introduce certain practices to teachers. Instead, the new programs should use teachers' current knowledge of the practices to promote continuous use with children who display challenging behaviors. The findings of this study indicate that teachers were less likely to use practice elements found in social-emotional learning curriculums; therefore, it may help to highlight how certain practices can be generalized across the classroom. Teachers were also grouping certain practices, which has implications for

teacher training which could leverage the practices teachers were already grouping together. Training efforts could focus on helping teachers use practices in conjunction. These findings provide baseline data about practice elements teachers may be using in their classrooms.

Study Strengths

This study has several strengths. A strength of this study was using TIMECS to measure practices elements delivered by teachers in BAU classrooms. TIMECS provides a better characterization of BAU as this measure is not tied to a particular program, allowing it to more broadly capture practice elements found in the literature. Another strength of the study is the observations. Observational coding is a strength because it is considered the gold standard, often viewed as less biased than self-report (Gershman et al., 2017; Sanetti & Collier-Meek, 2019). The study also includes a large sample size of 91 children and 650 observations, which allows for a better description of BAU classrooms.

Limitations

This study also has some limitations. Although observations were considered the gold standard, there are some drawbacks to using observations. Although the study coded an average of seven observations per child, observations may still not provide a complete picture of what is happening in the classroom as teachers were only observed for a limited time throughout the school year. Observational data may only capture some practice elements teachers use daily in their classrooms. Depending on when the observation was conducted, teachers may not have been using certain practice elements such as rules and precorrections, as these practice elements typically occur at the beginning of the day or before the transition to a new activity (Bicard, 2000; Hester et al., 2009). Another limitation of this study was using TIMECS measurement for observational coding because teachers may use other practices when interacting with young children who display challenging behaviors not captured by the TIMECS measure. For example, some literature suggests teachers engage in ignoring when addressing undesired behaviors (Feil et al., 2014). Such practices were not part of the TIMECS

measure because they can be difficult to observe (McLeod et al., 2017). Therefore, those practices were not captured in this measurement.

Related to the observational nature of the measure, it is important to consider who is conducting the observations. Although the coders for this study include a diverse representation of gender and racial/ethnic background, the coder's identities can still influence the results. Coders were not teachers; therefore, that could influence their understanding of the practices teachers deliver. Additionally, the race and ethnicity of the coders were not an exact match for the race and ethnicity of teachers. None of the coders or trainers were Black, while 35.2% of the teachers were Black. Since about 93% of the teachers were women, stereotypical bias about Black women may play a role in coding of the practice elements. These biases include perceiving Black women as more aggressive and hostile (Ashley, 2013). These biases may play a role in coding of certain practices that focuses on positive interaction with children such as Supportive Listening or building Teacher-Child Relationship. Although, two coders were sent to an observation, and regular reliability assessments were performed at weekly meetings to discuss coder drifts, since racist stereotypes were not explicitly discussed in the training nor in later meetings, there is still room for bias to occur in the observations.

Another limitation of this study is that it did not examine if those trained on BEST in CLASS prior to this data collection differed on any of the study aims. The sample bias estimate showed the teachers trained on BEST in CLASS significantly scored higher on five out of the six items in the BEST in CLASS intervention. However, further analysis was not conduct to examine if those teachers differed on any other items or if taking those teachers out the sample would significantly change the characterization of the BAU classrooms. Those teaches were kept as part of the whole analysis because teachers are in early childhood settings are often trained on other curriculums, I only knew which teachers were previously trained on BEST in CLASS because the principal investigators of BEST in CLASS and TIMECS are the same.

Teachers are trained on other curriculums as well as indicated in the methods, therefore teachers receiving training on various curriculums is part of the BAU for early childhood settings.

Lastly, another limitation of this study is that BAU classrooms can look very different. The practice elements' usage may depend on teacher training and other social and emotional learning resources provided in schools. Therefore, the findings from this study may not generalize to all early childhood classrooms.

Future Directions

Due to the limitations mentioned above, it is important to replicate this study in other early childhood classrooms to examine if similar findings are present across early childhood classroom. Future studies should also examine how practices used in BAU impacts child outcomes. This is particularly important because practices in interventions are used to decrease challenging behaviors. Therefore, knowing if what teachers currently use impacts child behavior is important. Future studies should also examine which teacher and child characteristics impact the usage of the practice elements. For example, do children with higher challenging behaviors receive less supportive relationships from their teachers?

Conclusion

The findings from this study showcase that teachers were using practice elements and combining them as they are found in the literature for children who display challenging behaviors. These findings have implications for quality improvement which focuses on using the current skills and building off those rather than re-teaching all practice elements, some of which teachers may already be using at sufficient dosage. Overall, these findings support that it is useful to examine the baseline practices delivered by teachers before implementing new ones.

