EFFECTIVENESS OF TEACHER-CHILD INTERACTION TRAINING (TCIT): A MULTIPLE PROBE DESIGN ACROSS CLASSROOMS IN A DAY-TREATMENT PRESCHOOL

A DISSERTATION
SUBMITTED TO THE FACULTY
OF
RUTGERS,
THE STATE UNIVERSITY OF NEW JERSEY

BY
RYAN JOHN MADIGAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY OCTOBER, 2011

APPROVED:

Robert H. LaRue, Ph.D., BCBA-D

Steven M.S. Kurtz, Ph.D., ABPP

DEAN:

Stanley Messer, Ph.D.
ABSTRACT

The current study assessed the effectiveness of Teacher-Child Interaction Training (TCIT), an adaptation of Eyeberg’s Parent-Child Interaction Therapy (PCIT), on teacher and child behaviors in a day-treatment preschool setting. The sample included 5 day-treatment classrooms in an urban, socioeconomically disadvantaged, and culturally diverse setting. The study utilized a concurrent multiple probe design across classroom settings (3 training groups consisting of 5 classrooms) to evaluate the effects of didactic and in-vivo coaching on teacher and child behaviors in the training and classroom settings. Results indicated that all teachers’ use of positive behaviors increased and negative behaviors decreased during pull-out sessions; all 5 teachers attained CDI and TDI mastery criteria. Results also indicated some evidence of spontaneous generalization of teachers’ use of Labeled Praises to the classroom setting, while other teacher behaviors did not generalize. Results on child behavior were variable and failed to demonstrate consistent improvements in the classroom setting; this finding is understood given the lack of generalization of teachers’ behaviors to the classroom. These findings provide initial support for the use of TCIT to improve teachers’ behavior management skills, as well as support for the feasibility of implementing TCIT with fidelity to the PCIT manual. Additionally, the study offers insight into the possible need for additional adaptations to train teachers in how and when to implement the TCIT skills under high stress in-vivo classroom conditions.
ACKNOWLEDGMENTS

In reflecting on the course of my graduate studies and dissertation, I would like to thank my committee chair, Dr. Robert LaRue. His insight and professional style in the teachings of Applied Behavior Analysis laid the foundation for my interests and development of clinical and research skills in behavior therapy and evidence based treatment. His sense of humor and timely levity provided a context in which I was able to more readily learn, take risks, and push myself to participate in new experiences that I may not have been so apt to take on (i.e., presenting a talk at a conference in the face of “internal behaviors” expressing the fight or flight response). Without his dedication and support, this dissertation would not have been possible.

I would also like to thank my committee member, Dr. Steven Kurtz, who accepted me as if I were his own graduate student at the New York University Child Study Center (NYUCSC). His clinical expertise, supervisory style, and faith in me as a clinician provided me with the ability to learn and expand on my clinical and research experiences in behavior therapy and evidence based treatment. His supportive yet challenging supervisory style provided me with the knowledge to learn, but more importantly with a framework of how to “learn to learn” in the field of clinical psychology. This of course, could not have been possible without his supportive personality style, sense of humor, and admirable appreciation for the use of Labeled Praises amongst his patients, colleagues, and trainees. This dissertation would not have been possible without his support, generosity of time, and faith in my ability to take on this project.

In addition, I would like to thank the faculty and graduate and undergraduate students I had the pleasure of working with at the NYUCSC, as well as the teachers and staff at the Little Red School House of the Astor Services for Children and Families. The
clinicians I was trained by, and along side, at the NYUCSC PCIT program were an
ingcredible group. I feel honored to have learned from, and to have spent time with, Drs.
Melanie Fernandez, Melissa Ortega, and Samantha Miller, who provided the support,
humor, empathy, and generosity which made this project possible. Yudelki Firpo donated
countless hours of dedicated and careful data collection and analyses that made this
project possible and undoubtedly more efficient. Drs. Rodney DiMotta, Polina Umylny,
and Arletha Kirby opened their school to us and set a tone that enabled us to engage with
their teachers and students which certainly enhanced the time I spent at the school.

I would also like to thank the faculty and students that I had the pleasure of
working, learning, and laughing with during the course of my graduate studies at GSAPP.
I especially want to thank Nathan Lambright and Jeffery Selman for being excellent
classmates, roommates, and great friends, particularly during our front porch study
breaks. I owe perhaps the greatest thanks to Dr. Sandy Harris, as she provided me with
the life changing opportunity to attend Rutgers and learn from the faculty, staff, and
students at the Douglass Developmental Disabilities Center; thank you for taking a risk
on me and for allowing me to learn and grow from my time with you.

