Incorporating evidence-based behavioral teacher training into Head Start mental health consultation: Description and initial outcomes of a large-scale program

Grace S. Hubel1 | John L. Cooley2 | Angela D. Moreland3

1Department of Psychology, The College of Charleston, Charleston, South Carolina
2Department of Psychiatry, Developmental Psychobiology Research Group, University of Colorado Anschutz Medical Campus, Aurora, Colorado
3National Crime Victims Research and Treatment Center, Medical University of South Carolina, Charleston, South Carolina

Correspondence
Grace S. Hubel, The College of Charleston, Charleston 29424, SC.
Email: hubelgs@cofc.edu

Abstract
Development of a multicomponent mental health consultation program for a countywide Head Start program is described. The consultation program incorporates strategies from the evidence-based practice, Teacher–Child Interaction Training (TCIT). Through large group professional development trainings before the school year, comprehensive feedback reports based on observations, and in-class coaching via modeling and performance feedback on teachers’ use of TCIT skills, the consultation program served 55 Head Start classrooms with 789 children. Outcome data are presented on teachers’ improvement in TCIT skill use and the relative effectiveness of two coaching methods (i.e., modeling and performance feedback) used during consultation. Results from multilevel modeling analyses indicated that teachers improved in the use of multiple observed TCIT skills between the initial and midyear assessment (i.e., increased frequency of labeled praises, reflections, behavioral descriptions, and commands that were complied with; decreased frequency of commands that children did not have an opportunity to comply with). Additionally, teachers who received coaching in the form of performance feedback, in comparison to modeling, exhibited greater gains in the frequency of labeled
praises and commands that resulted in compliance. Discussion focuses on implementation of evidence-based practice in large-scale preventative early interventions, study limitations, and directions for future research.

**KEYWORDS**
mental health consultation, preschool, teacher–child interaction training

## 1 | INTRODUCTION

The preschool years are a critical period in child socioemotional development. During this time, children begin developing capacities critical for success in school and other important spheres of life such as social problem-solving, emotion regulation, frustration tolerance, and the ability to inhibit behaviors and follow rules and directions (Berlin & Bohlin, 2002; Denham, 2006; Dowsett & Livesy, 2000; Jones, Rothbart, & Posner, 2003; Keenan & Wakschlag, 2002; Penela, Walker, Degnan, Fox, & Henderson, 2015). While these skills develop, it is typical for children to exhibit oppositional behaviors and emotional dysregulation. However, a small subset will display a pervasive and persistent pattern of socioemotional difficulties that is outside the developmental norm and impairs functioning (Bornstein, Hahn, & Haynes, 2010; Campbell, Shaw, & Gillion, 2000; Egger & Angold, 2006). Preschool-aged children can display clear and treatable characteristics of attention-deficit/hyperactivity disorder, oppositional defiant and conduct disorders, anxiety disorders, and depressive disorders (Dougherty et al., 2015; Egger & Angold, 2006; Whalen, Sylvester, & Luby, 2017). Overall, rates of psychological disorders for preschoolers range from 10% to 27%, which is similar to rates observed for older children and adults (Bekman et al., 2017; Bufferd, Dougherty, Carlson, & Klein, 2011; Egger & Angold, 2006). Among children diagnosed with early-onset mental health disorders, there is an increased risk for problems in adulthood, such as engagement in health-risk behavior, educational and occupational underachievement, criminality, and perpetration of violence against others (Chorozoglou et al., 2015; Moffitt, Caspi, Harrington, & Milne, 2002; Whalen, Belden, Tillman, Barch, & Luby, 2016).

Early identification and intervention are important for reducing the prevalence of mental health diagnoses and for fostering the self-regulation and adaptability skills necessary for lifelong well-being. Left unaddressed, difficulties identified in the preschool years often escalate and are exacerbated in later childhood by issues such as peer rejection, the formation of deviant peer groups, school-related stressors (e.g., academic failure and critical teachers), and high levels of family stress (Dodge & Pettit, 2003; Dodge et al., 2003, Milkie & Warner, 2011; Repetti, Taylor, & Seeman, 2002; Stearns, Dodge, & Nicholson, 2008). Moreover, while it is not always possible to prevent adversity from occurring during a child’s lifespan, promoting mental health in early childhood provides the tools necessary to cope with life stressors through self-regulation of attention, emotion, and behavior (Shonkoff, 2012).

Unfortunately, among the general population of preschoolers in the United States, about 75% of children in need of mental health services do not receive assistance (Dougherty et al., 2015). The disparity between the need for, and delivery of, preventative mental health is especially prevalent among children living in families that experience socioeconomic disadvantage (Kataoka, Zhang, & Wells, 2002). This disparity likely plays a role in the link demonstrated in multiple large-scale longitudinal studies between childhood poverty and poor adult health and well-being (Barch et al., 2016; Brooks-Gunn & Duncan, 1997; Evans & Cassells, 2014; Hair, Hanson, Wolfe, & Pollak, 2015).

School-based interventions represent one strategy for delivering preventive mental health services to preschool children. Providing training to preschool personnel regarding early signs and symptoms of mental health consultation, preschool, teacher–child interaction training
health issues can aid in early detection and prevention efforts. Additionally, school-based services help eliminate barriers that have been identified to delivery of services for children, including travel time and transportation problems, parents having to take time away from work, and lack of knowledge about where to seek help (Owens et al., 2002; Storch & Crisp, 2004). Given that some children and families may perceive stigma when seeking services in traditional mental health clinics or be mistrustful of mental health care providers (Owens et al., 2002), providing services in a trusted, familiar setting such as the school may make them more acceptable (Owens & Murphy, 2004). Finally, providing services in the school also provides an opportune environment for children to practice skills in real-life settings with peers and teachers, thereby increasing generalization to other settings and environments (Masia-Warner et al., 2005).

Head Start is a national, federally funded preschool program that serves socioeconomically disadvantaged families with 3- to 5-year-old children and is thus a program capable of reaching a large group of young children likely to experience disparity in access to preventative mental health services. The Head Start approach targets multiple domains of child competence as well as the broader contexts in which development occurs, through promotion of family well-being and community involvement (Yoshikawa & Zigler, 2000). As part of this multi-component approach, Head Start performance standards require that programs provide families with access to mental health consultants. Cohen and Kaufmann (2000) provide a conceptual definition of the consultation services that are typically provided to Head Start, using the term Early Childhood Mental Health Consultation (ECMHC). ECMHC is defined as a “problem-solving and capacity-building intervention that involves a collaborative relationship between a consultant with mental health expertise and one or more individuals in a preschool or childcare setting” (Cohen & Kaufmann, 2000, p. 4). There are two broad models of ECMHC used in Head Start, termed child/family focused and program focused (Perry, Dunne, McFadden, & Campbell, 2008). In child/family focused ECMHC, the mental health consultant works with teachers and staff to address the specific needs of an individual child or family. In program-focused consultation, the consultant works to improve the overall classroom environment and build teacher/staff capacity for addressing children's problematic behaviors. This manuscript presents a description of, and initial outcomes for, a program-focused ECMHC model tailored to the needs of a local Head Start program.