References

Abrahamse, M. E., Junger, M., van Wouwe, M. A., Boer, F., & Lindauer, R. J. (2016). Treating child disruptive behavior in high-risk families: A comparative effectiveness trial from a community-based implementation. *Journal of child and family studies*, *25*, 1605-1622. https://doi.org/10.1007/s10826-015-0322-4

Ashley, W. (2014). The angry black woman: The impact of pejorative stereotypes on psychotherapy with black women. *Social Work in Public Health*, *29*(1), 27-34. https://doi.org/10.1080/19371918.2011.619449

Barbarin, O. A. (2007). Mental health screening of preschool children: Validity and reliability of ABLE. *American Journal of Orthopsychiatry*, *77*(3), 402-418. https://doi.org/10.1037/0002-9432.77.3.402

Barra, S., Turner, D., Müller, M., Hertz, P. G., Retz-Junginger, P., Tüscher, O., Huss, M., & Retz, W. (2021). ADHD symptom profiles, intermittent explosive disorder, adverse childhood experiences, and internalizing/externalizing problems in young offenders. *European Archives of Psychiatry and Clinical Neuroscience, 272*(2), 257-269. https://doi.org/10.1007/s00406-020-01181-4

Bicard, D. F. (2000). Using classroom rules to construct behavior. *Middle School Journal*, *31*(5), 37-45. https://doi.org/10.1080/00940771.2000.11494651

Branson, D., & Demchak, M. (2011). Toddler teachers' use of teaching pyramid practices. *Topics in Early Childhood Special Education*, *30*(4), 196-208. https://doi.org/10.1177/0271121410369184

Brookman-Frazee, L. I., Taylor, R., & Garland, A. F. (2010). Characterizing community-based mental health services for children with autism spectrum disorders and disruptive behavior problems. *Journal of Autism and Developmental Disorders*, *40*(10), 1188-1201. https://doi.org/10.1007/s10803-010-0976-0

Campbell, C. (2011). *Adapting an evidence-based intervention to improve social and behavioral competence in Head Start children: Evaluating the effectiveness of teacher-child interaction* training (Doctoral Dissertation). The University of Nebraska-Lincoln. Retrieved from http://digitalcommons.unl.edu/psychdiss/37.

Castillo, W., & Gillborn, D. (2022). *How to "QuantCrit:" Practices and questions for education data researchers and users*. 22-546. EdWorkingPaper. https://doi.org/10.26300/v5kh-dd65

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290. https://doi.org/10.1037/1040-3590.6.4.284

Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology*, *77*(3), 566-579. https://doi.org/10.1037/a0014565

Collier-Meek, M. A., Johnson, A. H., Sanetti, L. H., & Minami, T. (2019). Identifying critical components of classroom management implementation. *School Psychology Review*, *48*(4), 348-361. https://doi.org/10.17105/SPR-2018-0026.V48-4

Connelly, L. M. (2016). Trustworthiness in qualitative research. *Medsurg Nursing, 25*(6), 435–437.

Cook, C. R., Williams, K. R., Guerra, N. G., Kim, T. E., & Sadek, S. (2010). Predictors of bullying and victimization in childhood and adolescence: a meta-analytic investigation. *School Psychology Quarterly*, *25*(2), 65-83. https://doi.org/10.1037/a0020149

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

Conroy, M. A., Sutherland, K. S., Algina, J. J., Wilson, R. E., Martinez, J. R., & Whalon, K. J. (2015). Measuring teacher implementation of the BEST in CLASS intervention program and corollary child outcomes. *Journal of Emotional and Behavioral Disorders*, *23*(3), 144-155. https://doi.org/10.1177/1063426614532949

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(7), 1-9. https://doi.org/10.7275/jyj1-4868

Daunic, A., Corbett, N., Smith, S., Barnes, T., Santiago-Poventud, L., Chalfant, P., Pitts, D., & Gleaton, J. (2013). Brief report: Integrating social-emotional learning with literacy instruction: An intervention for children at risk for emotional and behavioral disorders. *Behavioral Disorders*, *39*(1), 43-51. https://doi.org/10.1177/019874291303900106

de Lijster, J. M., van den Dries, M. A., van der Ende, J., Utens, E. M., Jaddoe, V. W., Dieleman, G. C., Hillegers, M.H., Tiemeier, H., & Legerstee, J. S. (2019). Developmental trajectories of anxiety and depression symptoms from early to middle childhood: A population-based cohort study in the Netherlands. *Journal of Abnormal Child Psychology*, *47*(11), 1785-1798. https://doi.org/10.1007/s10802-019-00550-5

De Pry, R. L., & Sugai, G. (2002). The effect of active supervision and pre-correction on minor behavioral incidents in a sixth grade general education classroom. *Journal of Behavioral Education, 11*(4), 255-267. https://doi.org/10.1023/A:1021162906622

Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *The Journal of Primary Prevention*, *28*(2), 67-91. https://doi.org/10.1007/s10935-007-0081-0

Driscoll, K. C., Wang, L., Mashburn, A. J., & Pianta, R. C. (2011). Fostering supportive teacher–child relationships: Intervention implementation in a state-funded preschool program.