Lastly, and certainly not least, I would like to thank my family who has provided
the definition of unconditional love and support throughout my life and most recently
during my graduate studies at Rutgers. I want to especially thank my family for providing
me with the opportunities that have made this possible. As a student of child psychology,
I would be remiss if I did not thank you for spending those early years reading to me,
setting firm yet warm limits, and of course providing just enough Labeled Praises over
the years to give me the confidence to pursue a doctorate in Clinical Psychology, but not
so much as to facilitate the development of a narcissistic personality. Thank you for all the support, encouragement, and perspective which I needed to complete this project.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
</tbody>
</table>

## CHAPTER

1. **INTRODUCTION**
   - Hypotheses and Predictions: 11

2. **METHOD**
   - Participants: 13
   - Measures: 13
   - Procedures: 16
   - Experimental Design: 20

3. **RESULTS**
   - Interrater Reliability and Time Requirements: 22
   - Labeled Praise: 23
   - CDI Do Skills and Don’t Behaviors: 25
   - Effective and Ineffective Command Sequences: 27
   - Generalization of Labeled Praise: 29
   - Student Behavior (REDSOCS): 32

4. **DISCUSSION**: 35
LIST OF TABLES

Table 1 Interrater Reliability..............................................................22
Table 2 Intervention Time Requirements............................................23
LIST OF FIGURES

Figure 1 Mean rates of in-training Labeled Praise.................................25
Figure 2 Mean rates of in-training CDI Do Skills and Don’t Behaviors.........27
Figure 3 Mean rates of in-training Effective and Ineffective Command Sequences.29
Figure 4 Generalization of in-classroom Labeled Praise............................31
Figure 5 Mean Teacher Report SESBI-R Intensity T-Scores.........................32
Figure 6 Mean in-classroom adaptive & maladaptive student behavior..........34
INTRODUCTION

Disruptive behavior in young children poses significant challenges to families, school personnel, and mental health professionals. Disruptive behavior disorders (DBDs), which include oppositional defiant disorder (ODD) and conduct disorder (CD), are primarily characterized by recurrent patterns of negativistic, defiant, and hostile behavior toward authority figures, and the violation of others’ basic rights or major age-appropriate social norms/rules, respectively (APA, 2000). DBDs affect as many as 16% of children and include significant impairment in social and academic functioning (APA, 2000). Studies documenting prevalence rates of psychiatric disorders among preschool children in particular have found that ODD is “by far” the most common disorder identified, present in 13-17% of children aged two through five, with eight percent of children exhibiting behavior disorders characterized as severe (Lavigne, Gibbons, Christoffel, Arend, Rosenbaum, & Binns et al., 1996; Lavigne, LeBailly, Hopkins, Gouze, & Binns, 2009). DBDs also represent the most common reason for mental health referrals of preschool children (Campbell, 2002). It is also important to consider the frequency with which disruptive behavior occurs in the general population of preschoolers because of developmental changes at this age. Indeed, most preschoolers demonstrate some of the specific behaviors that comprise a behavior disorder (Wakschlag, Leventhal, Briggs-Gowan, Danis, Keenan, & Hill, et al., 2005).

DBDs that emerge early in childhood have been shown to persist over time (Carter, Briggs-Gowan, & Davis, 2004; Wakschlag, Leventhal, Thomas, & Pine, 2007). In addition, there is an increased association between DBDs and depression, substance abuse, ADHD, aggression, school suspension, criminal behavior, personality disorders (Kronenberger and Meyer, 2001), low self-esteem, interpersonal conflicts with parents,
teachers, and peers, increased sexual behavior, sexually transmitted diseases, unplanned pregnancy, legal difficulties, and suicidal ideation, suicide attempts, and suicide completions (APA, 2000). Shaw, Gillion, Ingoldsby, and Nagin (2003) documented that specific risk factors for later, overt antisocial behavior can begin as early as when children are one to two years old. Lavigne, Arend, Rosenbaum, Binns, Christoffel, and Smith (1998) and Lavigne and colleagues (2009) suggested that preschool-aged children with ODD are more likely than typically developing preschool-aged children to develop anxiety and/or depression later. Lavigne and colleagues (2009) further posited that this etiology may be due to one of two factors: anxiety and depression may be present during preschool years but masked or manifested through symptoms of ODD, or symptoms of ODD may produce academic and social problems that give rise to symptoms of anxiety and depression.