The consultation program examined draws on available research evidence supporting Teacher–Child Interaction Training (TCIT) for use in preschool classrooms. TCIT is an adaptation of Parent–Child Interaction Therapy (PCIT), a manualized, empirically supported intervention in which caregivers are taught play-therapy and behavioral skills for improving relationships with children and managing child behavior (Eyberg, Nelson, & Boggs, 2008). Demonstrated efficacy and effectiveness of PCIT in improving parenting practices and reducing problematic behaviors in preschoolers has led to the development of several TCIT models (Filcheck, McNeil, Greco, & Bernard, 2004; Garbacz, Zychinski, Feuer, Carter, & Budd, 2014; Lyon et al., 2009; McIntosh & Rizza, 2000; Tiano & McNeil, 2006). Previous research supporting TCIT has examined training one teacher and the behavior of one child or children in one classroom at a time (Filcheck et al., 2004; McIntosh & Rizza, 2000). Other investigations of TCIT have examined training small groups of teachers to work with children in a few classrooms (Garbacz et al., 2014; Lyon et al., 2009; Tiano & McNeil, 2006). Findings from these previous evaluations of TCIT have included increases in positive teacher–child interactions, increased use of structured discipline routines, decreased child disruptive behavior, and increased child prosocial behavior (Filcheck et al., 2004; Garbacz et al., 2014; Lyon et al., 2009; McIntosh & Rizza, 2000; Tiano & McNeil, 2006).

Positive previous findings support TCIT as an intervention that could help improve classroom environments and build teacher’s skills for managing problematic behaviors. However, as is often a barrier in implementing evidence-based interventions in school settings, there were key differences between the conditions under which TCIT has previously been evaluated and the setting of the countywide Head Start program for which the current model was developed (Durlak & DuPre, 2008). In particular, the TCIT program needed to be modified so that many teachers could be trained simultaneously with a focus on improving their capacity for promoting the social-emotional well-being of all of the children in a countywide program. The current program description outlines the
needs of the Head Start program and how TCIT was adapted into a consultation model to meet those needs (McIntosh, Filter, Bennett, Ryan, & Sugai, 2010). An analysis of initial outcomes with two main aims is also presented. First, the analysis of initial outcomes explores the impact of the implementation of the program on teachers’ skills for improving relationships with children and managing classroom behavior. Further, as outlined in the program description below, the developed consultation program recognized the need to provide ongoing coaching for teachers on use of TCIT skills after introducing these skills via trainings at the beginning of the school year (Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005; Kraft, Blazar, & Hogan, 2018). The second aim of the analyses presented in this manuscript was to examine the relative effectiveness of two coaching methods used by consultants (i.e., modeling and performance feedback) to improve teacher use of TCIT skills.

2 | PROGRAM DESCRIPTION

The consultation program described in this manuscript was developed after Head Start leadership of a countywide program contacted a collaborative group of clinical researchers with expertise in preventative mental health at an academic medical center in the same county in the Southeastern United States. The Head Start leadership team recognized that mental health consultation was a required and needed service within their program, and that the Head Start program was currently unable to provide the service. At the time of development, the Head Start program had the need to provide mental health consultation to a large number of families (approximately 800) and a limited budget, as they did not have enough funding to hire a full-time psychologist. In response to these service and budgetary needs, a program was developed in which clinical psychology predoctoral interns completing training at the academic medical center would provide consultation services, with doctoral-level clinical psychologists providing training and weekly one-on-one or small group supervision.

In collaboration with Head Start leadership, the consultation program was developed to include four services that are required by Head Start performance standards related to mental health consultation and that were not currently being delivered (U.S. Department of Health and Human Services, 2016). It was identified that the following services should be built into the program: a training for teachers during annual professional development days by experts on techniques for managing child social, emotional, and behavioral concerns; annual screenings of all enrolled Head Start children to identify social and emotional concerns; an onsite, in-classroom consultant to assist teachers in promoting an environment of positive child social and emotional functioning; and a process through which families in need of individual mental health treatment could be referred to available resources in the community.

Design of the mental health consultation program focused on two important goals. One was addressing the specific deficits described above and identified by the program in its ability to provide the components of mental health consultation required by performance standards. Outside of the guidance that these performance standards provide, individual Head Start programs have wide discretion in their ability to tailor programing to meet the needs of families (Green, Everhart, Gettman, Gordon, & Friesen, 2003). Given this flexibility, design of the program also focused on incorporating evidence-based strategies from TCIT into the consultation model. The four components of the developed consultation program are as follows:

2.1 | Professional development trainings for teachers

Professional development trainings are delivered before the beginning of the school year to teachers in groups of approximately 40–55 during 1.5-hr time blocks. Procedures for delivering these trainings are described below. The content of the trainings was developed using the PCIT manual (McNeil & Hembree-Kigin, 2010) as well as insights from previous adaptations of PCIT for use in preschool settings (Filcheck et al., 2004; Garbacz et al., 2014; Lyon
et al., 2009; McIntosh & Rizza, 2000; Tiano & McNeil, 2006). The trainings are intended for delivery in conjunction with additional program components described below. Conceptually, and based on research examining the uptake of evidence-based practices in school settings, it was believed that the professional development trainings would be beneficial for increasing teacher’s knowledge and awareness of skills for building positive relationships with children and the behavioral principles that underlie TCIT (Coles, Owens, Serrano, Slavec, & Evans, 2015, Joyce & Showers, 2002). It was assumed that the next steps, observations with feedback reports and in-classroom coaching, would be necessary for teachers’ translation of skills from the training to the classroom and for promoting flexible use of the skills in a variety of situations and with a variety of children (Fixsen et al., 2005; Kraft et al., 2018; Schultz, Arora, & Mautone, 2015).

The training provides rationale for and opportunity for teachers to practice a set of skills for interactions with children that include: praise, reflection, imitation, description, and displaying enjoyment (i.e., PRIDE skills). The PRIDE skills are intended to help teachers build relationships with children and scaffold the development of prosocial student behaviors through social attention and modeling (Fernandez et al., 2015). The training explains that praise is most effective when it is labeled, that is when positive feedback about children’s behavior describes the specific behavior that the teacher would like the student to continue in the future (e.g., Wonderful job sharing your blocks with your friend!). Reflection is introduced as a skill through which teachers demonstrate attention to children’s verbalizations by repeating the child’s statements. For instance, if a child says “I’m drawing a house,” the teacher might say, “You are drawing a house!” In addition to providing social attention to children’s appropriate verbalizations, reflection also can allow the teacher to model appropriate grammar and verbal expression for the child. For instance, if a child is drawing and says, “I draw house red!” the teacher might say, “You drew a red house!”