Early Education & Development, *22*(4), 593-619. https://doi.org/10.1080/10409289.2010.502015

Duncan, N. G., Dufrene, B. A., Sterling, H. E., & Tingstrom, D. H. (2013). Promoting teachers' generalization of intervention use through goal setting and performance feedback. *Journal of Behavioral Education*, *22*(4), 325-347. https://doi.org/10.1007/s10864-013-9173-5

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272-299. https://doi.org/10.1037/1082-989X.4.3.272

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

Fleiss, J. (1981). Balanced incomplete block designs for interrater reliability studies. *Applied Psychological Measurement, 5*(1), 105-112. https://doi.org/10.1177/014662168100500115

Fleming, M., Fitton, C. A., Steiner, M. F., McLay, J. S., Clark, D., King, A., Mackay, D. F., & Pell, J. P. (2017). Educational and health outcomes of children treated for attention-deficit/hyperactivity disorder. *JAMA Pediatrics*, *171*(7), e170691-e170691. https://doi.org/10.1001/jamapediatrics.2017.0691

Floress, M. T., Berlinghof, J. R., Rader, R. A., & Riedesel, E. K. (2017). Preschool teachers' use of praise in general, at-risk, and special education classrooms. *Psychology in the Schools*, *54*(5), 519-531. https://doi.org/10.1002/pits.22014

Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers' views of implementation barriers and facilitators. *School Mental Health*, *1*(1), 26-36. https://doi.org/10.1007/s12310-008-9002-5

Fullerton, E. K., Conroy, M. A., & Correa, V. I. (2009). Early childhood teachers' use of specific praise statements with young children at risk for behavioral disorders. *Behavioral Disorders*, *34*(3), 118-135. https://doi.org/10.1177/019874290903400302

Gage, N. A., Wilson, J., & MacSuga-Gage, A. S. (2014). Writing performance of students with emotional and/or behavioral disabilities. *Behavioral Disorders*, *40*(1), 3-14. https://doi.org/10.17988/0198-7429-40.1.3

Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., & Ganger, W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services*, *61*(8), 788-795. https://doi.org/10.1176/ps.2010.61.8.788

Garland, A. F., Hawley, K. M., Brookman-Frazee, L., & Hurlburt, M. S. (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(5), 505-514. https://doi.org/10.1097/CHI.0b013e31816765c2

Garland, A. F., Plemmons, D., & Koontz, L. (2006). Research–practice partnership in mental health: Lessons from participants. *Administration and Policy in Mental Health and Mental Health Services Research*, *33*(5), 517-528. https://doi.org/10.1007/s10488-006-0062-2

George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge.

Gresham, F. M., Dart, E. H., & Collins, T. A. (2017). Generalizability of multiple measures of treatment integrity: Comparisons among direct observation, permanent products, and self-report. *School Psychology Review*, *46*(1), 108-121. https://doi.org/10.17105/SPR46-1.108-121

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). Multivariate data analysis New York. *NY: Macmillan*.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Hamm, J. V., Granger, K. L., & Van Acker, R. A. (2020). Classroom peer ecologies and cultures, and students with EBD: social dynamics as setting events for intervention. In T. W. Farmer, M. A. Conroy, E. M. Z. Farmer, & K. S. Sutherland (Eds.), *Handbook of Research on Emotional and Behavioral Disorders* (pp. 125-139). Routledge. https://doi.org/10.4324/9780429453106-9

Hamre, B., & Pianta, R. C. (2006). Student-teacher relationships. In G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, Prevention, and Intervention* (pp. 59–72). Bethesda, MD.

Hemmeter, M. L., Fox, L., Jack, S., & Broyles, L. (2007). A program-wide model of positive behavior support in early childhood settings. *Journal of Early Intervention*, *29*(4), 337-355. https://doi.org/10.1177/105381510702900405

Herschell, A. D., Taber-Thomas, S. M., Kolko, D. J., McLeod, B. D., & Jackson, C. B. (2019). Treatment-as-usual for child physical abuse in community mental health centers: therapist characteristics, client profiles, and therapy processes. *Journal of Emotional and Behavioral Disorders*, *28*(4), 223-234. https://doi.org/10.1177/1063426619866188

Hester, P. P., Hendrickson, J. M., & Gable, R. A. (2009). Forty years later—The value of praise, ignoring, and rules for preschoolers at risk for behavior disorders. *Education and Treatment of Children*, *32*(4), 513-535. https://doi.org/10.1353/etc.0.0067

Hogue, A., Bobek, M., Dauber, S., Henderson, C. E., McLeod, B. D., & Southam-Gerow, M. A. (2019). Core elements of family therapy for adolescent behavior problems: Empirical distillation of three manualized treatments. *Journal of Clinical Child & Adolescent Psychology*, *48*(1), 29-41. https://doi.org/10.1080/15374416.2018.1555762

Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy, 33*(2), 332-345. https://doi.org/10.1037/0033-3204.33.2.332

The IRIS Center. (2014). *Evidence-based practices (part 1): Identifying and selecting a practice or program.* Retrieved from https://iris.peabody.vanderbilt.edu/module/ebp_01/

Larsson, H. (2020). Developmental course of ADHD across adolescence into young adulthood. In S. P. Becker (Eds), *ADHD in adolescents: Development, assessment, and treatment* (pp. 21-35). New York: Guilford Publications.