Children with DBDs continue to go untreated. Lavigne and colleagues (1998) found that only 11-20% of children with a diagnosis received treatment, and more recent studies documented even lower rates, with only three percent of children with a diagnosis having received mental health services (Lavigne et al., 2009). Therefore, the risk for young children with DBDs to experience increasing difficulties in academic and social functioning remains if these children are not properly treated at crucial stages in their early development. The need for early intervention is clear, as early treatment may prevent the worsening of disruptive behavior and associated persistence. (Lavigne et al., 2009).

DBDs disrupt classroom functioning (Greenlee & Ogletree, 1993). One study found that 48% of teachers of young children indicated the presence of severe disruptive behaviors in their classes and 41% of teachers subsequently reported that the time needed
to attend to these behaviors decreased the time teachers were able to devote to learning (Hart, 1995). Given reported rates of disruptive behavior in the classroom, it is not surprising that teachers report feeling overwhelmed and inadequately prepared to manage such behavior. Merrett and Wheldall (1993) found that while 51% of teachers reported spending too much time on class control, 75% of teachers reported not being prepared to manage children with challenging behaviors, and 72% reported that they were dissatisfied with the level of training provided to deal with such behavior problems. Similarly, another study surveying teachers’ attitudes toward classroom management of student behavior problems found that stress related to lack of training in managing disruptive classroom behavior was the most influential factor in failure among novice teachers. Teachers emphasized the need to develop more skill and competence in curbing such behaviors (Greenlee et al., 1993). Martin, Linfoot, and Stephenson (1999) assessed teachers’ concerns about and responses to misbehavior in the classroom and found that teachers reported a lack of confidence in managing misbehavior, a significant desire to gain information designed to encourage positive behavior in the classroom, and a need for positively focused information/strategies. Teachers were also more likely to refer misbehaving children to other school personnel as well (Martin et al., 1999), which can result in added costs and reduced time spent on learning.

Despite the literature suggesting a need for positive approaches to behavior management and teachers’ desire to gain information on positive approaches, teachers’ likelihood of using non-physical punishment strategies as opposed to positive and negative reinforcement increased as a function of increases in their concerns about misbehavior (Martin et al., 1999). Teachers’ tendencies to employ punishment strategies
represent a significant concern given the relationship between coercive relationships and further maladaptive behaviors (Patterson & Reid, 1970).

In light of the literature suggesting a lack of teacher training and preparedness, it is crucial to develop interventions that are manageable and that can be implemented with sufficient integrity to be effective. In her review of interventions for preschool behavior problems, Campbell (2002) highlighted that teachers who feel overwhelmed by behavior management techniques are less likely to implement the interventions. Greene (1995) indicated that the time demands involved in implementing classroom intervention are particularly important to consider, as teachers tend to prefer time-efficient positive behavioral treatments that target groups rather than individually administered contingencies. Additionally, Martens, Witt, Elliott, and Darveaux (1985) found that teachers rated interventions requiring high amounts of time to implement as less acceptable. Therefore, it seems crucial to provide preschool teachers with additional training focused on efficient and effective interventions.

Despite the clear need for evidence-based treatments (EBTs) for disruptive behaviors in the classroom, little research exists in this area, and fewer studies have focused specifically on preschool classrooms. This lack of attention is particularly concerning given that interventions implemented prior to school age have been documented to have a higher probability of success (Shaw, et al., 2003).

In a review of the literature on psychosocial treatments for ADHD, Verduin, Abikoff, and Kurtz (2008) found that behavioral approaches, including parent training, contingency management, and school-based approaches such as a daily report card (DRC), had the best empirical support. Contingency management strategies have been shown to improve disruptive behaviors and symptoms of ADHD (Barkley, 2000) and
include frequent monitoring and high rates of positive and negative consequences. As such, these interventions become difficult and costly to replicate in community settings (Verduin et al., 2008). Despite this body of evidence, little is known about the efficacy of these interventions with preschool children (Verduin et al., 2008).

Further review of the literature on EBTs for child and adolescent DBDs identified 16 “well-conducted” studies (in accordance with the standards put forth by the APA Task Force on the promotion and dissemination of psychological procedures) that identify six EBTs for young children with DBDs. Most consisted primarily of parent training interventions (Eyberg, Nelson, & Boggs, 2008). The Incredible Years, which consists of three independent training programs (child, parent, and teacher) (Webster-Stratton, 2003), was identified. The child and parent training programs were found to be “probably efficacious,” as per the definitions set forth by the above APA Task Force (Chambless & Hollon, 1998; Eyberg et al., 2008). The teacher program met criteria for “possibly efficacious” (Chambless & Hollon, 1998) when combined with the child and/or parent protocols; it did not meet the necessary standards as a stand-alone intervention (Eyberg et al., 2008). Such findings highlight the gap that exists between EBTs for DBDs in the home versus the classroom setting.