Imitation is explained as a way of engaging in play with children at the child’s level, in a manner that demonstrates approval of their behavior but does not challenge them to change their behavior to be more consistent with expectations of adults. For instance, if a child is building a block tower that is three blocks high, the teacher is encouraged to also build a block tower that is about three blocks high (instead of showing the child how to build a tall tower). Teachers are also taught to describe children’s behavior during play interactions, for example, by saying, “You’re stacking the red block on top of the green block!” It is explained that describing children’s behavior reinforces the behavior through social attention. Rationale for use of description is also that it helps children to attend to the elements of their behavior and environment that the teacher describes, such as the color of the toy with which the child is playing (Gershenson, Lyon, & Budd, 2010). Finally, teachers are encouraged to demonstrate their enjoyment of their interactions with children through verbal and nonverbal behavior to build warmth and promote children’s engagement in teacher–child interactions.

The professional development training also provides strategies for reducing the number of unnecessary criticisms, questions, and commands that teachers use when interacting with children. For instance, teachers are encouraged to try using reflections and behavior descriptions to demonstrate interest in children and their behavior instead of asking conversational questions. Strategies for shaping children’s behavior (e.g., praising behaviors that are opposite of behaviors that they would like to see decrease and ignoring disruptive behaviors) are covered. It is emphasized that it is adaptive to increase the proportion of time and energy spent providing positive social attention to appropriate child behavior relative to attending to minor misbehavior. By doing this, it is likely that teachers will more often positively reinforce appropriate behavior and spend less time engaged in negative interactions with children (Gershenson et al., 2010).

The professional development training also covers strategies for preventing problem behaviors. Teachers are taught methods for effectively using commands (McNeil & Hembree-Kigin, 2010). Teachers are encouraged to use direct commands, which are clearly stated and specifically direct the behavior of the child (e.g., “Please tie your shoe”). The training also encourages teachers to decrease the use of indirect commands, which are intended to be directives but that could be interpreted by the child as optional or a question (e.g., “Will you please sit down?”). It is emphasized that only one command should be delivered at a time and that children should be allowed adequate time to comply with commands. Rationale is provided that preschool-aged children are often developmentally
unable to follow a multistep sequence of commands, and teachers may misinterpret children's behavior as non-compliance if they are not given adequate time to comply (Radley & Dart, 2016).

The structure of the professional development training involves presentation of an explanation of each skill, followed by examples of how to use the skill in the classroom. After each skill is presented, the consultants model the skill, with one consultant acting as a child and one consultant acting as a teacher. Then, teachers are given the opportunity to role-play using each of the skills themselves. Role-play exercises are conducted using volunteer teachers. In the role plays, teachers are provided with an example of a behavior that a typical preschool-aged child might display in the classroom (e.g., sharing a toy with a friend) and the teacher is asked to show how they might use the TCIT skill of focus (e.g., labeled praise) in response to that behavior (e.g., by saying, “Wonderful job sharing your blocks with your friend!”).

2.2 Initial and midyear classroom screenings

An initial screening of mental health and social and emotional concerns for all Head Start children in the first 90 days of programing is required by Head Start performance standards. In building a program with dual goals of meeting performance standards and incorporating TCIT strategies, we developed a process through which mental health consultants visit each classroom and conduct a screening of teacher’s use of TCIT skills for managing classroom behavior, children's response to classroom management techniques (e.g., whether children comply with commands), and children who appear to be displaying unusual levels of behavioral and/or emotional difficulties. The mental health consultants use an observational coding system (described below) to complete the observations. The observational coding system captures teacher TCIT skill use (these skills are described in Table 1) as well as children's responses to classroom management techniques. During the observation period, the observers also use clinical judgment to make note of children displaying unusual levels of behavioral and/or emotional difficulties. Examples of unusual levels of behavioral and/or emotional difficulties that would be flagged include aggression towards peers, withdrawn behavior that leads to few interactions with peers, or indications of unassessed/untreated developmental delays (e.g., lack of verbal language development).

Consultants, in collaboration with the doctoral-level supervisors, use initial screenings to identify classrooms and children as falling into the following three categories: (1) Classrooms not needing further consultation services at this time (this would be a classroom with few identified social and emotional concerns during the observation, with teachers who effectively use techniques for managing these concerns), (2) Classrooms containing children that display social and emotional concerns that might be better managed if teachers are provided with in-classroom coaching of TCIT skill use by consultants, and (3) Children in need of referrals for individual mental health assessment or treatment in the community. Classrooms containing children in need of referrals would typically be identified as in need of in-classroom coaching of TCIT skill by consultants (i.e., falling in both categories 2 and 3). This screening process allows for the consultants to develop triage categories to inform decision-making on how to divide the limited number of consultation hours across the large group of children served. Midway through the school year, a second assessment is conducted in all classrooms to assess teacher's change in use of TCIT skills, children's change in ability to respond to classroom management techniques, and to make note of whether initially identified children may have benefited from referrals to community providers or if additional children are in need of referrals.

Following each observation, mental health consultants complete comprehensive feedback reports on teachers' current use of TCIT skills, with recommendations for improved skill use. These feedback reports are delivered verbally to the teacher by the consultant and teachers are provided with a written summary of the feedback and recommendations. By using the observational coding system of TCIT skills in initial and midyear classroom screenings, the consultants gain insight into the teacher's current strengths and weaknesses in terms of use of TCIT skills, as well as any particularly challenging child behaviors that may be occurring in the classroom. Thus, the focus
of the consultant’s time with the teacher can be individualized. For instance, if a teacher frequently uses reflections during interactions with children, but does not praise appropriate child behavior often, the teacher would be provided with this feedback and the consultant would plan to focus on increasing the teacher’s use of praise during the upcoming consultations that use in‐classroom coaching. Using observation followed by coaching is consistent with research by Barnett et al. (2017) examining the promotion of changes in parent’s behavior via PCIT. This study indicated that responsive techniques for promoting parent skill use, which provide feedback following the use of a PCIT skill, were more effective than directive techniques, which tell a parent to use a skill in the future (Barnett et al., 2017). Further, providing teachers with both verbal and written feedback on their use of TCIT skills following observations is consistent with previous adaptations of PCIT to TCIT, which have paired verbal and written

<table>
<thead>
<tr>
<th>Teacher skill observed</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeled praise</td>
<td>Providing praise to a child, a group of children, or the entire classroom that includes a specific statement describing the behavior that the teacher would like to see increase in the future</td>
<td>“Great job sharing your toy with your friend!”</td>
</tr>
<tr>
<td>Unlabeled praise</td>
<td>Providing a praise to a child, a group of children, or the entire classroom that is unspecific regarding the behavior that the teacher would like to see increase in the future</td>
<td>“Nice work!”</td>
</tr>
</tbody>
</table>
| Reflections            | Responding to a verbalization by a child by repeating or paraphrasing what the child has said | Child: “I’m drawing a circle”  
Teacher “You’re drawing a circle” |
| Behavioral descriptions| The teacher observes the child’s behavior and describes either positive or neutral behavior of the child | Child: Paints with a paint brush  
Teacher: You’re painting on the paper |
| Negative talk          | Critical or sarcastic comments made by the teacher in reference to the attributes or behavior of a child, group of children, or the entire classroom of children | Teacher to the class: You all are being bad today!  
Teacher to a student: That was a stupid thing to do |
| Commands followed by:  | Commands are statements made by the teacher that direct the behavior of a child, group of children, or the classroom | Listen!  
Please find your seat  
Put this toy in the box |
| Child compliance       | When the child or the majority of the group of children to whom a command was delivered comply or attempt to comply with the command within 5 s | Teacher: Please find your seat  
Group of children: within 5 seconds, the majority of the group are walking towards the area where seats are located |
| Child noncompliance    | When the child or the majority of the group of children to whom a command was delivered do not comply or attempt to comply with the command within 5 s | No opportunity for child to comply  
When a child or the group of children to whom a command was delivered are not given adequate time and/or opportunity to comply |
| No opportunity for child to comply | When a child or the group of children to whom a command was delivered are not given adequate time and/or opportunity to comply | Teacher: Please this toy in the box  
Teacher: within 5 seconds, the teacher puts the toy in the box |
feedback to help teachers who are attempting to remember and use feedback while also accommodating many other demands in the preschool classroom (Lyon et al., 2009).