Kaiser, H. F. (1991). Coefficient alpha for a principal component and the Kaiser-Guttman rule. *Psychological Reports*, *68*(3), 855-858. https://doi.org/10.2466/pr0.1991.68.3.855

Kauermann, G., & Carroll, R. J. (2001). A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association*, *96*(456), 1387-1396. https://doi.org/10.1198/016214501753382309

Kena, G., Hussar W., McFarland J., de Brey C., Musu-Gillette, L., Wang, X., Zhang, J., Rathbun, A., Wilkinson Flicker, S., Diliberti M., Barmer, A., Bullock Mann, F., and Dunlop Velez, E. (2016). The condition of education 2016. *U.S. Department of Education, National Center for Education Statistics. Washington, DC*. Retrieved from http://nces.ed.gov/pubsearch

Kolko, D. J. (1996). Individual cognitive behavioral treatment and family therapy for physically abused children and their offending parents: A comparison of clinical outcomes. *Child Maltreatment*, *1*(4), 322-342. https://doi.org/10.1177/1077559596001004004

Kolko, & Swenson, C. (2002). *Assessing and treating physically abused children and their families: A cognitive-behavioral approach*. SAGE Publications, Inc. https://doi.org/10.4135/9781452231761

Margolin, G., Oliver, P. H., Gordis, E. B., O'hearn, H. G., Medina, A. M., Ghosh, C. M., & Morland, L. (1998). The nuts and bolts of behavioral observation of marital and family

interaction. *Clinical Child and Family Psychology Review*, *1*(4), 195-213. https://doi.org/10.1023/a:1022608117322

McIntosh, D. E., Rizza, M. G., & Bliss, L. (2000). Implementing empirically supported interventions: Teacher-child interaction therapy. *Psychology in the Schools*, *37*(5), 453-462. https://doi.org/10.1002/1520-6807(200009)37:5<453::AID-PITS5>3.0.CO;2-2

McLeod, B. D., Smith, M. M., Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2015). Measuring treatment differentiation for implementation research: The Therapy Process Observational Coding System for Child Psychotherapy Revised Strategies Scale. *Psychological Assessment*, *27*(1), 314-325. https://doi.org/10.1037/pas0000037

McLeod, B. D., Sutherland, K. S., Broda, M., Granger, K. L., Cecilione, J., Cook, C. R., Conroy, M. A., Snyder, P. A., & Southam-Gerow, M. A. (2021). Examining the correspondence between teacher-and observer-report treatment integrity measures. *School Mental Health 14*, 20-34. https://doi.org/10.1007/s12310-021-09437-7

McLeod, B. D., Sutherland, K. S., Martinez, R. G., Conroy, M. A., Snyder, P. A., & Southam-Gerow, M. A. (2017). Identifying common practice elements to improve social, emotional, and behavioral outcomes of young children in early childhood classrooms. *Prevention Science*, *18*(2), 204-213. https://doi.org/10.1007/s11121-016-0703-y

McLeod, B. D., & Weisz, J. R. (2010). The Therapy Process Observational Coding System for Child Psychotherapy Strategies Scale. *Journal of Clinical Child & Adolescent Psychology*, *39*(3), 436-443. https://doi.org/10.1080/15374411003691750

Michigan's MTSS Technical Assistance Center. Tier 2 Behavior | Michigan's MTSS Technical Assistance Center. (n.d.). Retrieved October 30, 2021, from https://mimtsstac.org/practices/behavior/tier-2-behavior.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Mvududu, N. H., & Sink, C. A. (2013). Factor analysis in counseling research and practice. *Counseling Outcome Research and Evaluation*, *4*(2), 75-98. https://doi.org/10.1177/2150137813494766

Nelson, B. B., Chung, P. J., Forness, S. R., Pillado, O., Savage, S., DuPlessis, H. M., Hayslip, W., & Kataoka, S. H. (2013). Developmental and health services in head start preschools: A tiered approach to early intervention. *Academic Pediatrics*, *13*(2), 145-151. https://doi.org/10.1016/j.acap.2012.12.007

Newborg, J. (2005). *Battelle Developmental Inventory, 2nd Edition, Examiner's Manual*. Riverside Publishing.