More recently, Teacher-Child Interaction Therapy (TCIT), which is based on Parent-Child Interaction Therapy (PCIT), has been evaluated for use in classrooms. PCIT is a well-established, evidence-based treatment of DBDs in young children (Eyberg et al., 2008). PCIT, largely influenced by Baumrind’s (1967, 1991) developmental theories and Hanf’s (1969) two-stage treatment model (as cited in Eisenstadt, Eyberg, McNeil, Newcomb, & Funderburk, 1993) aims to promote positive child outcomes by fostering an authoritative parenting style through a nurturing parent-child relationship and consistent
parental limit setting (Brinkmeyer & Eyberg, 2003). PCIT is a short-term parenting skills training program that targets parent-child interactions during two phases of treatment: Child Directed Interaction (CDI) and Parent Directed Interaction (Eyberg et al., 2008). In CDI, parents are taught to differentially attend to child behaviors though the use of the “PRIDE” skills (i.e., Labeled Praises, Reflections, Imitation, Behavior Descriptions, and Enthusiasm) and placing maladaptive behaviors on extinction (i.e., Active Ignoring) to establish and strengthen a warm relationship with their children. In PDI, parents are taught a clear and consistent method of implementing positively-stated commands and consequences to foster child compliance (Brinkmeyer & Eyberg, 2003). Following graduation from PCIT, studies have found an increase in parents’ use of positive parenting skills and a decrease in child disruptive and noncompliant behaviors from clinically significant levels to normal ranges (Eisenstadt et al., 1993; Schuhmann, Foote, Eyberg, & Boggs, 1998). Follow-up studies have also demonstrated impressive stability of treatment gains at both one and two years following treatment (Eyberg, Funderburk, Hembree-Kigin, McNeil, Wuerido, & Hood, 2001). Further, reductions in problem behavior have also been documented to generalize from the clinic setting to home and school (McNeil, Eyberg, Eisenstadt, Newcomb, & Funderburk, 1991).

Based on the gap between the number of children with DBDs who do not or cannot access mental health services, there is a compelling need for more efficient and effective interventions in the school setting. Researchers have identified two types of intervention strategies in the classroom: antecedent and consequence strategies (Dupaul and Weyandt, 2009). Antecedent strategies include manipulations of stimuli occurring prior to the behavior of interest, which have been paired with or signal the availability of certain consequences. Consequence strategies include changes in environment that occur
contingent upon a given behavior with the expectation of increasing, maintaining, or decreasing that behavior in the future (Cooper, Heron, Heward, 2007). A classroom version of PCIT could offer an efficient, system-wide intervention that uses both antecedent and consequence strategies and is accessible to the many children in need of services who may otherwise go untreated, while also providing preventative services to non-target children.

A few studies have assessed the effectiveness of TCIT in preschool classrooms. McIntosh, Rizza, and Bliss (2000) conducted a single-subject case study design to assess TCIT, which consisted of two phases: Child-Directed Interaction (five sessions) and Teacher-Directed Interaction (seven sessions). Each phase began with a didactic “teach” session and was subsequently followed by CDI and TDI coaching sessions that were predominantly held outside of the classroom. The teacher was also instructed to practice CDI and TDI skills on a daily basis with the child in the classroom for 5 minutes during the CDI phase and ten-minutes during the TDI phase. Observations of both teacher and child behavior indicated increased teacher use of praise, reflective statements, and descriptive statements and decreased child disruptive behavior and non-compliance relative to baseline. In addition, the assessment of behavior changes occurred in the treatment setting rather than the classroom (McIntosh et al., 2000). Although the study lacked experimental control and generalization of gains to the classroom were unclear, findings offered preliminary support for the use of TCIT in the classroom setting.

In another study of TCIT, Filcheck, McNeil, Greco, and Bernard (2004) compared their model of TCIT to a class-wide token economy approach in a preschool classroom of 17 children. The authors used an ABACC’ design where (A) represented strategies already employed by the teacher or baseline, (B) represented the class-wide
token economy, (C) represented the CDI phase of PCIT, and (C’) represented the PDI phase of PCIT. The teachers received a one-hour CDI didactic session and a one and one-half hour PDI didactic training, each of which was followed by a two-hour skills coaching session. Results demonstrated improvements in both the token economy and PCIT conditions. However, PCIT was demonstrated to be more effective in reducing teacher criticisms directed toward students and child non-compliance above and beyond the gains demonstrated during the token economy condition. As in the McIntosh et al., (2000) study, this study lacked experimental control and generalization of teachers’ skills throughout the entire school day, as observations were contained to circle time; however, the results provided positive preliminary support for further evaluation of TCIT in the classroom.