2.3 | In-classroom coaching by consultants

After identifying classrooms in need of in-class consultation during initial screenings, consultants visit classrooms during the school day. During these visits, the consultants further assess the antecedents and consequences of concerning issues with emotion and behavior displayed by children via direct observation of children and brief teacher interviews. Further, they provide teachers with coaching in addressing these concerns using TCIT skills. The frequency of in-class consultations varies based on classroom needs, with consultants typically visiting classrooms 1–2 times per month.

Consistent with the TCIT model, consultants use two strategies during in-class coaching. The first, modeling, is defined as the consultant using TCIT skills with children in the classroom after prompting the teacher to observe the consultant. For example, if a teacher is observed using an indirect command, such as “Don’t you think it’s time to clean up?” with a child, the consultant might use modeling to promote the use of direct commands by the teacher. To do this, the consultant might say, “It looks like you would like [child’s name] to clean up the toys, but you are having a hard time getting him/her to listen. A direct command might be a good strategy to use here. How about I try using a direct command and you can watch and see what happens.” Then, the consultant would use a direct command directed towards the child, such as “Please put the toys back in this box.” The second strategy, performance feedback, is defined as a consultation that includes each of the following in order: prompting the teacher to use TCIT skill(s) in the classroom, observing the teacher implementing one or more TCIT skills, and providing feedback at the moment to the teacher.

In previous adaptations of TCIT, modeling has often been used primarily during dyadic, out-of-classroom trainings to introduce TCIT skills (Garbacz et al., 2014; Gershenson et al., 2010; Tiano & McNeil, 2006). Conceptual support for the use of modeling is based in Bandura and Walters (1977), which explains that an individual’s learning of a new skill can be promoted when the individual observes another person perform the skill. The professional development trainings delivered as part of the current intervention used modeling of skills by consultants. However, these trainings were delivered to larger numbers of teachers over shorter periods of time than in previous adaptations. Therefore, it was thought that additional coaching via modeling might be helpful to improve the teachers’ skill use in the classroom.

Performance feedback is a core component of the originally developed PCIT model and has been used in classrooms in TCIT adaptations (see Fernandez, Gold, Hirsch & Miller, 2015 for a review). Performance feedback allows the consultant to respond to the teacher’s use of skills in the classroom environment and for immediate provision of reinforcement or corrective feedback (Barnett, Niec, & Acevedo-Polakovitch, 2014). Using performance feedback has been identified as a key aspect of success and generalization in interventions that train teachers in the use of behavioral classroom management skills (Snyder et al., 2012). Some previous TCIT models have delivered performance feedback to teachers outside the classroom with one or more children who also leave the classroom and participate in the session (e.g., Campbell, 2011). However, the cost of providing a substitute teacher so this type of coaching could occur, along with the logistics of this type of delivery, was not feasible for the current program. Thus, the current program tested an adaptation by which performance feedback is provided in vivo, during teachers’ regular classroom instruction.

When consultants visited classrooms and provided in-class coaching, they used clinical judgment to determine whether to use modeling or performance feedback. No specific criteria were provided to the consultants for determining when to use one strategy versus the other. In the training that the consultants received before delivering in-class coaching, the conceptual and research support for both modeling and performance feedback was described. Based on anecdotal reports from the consultants, an example of why modeling may have been used in
favor of performance feedback is if a child displayed a behavior that presented a good opportunity for use of a TCIT skill (e.g., if a child was observed sharing a toy, the consultant might choose to model use of labeled praise). In this example, modeling would have allowed the consultant to immediately reinforce the child's prosocial behavior as well as provide the teacher with an example of how to use a specific TCIT skill. On the other hand, an example of a time when performance feedback may have been used in favor of modeling is during activities that allowed for child-directed unstructured play and thus gave the teacher more opportunity to engage in one-on-one interactions with individual children than during more structured and teacher-directed activities.

2.4 | Referrals to community providers

In developing the consultation program, it was recognized that some children served by Head Start would likely display mental health needs that warranted individual treatment (Kazdin & Blase, 2011). After identifying children potentially in need of referrals for individual assessment or treatment in the community, the consultants work with teachers, other Head Start staff, and parents/caregivers as facilitators of access to these services. Children identified as in need of individual mental health treatment are referred to local providers with expertise in the delivery of services to preschool-age children, including providers who would accept Medicaid and who provide services on a sliding scale for families without insurance.

3 | CURRENT STUDY

The current study examines preliminary outcomes of the professional development trainings, classroom screening/observations with feedback, and in-classroom coaching components of the consultation program. Specifically, change in teacher use of each of the skills outlined in Table 1 across the period of time between an initial and a midyear classroom observation are examined. The first research question guiding the study was: Will teacher’s use of TCIT skills improve between the initial and the midyear classroom observation? Based on positive evidence from other TCIT programs, we hypothesized that teacher use of TCIT skills would improve across time. The role of type of coaching (modeling or performance feedback) in influencing change in teacher skills was also examined. The second research question guiding the study was: will one type of coaching (modeling or performance feedback) be more effective in influencing change in teacher skill? This aim of the study was considered exploratory, as previous research on TCIT has not examined differences between methods of promoting teacher skill use through consultation.

4 | METHOD

4.1 | Participants

The consultation program was delivered to 55 Head Start classrooms nested within 13 Head Start centers, which served 789 children. Each classroom was staffed by a lead and an assistant teacher. Lead teachers in Head Start are required to have a baccalaureate degree or higher in early childhood education or a related field. Assistant teachers are required to have an associates degree in child development or a related field. Of the 55 classrooms, 32 were classrooms operated independently by the Office of Head Start and contained only children eligible to participate in Head Start programming. The remaining 23 classrooms were Child Development centers operated jointly by the Office of Head Start and the local school district and contained both children eligible to participate in Head Start and children who were not eligible (e.g., children who enrolled in preschool services through the public school
system who lived in families with incomes above Head Start eligibility criteria). The Head Start center size ranged between 1 and 16 classrooms ($M = 4.23$, $SD = 3.79$, $Mdn = 3$). At the time of the initial observation, classroom sizes ranged between 9 and 20 students ($M = 15.04$, $SD = 2.72$, $Mdn = 15$). Classroom sizes at the midyear observation ranged between 11 and 20 students ($M = 15.44$, $SD = 2.10$, $Mdn = 15$).