O'Handley, R. D., Olmi, D. J., & Kennedy, A. (2019). Time-out procedures in school settings. In K. C. Radley & E. H. Dart (Eds.), Handbook of behavioral interventions in schools (pp. 482–500). Oxford University Press. https://doi.org/10.1093/med-psych/9780190843229.003.0025

Ocasio, K., Van Alst, D., Koivunen, J., Huang, C., & Allegra, C. (2015). Promoting preschool mental health: Results of a 3 year primary prevention strategy. *Journal of Child and Family Studies, 24,* 1800–1808. https://doi.org/10.1007/s10826-014-9983-7

Patton, M. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research, 34*(5 Pt 2), 1189-1208.

Quin, D. (2017). Longitudinal and contextual associations between teacher–student relationships and student engagement: A systematic review. *Review of Educational Research*, *87*(2), 345-387. https://doi.org/10.3102/0034654316669434

Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS: Identifying areas for enhancement. *Journal of Positive Behavior Interventions*, *15*(1), 39-50. https://doi.org/10.1177/1098300712459079

Reinke, W. M., Stormont, M., Webster-Stratton, C., Newcomer, L. L., & Herman, K. C. (2012). The Incredible Years Teacher Classroom Management program: using coaching to support generalization to real-world classroom settings. *Psychology in the Schools*, *49*(5), 416-428. http://dx.doi.org/10.1002/pits.21608

Riley, W. J., Moran, J. W., Corso, L. C., Beitsch, L. M., Bialek, R., & Cofsky, A. (2010). Defining quality improvement in public health. *Journal of Public Health Management and Practice*, *16*(1), 5-7. https://doi.org/10.1097/PHH.0b013e3181bedb49

Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. (2017). Affective teacher–student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review*, *46*(3), 239-261. https://doi.org/10.17105/SPR-2017-0035.V46-3

Roorda, D. L., Koomen, H. M., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review Of Educational Research*, *81*(4), 493-529. https://doi.org/10.3102/0034654311421793

Rones, M., & Hoagwood, K. (2000). School-based mental health services: A research review. *Clinical child and family psychology review*, *3*(4), 223-241. https://doi.org/10.1023/a:1026425104386

Sanetti, L. M. H., & Collier-Meek, M. A. (2019). *Supporting successful interventions in schools: Tools to plan, evaluate, and sustain effective implementation*. The Guilford Press.

Sec. 300.8 (C) (4). Individuals with Disabilities Education Act. (2017, May 2). Retrieved November 3, 2021, from https://sites.ed.gov/idea/regs/b/a/300.8/c/4.

Schweinhart, L. J., & Weikart, D. P. (1997). The high/scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12*(2), 117-143. https://doi.org/10.1016/S0885-2006(97)90009-0

Scott, T. M., Alter, P. J., & Hirn, R. G. (2011). An examination of typical classroom context and instruction for students with and without behavioral disorders. *Education and Treatment of Children*, *34*(4), 619-641. https://doi.org/10.1353/etc.2011.0039

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420-428. https://doi.org/10.1037//0033-2909.86.2.420

Sibley, M. H., Mitchell, J. T., & Becker, S. P. (2016). Method of adult diagnosis influences estimated persistence of childhood ADHD: A systematic review of longitudinal studies. *Lancet Psychiatry*, *3*(12), 1157-1165. https://doi.org/10.1016/S2215-0366(16)30190-0

Sibley, Swanson, J. M., Arnold, L. E., Hechtman, L. T., Owens, E. B., Stehli, A., Abikoff, H., Hinshaw, S. P., Molina, B. S. ., Mitchell, J. T., Jensen, P. S., Howard, A. L., Lakes, K. D., Pelham, W. E., Vitiello, B., Severe, J. B., Eugene Arnold, L., Hoagwood, K., Richters, J., … Stern, K. (2017). Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity. *Journal of Child Psychology and Psychiatry*, *58*(6), 655–662. https://doi.org/10.1111/jcpp.12620

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*(3), 351-380. https://doi.org/10.1353/etc.0.0007

Smith, S. C., Lewis, T. J., & Stormont, M. (2011). The effectiveness of two universal behavioral supports for children with externalizing behavior in Head Start classrooms. *Journal of Positive Behavior Interventions*, *13*(3), 133-143. https://doi.org/10.1177/1098300710379053

Smith, M. M., McLeod, B. D., Southam-Gerow, M. A., Jensen-Doss, A., Kendall, P. C., & Weisz, J. R. (2017). Does the delivery of CBT for youth anxiety differ across research and practice settings? *Behavior Therapy*, *48*(4), 501-516. https://doi.org/10.1016/j.beth.2016.07.004

Snyder, P. A., Hemmeter, M. L., Fox, L., Bishop, C. C., & Miller, M. D. (2013). Developing and gathering psychometric evidence for a fidelity instrument: The Teaching Pyramid Observation Tool–Pilot version. *Journal of Early Intervention*, *35*(2), 150-172. https://doi.org/10.1177/1053815113516794

Sprague, J. R., & Golly, A. (2005). *Best behavior: Building positive behavior support in schools*. Sopris West Educational Services.