Tiano and McNeil (2006) compared the use of PCIT skills in four classrooms to three no-treatment classroom control groups. The first training phase consisted of didactics with teachers, assistants, and volunteers in classroom behavior management skills (e.g., strategic attention, selective ignoring, redirection, praise). Training was followed by in-classroom coaching until staff members met a predetermined mastery criterion. The second training phase continued with instruction of classroom behavior management skills (e.g., giving effective commands, handling disruptive behavior, time-out) followed by live coaching of teachers in the classroom until predetermined mastery criteria were met. The study further outlined specific modifications to PCIT for implementation in a Head Start classroom. Modifications included didactic instruction delivered in groups, live coaching in the classroom, training teachers to use PCIT skills to modify group rather than individual behavior, and time-out referred to as the “Thinking Chair.” Results indicated that child inappropriate behavior improved regardless of
treatment group and that teachers in the intervention group used more Labeled Praise than controls after treatment. There were no differences in teachers’ use of Unlabeled Praise or Criticism (Tiano & McNeil, 2006).

Lyon, Gershenson, Farahmand, Thaxter, Behling, and Budd, (2009) further evaluated the effects of TCIT and attempted to expand on past adaptations. The study’s adaptations included: a targeted sample of ethnic minority children and teachers, and emphasis on increased number of training sessions and PCIT skills, observations of teacher behaviors conducted across a number of classroom situations, and a two-stage non-concurrent multiple baseline design across classrooms. Each stage consisted of weekly training of all six teachers from two classrooms for a total of nine 1.5 hour sessions (four CDI, four TDI, and one graduation session). Teachers also received 20 minutes of individualized coaching on their CDI and TDI skills one to three times per week. Coaching began after the third CDI and TDI group sessions. Results demonstrated small to moderate effects in teachers’ use of positive behaviors. In two classrooms, teachers’ use of the CDI skills increased following the introduction of the PRIDE skills and initiation of self-monitoring homework taught during the first CDI Teach session. Although the results of this study offer preliminary data to support the use of TCIT in a preschool setting, the findings are somewhat limited in external and internal validity. The use of a single school limits generalizability of results. A non-concurrent multiple baseline design fails to control for history effects. Additionally, CDI and TDI didactics and coaching phases were time-limited rather than based on mastery criteria as originally intended in PCIT. Thus, the lack of significant teacher behavior change could have been due, in part, to the fact that teachers were not coached to overlearn the skills adequately, an important component of PCIT.
Fernandez and Kurtz (2009) further evaluated the effects of TCIT in four classrooms, two at a day-treatment preschool and two at a therapeutic nursery. Initially, teachers and classroom support staff were provided with eight weekly didactic “teach” session involving CDI and TDI skill use. Each “teach” session lasted one to one and one-half hours. Additionally, teachers were provided with up to six 20- to 30-minute, in-vivo coaching sessions to reinforce their use of CDI and TDI skills. In-classroom behavioral observations were conducted pre- and post- treatment on teachers’ use of CDI “Do” (i.e., Labeled Praises, Reflections, Behavior Descriptions) and “Don’t” skills (i.e., Questions, Commands, Criticisms) using the Dyadic Parent-Child Interaction Scale (DPICS; Eyberg, Nelson, Duke, & Boggs, 2005). Classroom observations were also conducted on child behavior (i.e., appropriate/inappropriate, on-task/off-task, and compliance) using the Revised Edition of the School Observation Coding System (REDSOCS; Jacobs, Boggs, Eyberg, Edwards, Durning, & Querido et al., 2002). Results indicated that teachers at one site increased their use of total “Do” skills while teachers at the second site showed minimal changes. Teachers in three of the four classrooms demonstrated a decrease in their use of total “Don’t” skills. They also demonstrated an increase in their use of Direct Commands and a decrease in their use of Indirect Commands. Regarding child behavior, off-task behavior decreased across all classrooms, inappropriate behavior decreased in two classrooms; noncompliance decreased in one classroom, while it increased in two classrooms and remained the same in the fourth classroom. While the results were variable and the design lacked experimental control, the study offers preliminary support for the use of TCIT in a preschool setting.