4.2 | Procedure

Clinical psychology predoctoral interns delivered the consultation services. Two interns at a time served as consultants on 6-month rotations. The predoctoral interns in the consultant role had extensive training in behavior management techniques and received additional training in the consultation intervention. Before beginning their work with the teachers, the consultants participated in a 1.5-day (12 hr) training on the specifics of the consultation program and its incorporation of strategies from TCIT. Two doctoral-level psychologists oversaw the mental health consultation program, provided training for the consultants before the beginning of their work with teachers, and provided weekly supervision.

The first team of two consultants (consultants A and B) delivered the large group teacher professional development training in TCIT skill use before initial observations of teachers’ skills. During the professional development training, the consultants informed teachers that they would be visiting their classrooms to conduct observations of children’s behavior and assess needs for ongoing coaching in TCIT skills. They also provided a brief overview of structure and nature of the coaching services that would be provided in the classroom during the upcoming school year. Two large group professional development trainings were delivered to different groups of teachers; however, the content, examples, and role-play exercises used in the trainings were identical. These same two consultants conducted initial observations and provided coaching on TCIT skills. To ensure fidelity to the TCIT model used to structure observations and consultation, weekly supervision was held with the consultants and the supervisors for approximately 90 min per week to discuss the coding system, use of coaching to increase teacher TCIT skill use, and any implementation challenges.

Midway through the school year, two new consultants completed midyear assessments (consultants C and D). Consultants C and D provided later intervention that was not a focus of this investigation. During the time period of focus in the current study, consultant A was responsible for service delivery to 25 classrooms and consultant B was responsible for 30 classrooms. Before the midpoint assessment, consultant C assumed responsibility for the classrooms of consultant A and D assumed responsibility for the classrooms of consultant B.

Initial observations were conducted for 30-min time periods. Midyear observations were conducted for 15-min time periods. Observations used the measure described below and focused on the behavior of both the lead and assistant teachers, as well as children’s responses to commands delivered by teachers. The initial observations were completed within 90 days of the first day of the Head Start school year. In individual classrooms, the midyear observation occurred between 4.75 and 8 months after the initial observation ($M = 5.63$, $SD = 0.78$, $Mdn = 5.25$). The consultation program was a clinical services program and the evaluation relied on data gathered that was essential to service delivery. The clinical data collected were analyzed in a deidentified database, so the institutional review board of [masked for peer review] did not require informed consent from teachers who participated in the consultation program.

4.3 | Observational coding system

The observational coding system used for initial and midyear assessments was adapted from the Dyadic Parent Child Interaction Coding System—III (Didactic Parent Interaction Coding System—3rd Edition, DPICS-III; Eyberg, Nelson, Duke, & Boggs, 2005). This observational coding system was originally developed for use with parent-child dyads. Numerous studies have been conducted documenting the reliability of the DPICS, which are summarized in the DPICS-III manual (Eyberg et al., 2005). Validity of the DPICS, including the ability to discriminate between
interactions of treatment referred and nontreatment referred parent–child dyads, convergent validity, and sensitivity to treatment effects, has also been demonstrated (see Eyberg et al., 2005, for a review). Consultants were trained in the use of the coding system for clinical purposes before conducting observations. The training involved review of the DPICS coding manual and completing coding exercises. Further, consultants met with the two doctoral-level psychologists who oversaw the mental health consultation program on a weekly basis for continued training and review of observational coding issues to maximize reliable use of the coding system. The coding categories utilized in the assessments were labeled praise, unlabeled praise, reflections, behavioral descriptions, and negative talk. Also coded were commands: followed by child compliance, followed by child noncompliance, and followed by no opportunity for the child to comply. These coding categories are displayed and defined in Table 1. Frequencies of use of each skill for each coding category were recorded during the observational assessment. Due to the varying duration of the initial and midyear observations, dependent variables were created according to the percentage of each skill recorded out of the total number of skills coded during each observation.

4.4 Collection of clinically relevant data

When consultation services were delivered in classrooms, the consultant made clinical progress notes, including if the consultation was delivered via modeling or performance feedback on TCIT skills. In cases, when modeling and performance feedback interventions were delivered in the same classroom on the same day, it was recorded that both a modeling and a performance feedback intervention was delivered. The progress notes were used as a way for consultants to self-monitor their fidelity to the TCIT model and were discussed in weekly supervision. Data from these progress notes were later extracted to a data set that tracked program implementation.

4.5 Data analytic plan

Descriptive statistics were initially estimated using IBM SPSS Statistics (Version 24; IBM Corp., 2016). Independent samples t tests were also conducted to examine potential differences between teachers’ behaviors in the Head Start and Child Development classrooms during the initial observation. Changes in teachers’ behaviors from the initial to the midyear observation were subsequently evaluated using a series of multilevel models within SAS (University Edition; SAS Institute Inc., 2014). Given the clustered longitudinal design, observations at level 1 were nested within classrooms at level 2 and schools at level 3. Model parameters were estimated using restricted maximum likelihood and the PROC MIXED procedure. Time was represented in the number of months between observations and centered such that the intercept corresponded to the initial observation. The significance of random effects was tested with likelihood ratio tests, whereas the significance of fixed effects was assessed by their individual Wald test p values using Satterthwaite denominator degrees of freedom. Effect sizes were measured by the pseudo-\(R^2\) values for the proportion reduction in the level-1 residual variance.

Multilevel models were built using a hierarchical approach. Empty means models were first estimated to partition the variance in teachers’ behaviors across levels. Next, a level-2 random intercept variance for classrooms was added to the model, which was followed by the addition of a level-3 random intercept variance for schools. A fixed linear effect of time was then added to the model to determine if there were significant changes in teachers’ behaviors from the initial to the midyear observation. Finally, for teacher behaviors that exhibited significant change over time, variables representing the classroom type (0 = Head Start, 1 = Child Development), mental health consultant who delivered the interventions (0 = consultant A, 1 = consultant B), number of modeling interventions, and number of performance feedback interventions were added to the model as fixed effects.
Overall, 789 children were observed within 55 classrooms and 13 schools. At the time of initial observation, 155 children were flagged with behavioral, emotional, and/or developmental concerns in the classrooms (\(M = 2.10, \ SD = 1.42, \ range = 0–5\)). At the midyear observations, 99 children remained flagged and 79 children were newly flagged, for a total of 178 flagged children (\(M = 2.96 \) children in each classroom, \(SD = 1.77, \ range = 0–7\)). Overall, 22% of children displayed concerns in more than one area, with reasons for being flagged including behavioral concerns (70%), development/language delay (25%), emotional concerns (16%), and social concerns (12%).

After the initial observation, 29 classrooms (52.7%) were identified as not needing further consultation services, whereas 26 classrooms (47.3%) were identified that contained children that displayed social and emotional concerns that would be better managed if teachers were provided with in-classroom coaching of TCIT skill use by consultants. During the time period between the initial and the midyear observations, the mental health consultants delivered a total of 74 interventions in which they modeled the use of behavior management techniques and 44 interventions in which they delivered performance feedback for teachers on the use of behavior management techniques across all of the classrooms. Within each classroom, the mental health consultants delivered between 0 and 11 modeling interventions (\(M = 1.35, \ SD = 2.41, \ Mdn = 0\)) and between 0 and 5 performance feedback interventions (\(M = 0.80, \ SD = 1.18, \ Mdn = 0\)).