Steinberg, M. P., & Lacoe, J. (2017). What do we know about school discipline reform? Assessing the alternatives to suspensions and expulsions. *Education Next*, *17*(1), 44-53.

Strain, P. S., & Timm, M. A. (2001). Remediation and prevention of aggression: An evaluation of the Regional Intervention Program over a quarter century. *Behavioral Disorders*, *26*(4), 297-313. https://doi.org/10.1177/019874290102600409

Stormont, M., Lewis, T. J., Beckner, R., & Johnson, N. W. (2008). *Implementing systems of positive behavior support systems in early childhood and elementary settings*. Corwin Press.

Sugai, Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., Scott, T., Liaupsin, C., Sailor, W., Turnbull, A. P., Turnbull, H. R., Wickham, D., Wilcox, B., & Ruef, M. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of positive behavior interventions*, *2*(3), 131-143. https://doi.org/10.1177/109830070000200302

Sultan, R. S., Liu, S. M., Hacker, K. A., & Olfson, M. (2021). Adolescents with attention-deficit/hyperactivity disorder: Adverse behaviors and comorbidity. *Journal of Adolescent Health*, *68*(2), 284-291. https://doi.org/10.1016/j.jadohealth.2020.09.036

Sutherland, K. S., McLeod, B. D., Granger, K., Nemer, S. L., Kunemund, R. L., Conroy, M. A., Johnson, A., & Miles, C. (2019). Adapting an evidence-based early childhood tier 2 program for early elementary school. *The Elementary School Journal*, *119*(4), 542-561. https://doi.org/10.1086/703103

Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD: A review. *Remedial and Special Education*, *22*(2), 113-121. https://doi.org/10.1177/074193250102200205

Sutherland, K. S., Wehby, J. H., & Yoder, P. J. (2002). Examination of the relationship between teacher praise and opportunities for students with EBD to respond to academic requests. *Journal of Emotional and Behavioral Disorders*, *10*(1), 5-13. https://doi.org/10.1177/106342660201000102

Tabachnick, B. G., & Fidell, L. S. (2001). Principal components and factor analysis. *Using multivariate statistics*, *4*(1), 582-633.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics: International edition*. Pearson.

Theodore, L. A. (2020). Linking assessment results to evidence-based interventions: best practices for preschool children. In V. C. Alfonso, B. A. Bracken & R. J. Nagle (Eds.), *Psychoeducational Assessment of Preschool Childre*n (5th ed., pp. 421-439). Routledge. https://doi.org/10.4324/9780429054099

Thompson, B & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, *56*(2), 197–208. https://doi.org/10.1177/0013164496056002001

Upshur, C. C., Heyman, M., & Wenz-Gross, M. (2017). Efficacy trial of the second step early learning (SSEL) curriculum: Preliminary outcomes. *Journal of Applied Developmental Psychology, 50,* 15–25. https://doi.org/10.1016/j.appdev.2017.03.004

Vo, A. K., Sutherland, K. S., & Conroy, M. A. (2012). BEST in CLASS: A classroom-based model for ameliorating problem behavior in early childhood settings. *Psychology in the Schools*, *49*(5), 402-415.

Walker, H. Severson, H., & Feil, E. (1995). *Early Screening Project: A proven child find process.* Sopris West Inc.

Walker, H., Horner, R., Sugai, G., Bullis, M., Sprague, J., Bricker, D., & Kaufman, M. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional and Behavioral Disorders*, *4*(4), 194-209. https://doi.org/10.1177/106342669600400401

Williford, A. P., LoCasale-Crouch, J., Whittaker, J. V., DeCoster, J., Hartz, K. A., Carter, L. M., Wolcott, C. S., & Hatfield, B. E. (2017). Changing teacher–child dyadic interactions to improve preschool children's externalizing behaviors. *Child Development*, *88*(5), 1544-1553. https://doi.org/10.1111/cdev.12703

Williford, A. P., & Vitiello, V. E. (2020). Who's in charge? Child behavior predicts teacher subsequent classroom management practice for preschoolers reported to display disruptive behavior. *School Psychology*, *35*(5), 299-310. https://doi.org/10.1037/spq0000402

Yoder, M. L., & Williford, A. P. (2019). Teacher perception of preschool disruptive behavior: Prevalence and contributing factors. *Early Education and Development*, *30*(7), 835-853. https://doi.org/10.1080/10409289.2019.1594531

Zuberi, T., & Bonilla-Silva, E. (Eds.). (2008). *White logic, white methods: Racism and methodology*. Rowman & Littlefield Publishers.