Through the ongoing relationship between the New York University Child Study Center (NYUCSC) and a day-treatment preschool, the current study sought to assess the
effectiveness of TCIT in a preschool setting involving training of five head teachers. It should be noted that this study was a continuation of the program evaluation conducted by Fernandez and colleagues (2009) described above, specifically targeting the preschool day-treatment program. The study also sought to expand on prior models of TCIT in several ways. First, the TCIT training procedure maintained, to a greater extent, the integrity of the PCIT treatment protocol (i.e., treatment conducted during individual pull-out sessions, treatment delivered in a TCIT-adapted treatment room, live coding and coaching of skills in-training ). Second, experimental control was maintained through the use of a concurrent multiple probe design across classrooms to control for threats to internal validity. Third, teacher and child classroom behavior was coded parallel to treatment sessions to assess generalization of skills to the classroom. Lastly, to enhance motivation for participation and adherence and to tailor treatment to the needs of each head teacher, brainstorming sessions with all teachers to assess individual classroom concerns and to collect Sutter-Eyberg Student Behavior Inventory (SESBI-R; Eyberg, & Pincus, 1999) teacher report forms.

Hypotheses and Predictions:

Primary Hypothesis: Efficacy. Teachers receiving TCIT will increase their in-training use of Labeled Praises (LP) relative to their own baseline rates of LPs, and relative to concurrent in-training baseline trends in subsequent control classrooms.

Secondary Hypotheses:

Hypothesis two: Teachers receiving TCIT will increase their in-training use of positive behaviors [e.g., Labeled Praises (LPs), Reflections (RF), and Behavior Descriptions (BD)] and decrease their use of in-training negative behaviors [e.g., Criticisms (NT), Commands (C), and Questions (Q)], relative to their baseline rates of
positive and negative behaviors, and relative concurrent in-training baseline trends in subsequent control classrooms (Appendix A).

Hypothesis three: Teachers receiving TCIT will increase their in-classroom use of positive behaviors and decreased their in-classroom use of negative behaviors relative to their own baseline rates of positive and negative behaviors, and relative concurrent in-classroom baseline trends in subsequent control classrooms.

Hypothesis four: Children in classrooms with teachers receiving TCIT will demonstrate decreased rates of in-classroom maladaptive behavior and increased rates of in-classroom adaptive behavior relative to their own baseline rates of adaptive and maladaptive behaviors, and relative concurrent in-classroom baseline trends in subsequent control classrooms.

Hypothesis five: Teachers’ report of student problem behavior will decrease across the three assessment phases (baseline, post-CDI, and post-TDI).
METHOD

Participants

The participants in the current investigation included five classroom teachers in a preschool setting with an ongoing relationship with NYU involving institution-wide training of staff. The preschool is a New York City day-treatment program serving preschoolers with DBDs. The NYU staff included members of the clinical staff of the ADHD and Disruptive Behaviors Institute at the New York University Child Study Center. The NYU Child Study Center staff serving the above client included its director, three predoctoral psychology externs, and three undergraduate research assistants. The director and staff had considerable clinical experience with children with disruptive behavior disorders and in utilizing the Parent-Child Interaction Therapy protocol (Eyberg et al., 1999). The client was a New York City day-treatment preschool who sought consultation from the NYU Child Study Center staff two years prior regarding classroom management. Data collected on the head teachers and students in all five classrooms were used to continue informing clinical treatment as usual and were compiled anonymously to inform educational program development and planning (i.e., treatment did not target specific children, nor were data used to inform individual student treatment plans or diagnoses).

Measures

_Dyadic Parent-Child Interaction Coding System-Third Edition 3.06 (DPICS-III_ 3.06; Eyberg et al., 2005). The DPICS is a behavioral coding system that measures the quality of parent-child social interactions during three five-min observations (Child-Led Play, Parent-Led Play, and Clean-Up) that vary in degree of parental control required. Teacher behaviors that were recorded included Labeled Praise, Unlabeled Praise,
Reflections, Behavior Descriptions, Direct Commands, Indirect Commands, Questions, Criticism (Appendix A), and Effective Command Sequences (Appendix B). Each of the teacher behaviors was coded for frequency of occurrence during three five-minute baseline tasks; five two-minute coding segments across multiple settings within each individual classroom; and during five-minute coding of teacher behavior at the onset of each treatment session. Observations in each classroom were conducted as often as once per week by one graduate and/or one undergraduate research assistant and were approximately 10-12 minutes in length. Scheduling conflicts, due to school vacations, periodically interfered with coders’ ability to conduct observations each week. Observers were trained in the DPICS-III through didactic training meetings that consisted of reviewing the DPICS-III manual in detail, practicing coding from mock circle-time videotapes, and completing homework assignments and reviews of quizzes from the DPICS-III workbook. The training clinician, a predoctoral psychology extern, conducted observations on teacher behaviors during pull-out baseline tasks and during pull-out in-training observations.