Results from a series of independent samples \(t\) tests indicated that during the initial observation, teachers in the Head Start classrooms exhibited more reflections and behavioral descriptions, whereas the teachers in the Child Development classrooms exhibited more commands that resulted in compliance and fewer commands that resulted in noncompliance (see Table 2). Accordingly, classroom-type was controlled for in subsequent analyses.

### 5.2 | Multilevel models

#### 5.2.1 | Labeled praises

Compared to a single-level model, results indicated that the addition of a level-2 random intercept variance for classrooms did not significantly improve model fit, \(-2\Delta LL(1) = 0.872, p = .350\). However, the inclusion of a level-3

| TABLE 2 | Differences between teachers’ behaviors in head start and child development classrooms during the initial observation |
|-------------------------------|-----------------|----------------------------------|
| Labeled praises               | Head start (%)  | Child development (%)            |
|                               | 1.6             | 2.8                              |
| Unlabeled praises             | 14.8            | 15.2                             |
| Reflections                   | 8.5             | 4.7                              |
| Behavioral descriptions       | 3.6             | 1.1                              |
| Negative talk                 | 9.1             | 8.4                              |
| Commands—compliance           | 25.3            | 32.5                             |
| Commands—noncompliance        | 4.3             | 2.6                              |
| Commands—no opportunity       | 32.8            | 32.6                             |

Note: Bold estimates represent statistically significant differences in the percentage of each behavior exhibited by teachers during the initial observation. 

\(^a\)Represents independent samples \(t\) test estimates with equal variances not assumed.
random intercept variance for schools did significantly improve model fit, $-2\Delta\text{LL}(1) = 10.916, p = .001$. A fixed linear effect of time was then added to the model, which indicated that there were significant increases in the percentage of labeled praises used by teachers from the initial to the midyear observation, $b = 0.006, SE = 0.002, p = .007$ (see Figure 1); this effect accounted for 6.3% of the level-1 residual variance. When the other fixed effects were added to the model, results indicated that performance feedback interventions predicted greater

![Figure 1](image-url)
increases in the use of labeled praises over time, and teachers who interacted with consultant A exhibited greater gains than those who interacted with consultant B (see Table 3). Together, these effects accounted for an additional 11.9% of the level-1 residual variance.

### 5.2.2 | Unlabeled praises

Compared to a single-level model, the addition of a level-2, as well as a level-3 variance, resulted in a nonpositive definite G matrix, suggesting that there were no significant differences in the use of unlabeled praises across classrooms or schools. Accordingly, these effects were removed from the model. A fixed linear effect of time was then added to the model, which indicated that there were nonsignificant decreases in the percentage of unlabeled praises used by teachers from the initial to the midyear observation, $b = -0.005$, $SE = 0.003$, $p = .092$ (see Figure 1); this effect accounted for 1.7% of the residual variance.

### 5.2.3 | Reflections

Compared to a single-level model, the addition of a level-2, as well as a level-3 variance, resulted in a nonpositive definite G matrix, suggesting that there were not significant differences in the use of reflections across classrooms or schools. Accordingly, these effects were removed from the model. A fixed linear effect of time was then added to the model, which indicated that there were significant increases in the percentage of reflections used by teachers from the initial to the midyear observation, $b = 0.009$, $SE = 0.003$, $p = .012$ (see Figure 1); this effect accounted for 4.8% of the level-1 residual variance. When the other fixed effects were added to the model, results indicated that teachers who interacted with consultant B exhibited greater gains than those who interacted with consultant A (see Table 3). Together, these effects accounted for an additional 18.7% of the level-1 residual variance.

---

**TABLE 3** Predictors of linear change in labeled praises, reflections, and behavioral descriptions

<table>
<thead>
<tr>
<th></th>
<th>Labeled praises</th>
<th>Reflections</th>
<th>Behavioral descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b$</td>
<td>$SE$</td>
<td>$p$</td>
</tr>
<tr>
<td><strong>Level-2: Random intercept</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom type</td>
<td>-0.003</td>
<td>0.019</td>
<td>.882</td>
</tr>
<tr>
<td>Consultant</td>
<td>0.005</td>
<td>0.031</td>
<td>.870</td>
</tr>
<tr>
<td>Modeling</td>
<td>0.001</td>
<td>0.007</td>
<td>.930</td>
</tr>
<tr>
<td>Performance feedback</td>
<td>0.000</td>
<td>0.013</td>
<td>.976</td>
</tr>
<tr>
<td><strong>Level-1: Fixed linear slope</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.007</td>
<td>0.004</td>
<td>.038</td>
</tr>
<tr>
<td>Classroom type $\times$ time</td>
<td>0.005</td>
<td>0.004</td>
<td>.213</td>
</tr>
<tr>
<td>Consultant $\times$ time</td>
<td>-0.012</td>
<td>0.004</td>
<td>.004</td>
</tr>
<tr>
<td>Modeling $\times$ time</td>
<td>-0.001</td>
<td>0.001</td>
<td>.262</td>
</tr>
<tr>
<td>Performance feedback $\times$ time</td>
<td>0.007</td>
<td>0.003</td>
<td>.023</td>
</tr>
</tbody>
</table>

Note: Bold estimates represent statistically significant paths ($p < .05$); classroom type (0 = head start, 1 = child development); consultant (0 = consultant A, 1 = consultant B).
5.2.4 | Behavioral descriptions

Compared to a single-level model, results indicated that the addition of a level-2 random intercept variance for classrooms did not significantly improve model fit, $-2\Delta\text{LL}(1) = 0.695, p = .404$. Further, the inclusion of a level-3 random intercept variance for schools did not significantly improve model fit, $-2\Delta\text{LL}(1) = 1.078, p = .299$. Accordingly, these effects were removed from the model. A fixed linear effect of time was then added to the model, which indicated that there were significant increases in the percentage of behavioral descriptions used by teachers from the initial to the midyear observation, $b = 0.004, SE = 0.002, p = .023$ (see Figure 1); this effect accounted for a 3.9% of the level-1 residual variance. When the other fixed effects were added to the model, none significantly predicted change in behavioral descriptions over time (see Table 3). Together, these effects accounted for an additional 7.7% of the level-1 residual variance.

5.2.5 | Negative talk

Compared to a single-level model, results indicated that the addition of a level-2 random intercept variance for classrooms did not significantly improve model fit, $-2\Delta\text{LL}(1) = 2.813, p = .094$. However, the inclusion of a level-3 random intercept variance for schools did significantly improve model fit, $-2\Delta\text{LL}(1) = 20.435, p < .001$. A fixed linear effect of time was then added to the model, which indicated that there were nonsignificant increases in the percentage negative talk observed for teachers from the initial to the midyear observation, $b = 0.003, SE = 0.003, p = .266$ (see Figure 1); this effect accounted for 0.2% of the level-1 residual variance.