Table 1

Teacher Sample Demographic and Education Data

Variable	M (*SD*) or % (*n* = 54)
Female	92.6
Male	7.4
Age	
18 – 25	1.9
26 – 35	40.7
36 – 45	16.7
46 – 55	22.2
> 55	14.8
Prefer not to answer	3.7
Race	
Black	35.20
White	61.10
Native American/American Indian	–
Asian/Pacific Islander	–
Multiracial	3.70
Highest level education	
High School Diploma	3.7
Associates Degree	3.7
Bachelor's Degree	38.9
Master's Degree	48.1
Doctoral Degree	1.9
Other	3.7
Years teaching	12.99 (9.50)
Years teaching early childhood	7.69 (7.98)

Table 2

Child Demographic Data

Variable	*M (SD)* or % (n=91)
Age	4.31 (.67)
Female	54.9
Male	25.3
Not reported	19.8
Race	
Black	45.1
White	8.8
Asian/Pacific Islander	1.1
Native American/American Indian	1.1
Other/Multiracial	5.5
Unknown/Not Reported	38.4
Ethnicity	
Latinx	7.7

Table 3

t-Test Results Comparing Business-as-Usual and BEST in CLASS Groups on BEST in CLASS Items

	BAU		BiC				
	M	*SD*	*M*	*SD*	*df*	*t*	*p*
Rules	1.32	.67	1.64	.92	648	-4.95	<.001**
Precorrection	1.48	.68	1.77	.79	648	-4.99	<.001**
Opportunity to Respond	4.59	.99	4.82	1.08	648	-2.66	.008**
Praise	2.73	1.02	3.15	1.15	648	-4.74	<.001**
Error Correction	2.44	1.05	2.64	1.10	648	-2.24	.025*
Instructive Feedback	1.52	.70	1.50	.72	648	.10	.917

Note. BAU = Business-as-Usual; BiC = BEST in CLASS.
*$p < .05$. **$p < .01$.

Table 4

TIMECS item name, definition, descriptive data and interrater reliability

Item name	Definition	Skewness	Kurtosis	M (*SD*)	Range	ICC(2,2)
Social Skills	Teacher provides instruction on strategies that can facilitate positive social interactions with their peers (e.g., friendship skills, social etiquette, sharing, taking turns) or adults.	1.33	1.54	1.66(.85)	5.00	.88
Emotion Regulation	Teacher provides instruction focused on helping to identify, label, or regulate his/her emotions.	2.38	6.05	1.33(.68)	4.00	.89
Problem Solving	Teacher provides instruction designed to generate solutions to social, behavioral, or pre-academic problems.	4.35	25.89	1.13(.42)	4.50	.83
Promoting Behavioral Competence	Instruction that focuses on promoting positive behavior (e.g., engagement) during instructional activities.	-.19	-.33	3.95(1.05)	5.50	.80
Teacher-Child Relationship	Teacher behavior that conveys warmth, closeness, and interest when listening to and interacting.	.60	-.28	2.49(1.25)	6.00	.87
Rules	Teacher uses guidelines to teach the rules and behavioral expectations of the classroom.	1.90	3.07	1.43(.77)	4.00	.90
Narrating	Teacher provides a verbal description of behavior.	1.48	1.63	1.46(.66)	3.00	.80
Supportive Listening	Teacher actively demonstrates understanding of the topic.	1.47	2.14	1.73(.94)	5.00	.83
Choices	Teacher provides an opportunity to select between two or more options related to instructional activities.	3.32	12.17	1.11(.32)	2.50	.68
Monitoring	Teacher actively monitors.	-.24	-.32	4.80(.92)	5.00	.69
Modeling	Teacher demonstrates, or has a peer demonstrate, a specific behavioral or pre-academic skill to promote learning.	.74	-.33	2.14(1.09)	4.50	.81
Rehearsal	Teacher encourages practice of a behavioral skill (e.g., during interactions with peers).	3.44	13.66	1.20(.50)	3.50	.80
Precorrection	Teacher uses prompts prior to the occurrence of a behavior to remind of appropriate behavior and correct responding (e.g., reminding of rules, expectations).	1.17	.65	1.58(.73)	3.50	.77

Opportunities to Respond	Teacher uses questions or prompts (i.e., gestural, verbal, visual, physical) that seek an active, observable, and specific response.	-.49	-.02	4.67(1.03)	6.00	.72
Visual Cueing	Teacher uses visual cues to prompt for appropriate behavioral responses or consequences.	1.22	.94	1.77(.93)	4.00	.79
Premack Principle	Teacher uses a more reinforcing behavior (e.g., playtime) to reinforce less probable behaviors (e.g., lesson time).	3.02	9.38	1.15(.41)	2.50	.80
Tangible Reward	Teacher gives a tangible/representative reward in response to an appropriate social, emotional, or behavioral response.	3.87	15.03	1.11(.39)	2.50	.89
Time-out	Teacher removes a child from a preferred activity for a specified period of time following a problem behavior.	7.66	66.55	1.04(.25)	3.00	.95
Praise	Teacher provides positive verbal statements of approval in response to an appropriate social, emotional, behavioral, or pre-academic response.	.06	-.64	2.87(1.08)	4.50	.82
Error Correction	Teacher provides corrective feedback following an incorrect response or undesirable behavior.	.30	-.48	2.51(1.07)	5.50	.79
Instructive Feedback	Teacher provides extra instructional information while responding to correct response or appropriate behavior.	1.34	1.73	1.51(.71)	3.00	.76

Note. TIMECS = Treatment Integrity Measure for Early Childhood Settings; ICC= interclass correlation coefficient.