Revised Edition of the School Observation Coding System (REDSOCS; Jacobs et al., 2002). The REDSOCS is a 10-second partial interval coding system used to measure student classroom behavior. Participants were observed for 10-second partial intervals for a total of one minute per child, for a total of five to six minutes per classroom/per week. Observations occurred across multiple activities and settings within the classroom. One graduate and/or one undergraduate research assistant conducted weekly observations; observations targeted five to six children selected at random each week, totaling five to six minutes per classroom. Observers were trained in the REDSOCS through didactic training meetings that consisted of reviewing the REDSOCS manual in detail, practicing
coding from mock circle-time videotapes, and completing homework assignments and
review of quizzes from the REDSOCS manual. Percentages of child compliance to
commands, appropriate, inappropriate, off-task, on-task, aggressive, and disruptive
behaviors were obtained by dividing the number of intervals coded by the total number of
intervals. Of note, aggression and disruption were added to traditional REDSOCS
categories to differentiate more severe “inappropriate” behaviors. Traditional REDSOCS
coding includes aggressive and disruptive behaviors in the category of “inappropriate”
behavior, which also includes mild behaviors such as tapping a pencil. Observers coded
teacher behaviors using the DPICS, and child behaviors using the REDSOCS, in
alternating sequence, such that the primary teacher was coded for two minutes, followed
by coding child one for one minute, followed by coding the primary teacher for two
minutes, followed by coding child two for one minute, etc.

*Sutter-Eyberg Student Behavior Inventory-Revised* (SESBI-R; Eyberg & Pincus,
1999). The SESBI-R is a 36-item teacher report inventory of classroom conduct problem
behaviors. Teachers were instructed to indicate how often each behavior-item currently
occurs, yielding an intensity score, and whether the behavior is currently a problem for
the teacher, yielding a problem score. The mean intensity score for three to five year-old
children in the standardized sample was 100.9 (SD = 47.6), and the mean problem score
was 6.0 (SD = 8.8). High internal consistency, inter-teacher agreement, and test-retest
stability have been demonstrated for the Intensity Problem Scales. Significant
correlations have been found between the SESBI-R and several school rating scales
including the RCTRS and the Behar Preschool Behavior Questionnaire (Funderburk &
eyberg, 1989). The SESBI-R was administered to each lead teacher at baseline (prior to
Procedures

Interrater agreement. Interrater agreement data were collected for a minimum of 20% of the REDSOCS in-classroom observations. Interrater agreement data were collected for a minimum of 75% of the DPICS baseline and in-training coding observations. Baseline and in-training DPICS coding were audiotaped and coded by a blind observer for reliability; target interrater agreement for in-classroom DPICS and REDSOCS and in-training DPICS was set at a minimum of 80%. It was determined that if interrater reliability fell below 80%, a second blind observer would code discrepant audio files to ensure accuracy, and the discrepant coder would receive immediate supplemental DPICS training by a master PCIT coder until 80% reliability was restored.

Training Process: Prior to training implementation, each classroom was randomly assigned to one of the three training groups. Group one and two each consisted of two classrooms, and group three consisted of the final, fifth classroom. The training clinician then met as a group with all five lead teachers to discuss training goals, assessed variables from TCIT coaching conducted during previous years that the teachers found helpful or discouraging, and clarified misconceptions to improve transparency and motivation for training.

Subsequently, data were collected to establish pre-training baselines of student and teacher behaviors until a stable trend of Labeled Praises was established (minimum of two data points per teacher). Baseline data were collected on teacher behaviors “in-training” (DPICS Child-Led condition during pull-out sessions with trainer, teacher, and an individual student). Baseline data were also collected on teacher behaviors “in-
classroom” (DPICS observations in the classroom setting throughout school day). Baseline data were collected on child behaviors “in-classroom” as well (REDSOCS observations in the classroom setting throughout the school day). DPICS baseline tasks and DPICS coding were conducted during training sessions were audiotaped and coded for interrater reliability by a blind coder. Classroom baseline observations were collected by parallel coders, blind to each others’ results, to assess for interrater reliability (coders shared an audio timer that was split through a headphone extension cord to ensure coders were blind to each others codes). A minimum of 80% interrater reliability was determined to be acceptable for the current study. The SESBI-R was also administered to all five lead teachers for each student in their class, prior to training, following the completion of CDI, and following the completion of TDI.