5.2.6 | Commands—Compliance

Compared to a single-level model, results indicated that the addition of a level-2 random intercept variance for classrooms did not significantly improve model fit, $-2\Delta\text{LL}(1) = 0.618, p = .432$. However, the inclusion of a level-3 random intercept variance for schools did significantly improve model fit, $-2\Delta\text{LL}(1) = 8.462, p = .015$. A fixed linear effect of time was then added to the model, which indicated that there were significant increases in the percentage

| TABLE 4 | Predictors of linear change in commands—compliance and commands—no opportunity |
|-----------------|------------------|-----------------|------------------|------------------|------------------|
|                 | Commands—compliance |                 | Commands—no opportunity |                 |                 |
|                 | $b$    | $SE$    | $p$    | $b$    | $SE$    | $p$    |
| Level-2: Random intercept |                 |                 |                 |                 |                 |
| Classroom type | 0.067  | 0.033  | .444  | -0.015 | 0.037  | .690  |
| Consultant     | 0.037  | 0.038  | .339  | 0.024  | 0.038  | .530  |
| Modeling       | 0.015  | 0.012  | .189  | -0.019 | 0.013  | .127  |
| Performance feedback | -0.043 | 0.022  | .060  | 0.030  | 0.025  | .237  |
| Level-1: Fixed linear slope |                 |                 |                 |                 |                 |
| Time           | 0.005  | 0.006  | .459  | -0.028 | 0.008  | .201  |
| Classroom type × time | 0.003  | 0.008  | .719  | 0.003  | 0.009  | .766  |
| Consultant × time | 0.015  | 0.008  | .063  | -0.014 | 0.009  | .128  |
| Modeling × time | -0.008 | 0.003  | .005  | 0.008  | 0.003  | .020  |
| Performance feedback × time | 0.014  | 0.005  | .012  | -0.011 | 0.007  | .103  |

Note: Bold estimates represent statistically significant paths ($p < .05$); classroom type (0 = head start, 1 = child development); consultant (0 = consultant A, 1 = consultant B).
of provision of commands that were complied with from the initial to the midyear observation, $b = 0.014$, $SE = 0.004$, $p = .001$ (see Figure 1); this effect accounted for 10.8% of the level-1 residual variance. When the other fixed effects were added to the model, results indicated that teachers in Child Development classrooms provided more commands that were complied with at the initial assessment as compared to teachers in the Head Start classrooms (see Table 4). Further, performance feedback interventions predicted greater increases in the use of commands that were complied with, whereas modeling interventions predicted fewer increases over time (see Table 4). Together, these effects accounted for an additional 22.6% of the level-1 residual variance.

### 5.2.7 | Commands—Noncompliance

Compared to a single-level model, results indicated that the addition of a level-2 random intercept variance for classrooms did not significantly improve model fit, $-2\Delta LL(1) = 0.022$, $p = .883$. Further, the inclusion of a level-3 random intercept variance for schools did not significantly improve model fit, $-2\Delta LL(1) = 4.376$, $p = .112$. Accordingly, these effects were removed from the model. A fixed linear effect of time was then added to the model, which indicated that there were nonsignificant increases in the percentage of commands delivered by teachers that were not complied with from the initial to the midyear observation, $b = 0.002$, $SE = 0.002$, $p = .253$ (see Figure 1); this effect accounted for 0.3% of the level-1 residual variance.

### 5.2.8 | Commands—No opportunity

Compared to a single-level model, the addition of a level-2, as well as a level-3 variance, resulted in a nonpositive definite G matrix, suggesting that there were not significant differences in the use of commands followed by no opportunity to comply across classrooms or schools. Accordingly, these effects were removed from the model. A fixed linear effect of time was then added to the model, which indicated that there were significant decreases in the percentage of commands delivered by teachers that students did not have an opportunity to comply with, $b = -0.032$, $SE = 0.004$, $p < .001$ (see Figure 1); this effect accounted for 32.3% of the level-1 residual variance. When the other fixed effects were added to the model, results indicated that modeling interventions predicted fewer decreases in the use of commands that students did not have an opportunity to comply with (see Table 4). Together, these effects accounted for an additional 4.7% of the level-1 residual variance.

### 6 | DISCUSSION

This study examines the initial outcomes of a large-scale mental health consultation program for a countywide Head Start program. The consultation program used large group teacher professional development trainings, observations with feedback reports for teachers, and in-classroom coaching. We examined clinically relevant data collected by the consultants to evaluate whether teacher use of key TCIT skills improved across the period of time between the initial and midyear classroom observation. Teachers demonstrated significant increases in the percentage of use of labeled praise, reflections, behavioral descriptions, and commands that were followed by child compliance. Teachers also demonstrated significant decreases in use of commands followed by no opportunity for compliance (e.g., delivering a command to complete a task and then completing the task for the child). These findings provide initial support for the consultation model, indicating that teachers were able to perform the skills and increase their utilization of them through this universal intervention. Further, changes in behavior were seen in a relatively short time frame (6-months), which provides additional support for the usefulness and effectiveness of the consultation model.
The positive results of this study in terms of improvement in teacher skill use are consistent with the previous literature on TCIT, which has consistently demonstrated that a variety of TCIT adaptations are capable of improving teacher TCIT skill use (Filcheck et al., 2004; Gershenson et al., 2010; Lyon et al., 2009; Tiano & McNeil, 2006). The relationship between the consultation program and teacher use of the skills examined is important, as the skills examined are those that the theoretical model underpinning the intervention proposes will improve relationships between adults and children and prevent problematic child behaviors (McNeil & Hembree-Kigin, 2010). This study expands on previous TCIT literature by demonstrating that training teachers in large groups, assessing and providing feedback on their skills use, and then delivering coaching on an individual, needs-based basis is a way of expanding the reach of TCIT. This is important because TCIT has previously been examined for training one teacher at a time or small groups (e.g., 4) of teachers to work with children in one or a few classrooms. Expanding the reach of empirically supported interventions such as TCIT may help reduce the disparity between the number of young children in need of preventative intervention and the number of trained and available providers to deliver these services (Dougherty et al., 2015; Kataoka et al., 2002).

The study also aimed to evaluate if one type of consultation (modeling or performance feedback) would be more effective in influencing change in teacher skill. Results indicated that performance feedback, in comparison to modeling, was related to greater increases in teachers’ use of labeled praise and commands followed by child compliance. Further, the use of modeling predicted fewer increases in the use of commands followed by compliance as well as fewer decreases in the use of commands with which preschool students did not have the opportunity to comply. Based on previous research in both the areas of dissemination of evidence-based treatments in psychology and teacher professional development, it was expected that additional in-classroom coaching after the professional development training in TCIT skills would be necessary to promote some teachers’ use of the skills in the classroom (Fixsen et al., 2005; Herschell, McNeil, & McNeil, 2004; Kraft et al., 2018; Kretlow & Bartholomew, 2010). However, while the importance of ongoing coaching is recognized, the method by which this coaching is delivered has not been examined in research on the implementation of TCIT in classrooms. In the current consultation model, teachers are exposed to additional training in how to correctly use TCIT skills when consultants use modeling, however they are not provided with specific feedback on their own use of the skills. The level to which individualized feedback on behavior is delivered has been shown to be a predictor of teachers’ ability to use other evidence-based behavior management strategies in preschool classrooms (Snyder, Hemmeter, & Fox, 2015). Consistent with this study, in vivo performance feedback appeared to be a more effective method than modeling for helping teachers improve in the use of labeled praise and the provision of effective commands.