Table 5

Results from Hypothesis 1 and Hypothesis 2

Item name	% across obs	M (*SD*)	M (*SD*) when observed
Social Skills	50.0%	1.66(.85)	2.32(.75)
Emotion Regulation	25.2%	1.33(.68)	2.31(.74)
Problem Solving	12.6%	1.13(.42)	2.06(.65)
Promoting Behavioral Competence	99.1%	3.95(1.05)	3.98(1.02)
Teacher-Child Relationship	76.8%	2.49(1.25)	2.94(1.08)
Rules	30.9%	1.43(.77)	2.39(.77)
Narrating	41.7%	1.46(.66)	2.09(.59)
Supportive Listening	51.8%	1.73(.94)	2.40(.86)
Choices	12.2%	1.11(.32)	1.89(.39)
Monitoring	95.5%	4.80(.92)	4.80(.92)
Modeling	69.1%	2.14(1.09)	2.65(.94)
Rehearsal	19.2%	1.20(.50)	2.01(.70)
Precorrection	49.5%	1.58(.73)	2.16(.62)
Opportunities to Respond	99.7%	4.67(1.03)	4.67(1.01)
Visual Cueing	54.8%	1.77(.93)	2.40(.83)
Premack Principle	15.5%	1.15(.41)	1.99(.51)
Tangible Reward	8.8%	1.11(.39)	2.25(.57)
Time-out	2.9%	1.04(.25)	2.34(.67)
Praise	90.8%	2.87(1.08)	3.06(.95)
Error Correction	82.6%	2.51(1.07)	2.83(.90)
Instructive Feedback	45.1%	1.51(.71)	2.14(.63)

Note. Obs= Observations.

Table 6

Correlations between TIMECS Items that displayed normal distribution

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Social Skills	1													
2. Emotion Regulation	.25**	1												
3. Promoting Behavioral Competence	.27**	.25**	1											
4. Teacher-Child Relationship	.26**	.13**	.11**	1										
5. Rules	.22**	.21**	.38**	-.03	1									
6. Narrating	.15**	.15**	.41**	.09	.26**	1								
7. Supportive Listening	.21**	.08*	.05	.66**	-.12**	-.01	1							
8. Monitoring	.27**	.12**	.55**	.35**	.21**	.27**	.23**	1						
9. Modeling	.01*	15**	.47**	.09*	.25**	.29**	-.01	.34**	1					
10. Precorrection	.14**	.10*	.39**	.01	.35**	.29**	-.01	.20**	.31**	1				
11. Opportunities to Respond	.20**	.17*	.69**	.21**	.31**	.36**	.01**	.68**	.49**	.34**	1			
12. Visual Cueing	-.02	.16**	.44**	-.07	.17**	.19**	-.07	.26*	.46**	.28**	.38**	1		
13. Praise	.15**	.15**	.57**	.15**	.31**	.45**	.02	.47**	.43**	.37**	.61**	.29**	1	
14. Error Correction	.12**	.10**	.46**	.06	.11**	.25**	.03	.24**	.40**	.27**	.43**	.33**	.33**	1
15. Instructive Feedback	.05	.06	.33**	.01	.12**	.17*	.07	.39**	.28**	.18**	.47**	.30**	.33**	.29*

Note. TIMECS = Treatment Integrity Measure for Early Childhood Settings; * = $p < .01$, ** = $p < .05$.

Table 7

Results of Exploratory Factor Analysis for TIMECS Items

	Factor 1: Classroom Management	Factor 2: Supportive Relationship
Factor Eigenvalue	4.658	1.827
Factor Variance	70.48%	29.52%
Factor 1: Behavioral Management		
Promoting Behavioral Competence	**.813**	.099
Rules	**.404**	-.080
Narrating	**.477**	.052
Monitoring	**.687**	.384
Modeling	**.598**	.043
Precorrection	**.459**	-.033
Opportunities to Respond	**.859**	.216
Visual Cueing	**.495**	-.101
Praise	**.711**	.117
Error Correction	**.513**	.035
Instructive Feedback	**.478**	.044
Factor 2: Supportive Relationship		
Teacher-Child Relationship	.188	**.870**
Supportive Listening	.076	**.752**

Note. TIMECS = Treatment Integrity Measure for Early Childhood Settings.
Bold font indicates the primary factor on which the given item loads.