In examining teachers’ improvement in specific TCIT skills, it appeared that teachers who worked with one consultant (“consultant A”) exhibited greater improvement in the use of labeled praise, while teachers who worked with another (“consultant B”) exhibited greater improvement in the use of reflections. On one hand, we are not aware of the literature on PCIT or TCIT that supports the relative superiority of any particular skill for improving child behavior or emotional functioning. On the other hand, these results indicate that some consultants may have more familiarity or skill in promoting teachers’ use of particular TCIT skills. In the current and similar consultation models, methods for monitoring fidelity by supervisors might improve consultants’ consistency in the use of components of evidence-based practice in the applied setting (Webster-Stratton, Reinke, Herman, & Newcomer, 2011).

### 6.1 Limitations

The results presented in this manuscript should be considered in light of several study limitations, in particular, the lack of a control comparison group. The research presented examines outcomes of a clinical service endeavor that was implemented with the intention of monitoring outcomes; thus, it was not pragmatically
possible to withhold delivery of the consultation service from any of the participants in the local Head Start program. Due to this limitation, teacher development over time (independent of the benefits of the consultation program) cannot be ruled out as an explanation for the improvements noted in teacher use of TCIT skills. Further, because this was an evaluation of a clinical service endeavor, the available data for analysis was limited to those that were clinically relevant and therefore collected by consultants to monitor progress and improve the service. When consultants met with teachers, they did not ask teachers to self-report their own demographics such as gender, ethnicity, education, or years of teaching experience, nor did they collect demographic information for classroom children. Thus, the analyses conducted did not test whether these demographics were related to the effectiveness of the intervention.

In this initial evaluation of the program, impacts in areas such as individual child behavior, teacher satisfaction, and teacher-consultant alliance were not measured. While this evaluation provides initial evidence that the consultation program led to improvements in teacher's use of skills for promoting preschool children's well-being and reducing problematic child behavior, it is not known to what extent the changes in teacher behavior impacted child behavior. Further, in the current evaluation, we were not able to assess whether factors such as individual teacher's satisfaction or their alliance with the consultants influenced the likelihood that they would successfully use TCIT skills. Additionally, the same consultants who delivered the large group teacher professional development trainings conducted initial observations and, thus, these data may have been impacted by the consultants' expectation that the training they delivered would lead to change in teacher behavior. This limitation was addressed to some degree when separate consultants, new to the rotation, implemented midyear assessments.

An additional limitation of this evaluation is that no inter-rater reliability data were obtained on the observations conducted by the consultants. As each consultant was individually responsible for observing and working with at least 25 classrooms, it was not possible in terms of time and resources for observations to be completed in the same classroom by two consultants simultaneously to assess inter-rater reliability. However, one way to improve this intervention in the future could be to use a supervision model whereby the supervisor conducts a percentage of observations simultaneously with the consultant to monitor inter-rater reliability. Finally, no formalized fidelity observations, such as a live or video-recorded observation of consultant's work with teachers rated by an independent observer, were incorporated. Consultants self-monitored their use of modeling and performance feedback on TCIT skills in progress notes and weekly supervision focused on fidelity to the model. Monitoring consultant's fidelity to a model is often impractical or cumbersome in implementation research in applied settings (Herschell, 2010). However, less time-intensive methods for monitoring fidelity, such as asking teachers to rate the consultants on the level to which they incorporate TCIT skills into consultation, might improve this intervention and its evaluation in the future.

6.2 | Future directions and summary

The consultation program presented continues to develop, with the involved Head Start leadership expressing a high level of satisfaction with the program and support of the work into the foreseeable future. Gathering outcome data on the effects of the intervention on child behavior across time will be essential to establish that the current design of the model is an effective method of impacting children's socioemotional development. Strategically planning to track child outcomes longitudinally would be particularly informative in determining if this type of low intensity, broad-reaching, preventative intervention is capable of reducing the prevalence of impairing mental health diagnoses in later childhood. The current support of the Head Start program, which closely collaborates with the local school system, is a critical first step in forming the partnerships that would be needed for longitudinal tracking of outcomes. Overall, evaluation with greater standardization of procedures and incorporation of a control group is necessary to further assess the effectiveness and refine the described model. Furthermore, although the support of Head Start leadership is promising, we have not assessed teacher satisfaction or alliance with
consultants, each of which are likely important to long-term sustainability (Handler et al., 2007; Johnson, Pas, & Bradshaw, 2016). Working with teachers to understand their perceptions of the program and overcome barriers to their active participation will be an important next step in this line of work.

While incorporating continued evaluation into delivery will be necessary to make the model as effective as possible, the needs of the Head Start program and the families served will likely make it impractical to design a traditional randomized controlled trial. However, in recent years, approaches have been described within dissemination and implementation science that provide methods for rigorously evaluating interventions as they are implemented in community settings. For example, the current consultation model uses a triage approach, wherein some teachers are provided with greater amounts of consultation because of demonstration of need for increased training during the initial observation. Given this approach, future evaluation could compare the effectiveness of two levels of intensity for delivery of consultation services, especially if a standardized method of delivering lower intensity services (e.g., a web-based or video follow-up to the initial professional development training) was also developed in the future (Glasgow et al., 2012). This type of evaluation method could also be used to test the relative added benefit of incorporating strategies to improve teacher satisfaction and alliance with consultants.

In summary, this manuscript describes and presents initial outcomes of a consultation program designed to improve Head Start teacher’s ability to promote healthy socioemotional development in the classroom. A strength of the presented model is its ability to disseminate evidence-based strategies to a large Head Start program within the constraints of the program’s needs and budget. Further, the initial analysis of outcomes, which examined the relative utility of performance feedback compared with modeling, provides insight into the need for and type of coaching necessary to improve uptake of skills by teachers following professional development training in evidence-based practice. A great deal of disparity exists between the number of preschool children who will exhibit difficulties with socioemotional development and the number who are likely to receive help from a trained professional (Kazdin & Blase, 2011). Given this disparity, the type of service delivery model presented is needed to expand the reach of evidence-based practice to populations such as the economically disadvantaged families served by Head Start.

ACKNOWLEDGEMENT

John Cooley, NIMH T32 Postdoctoral Fellow, Developmental Psychobiology Research Group, University of Colorado Anschutz Medical Campus, was supported in part by a training fellowship from NIMH (T32 MH015442).

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Grace S. Hubel http://orcid.org/0000-0002-3084-8905

REFERENCES


IBM Corp (2016). IBM SPSS Statistics for Windows (Version 24.0) [Computer software]. IBM Corp.