Following baseline data collection, Group one, which consisted of two classrooms, received training during pull-out sessions while blind coders conducted weekly classroom observations of teacher and child behavior across all five classrooms. Thus, classroom observations for the training classrooms served as generalization probes, while the classroom observations for the non-training classrooms served as concurrent baselines, controlling for factors external to the TCIT intervention. A minimum of 20% of classroom observations were double coded for interrater reliability. Also, following CDI Mastery, in-classroom data on CDI Don’t Behaviors coded but not included in data analyses, as TDI procedures trained teachers to implement Direct Commands (i.e., CDI Don’t Behavior data would not be comparable from CDI to TDI).

Training was largely conducted in accordance with the procedures outlined in the PCIT manual (Eyberg and the Child Study Laboratory, 1999). The following components of PCIT were adapted from the PCIT manual: First, each teacher was instructed to
identify one child for whom they would like to be coached with each week. PCIT protocol dictates that the same child in the dyad remains constant throughout the duration of training. Due to student absences and academic schedules, which would have interrupted teacher training, this aspect of the protocol could not be retained; instead, teachers were coached with a student of their choice each week. However, it should be noted that this deviation was accounted for in two ways. One goal of traditional CDI is to enhance attachment between a child and caretaker. While teachers were encouraged to remain with the same child, variations did occur. Thus, teacher-student attachment may not develop in the same course as is expected in traditional PCIT (i.e., attachment may not improve until a teacher begins using CDI consistently in the classroom with all students). Secondly, when TDI was introduced all students in the classroom were provided with an overview of the time-out procedure to ensure clarity of expectations. This is an important distinction, since traditional PDI dictates that children are made aware of the compliance and time-out expectations.

Second, live coaching was conducted with the therapist in the training room rather than using a “bug-in-ear” approach because no two-way mirror or video surveillance were available.

Third, while CDI coaching sessions followed the outlines provided in the PCIT manual, instruction around the use of Questions was adapted to meet the needs of an academic setting. During CDI teach, teachers were taught to differentiate between general questions and academic questions (i.e., “questions that assess a child’s knowledge of a specific academic subject matter”). Teachers were still coached to avoid all questions during CDI coaching sessions, as per the PCIT manual; they were instructed to utilize only educational questions during classroom activities. Thus, CDI mastery criteria
remained consistent with the guidelines outlined in the PCIT manual (i.e., teachers were required to provide a minimum of 10 of each of the three CDI “Do” skills, and three or fewer CDI “Don’t” behaviors.

Fourth, to permit necessary follow-through with time-out procedures as outlined in the PCIT manual, the time-out procedure and time-out backup room procedure was adapted to accommodate administrative restrictions concerning implementation concerning physical prompts. Regarding the time-out procedure, the “sit on the chair” consequence was replaced with a “contextually relevant consequence,” in which the teacher was able to remove “positive reinforcement” without physical prompts to the child. For example, the consequence paired with a command to “color on the paper” would be the removal of the coloring activity [i.e., “If you don’t color on the paper, coloring is finished” (contextually relevant consequence)]. Additionally, the training room was used as a time-out backup room for the purposes of implementing a secondary consequence during training sessions. The teacher was coached to remove the furniture and toys from the training room upon exiting the room to begin the time-out backup room sequence. Regarding time-outs issued in the classroom, teachers were coached to issue the statement, “time-out begins when you are sitting on the chair.” Teachers were subsequently coached to actively ignore all other behavior. The classroom was used as a time-out backup room if a student broke a “house rule” and thereby presented a danger to themselves or other students in the classroom; the remaining students were subsequently escorted out of the classroom to a temporary classroom setting or adjacent classroom.

Finally, at the completion of training teachers completed the third SESBI-R. Teachers were also provided with the opportunity to discontinue training (depending on
SESBI-R scores) or to continue coaching in the training room with multiple students or in their respective classrooms.

**Experimental Design**

Using a concurrent multiple probe design, the current study compared the effects of the independent variable, TCIT, on the primary dependent variable, teachers’ use of Labeled Praise, across classroom settings (Baer, Wolf, and Risley, 1968; Kazdin, 1982; Cooper, et al., 2007). Data were also collected on secondary dependent variables. However, it should be noted that changes in training phase were dependent upon Labeled Praise data only. All five classroom teachers were randomly assigned to one of three training groups.

Multiple probe designs assess and compare, through visual graphic inspection, data collected on operationalized behaviors’ mean, trend, level, and latency of change of operationalized behaviors over time. Initially, concurrent baseline data is collected for all groups until a stable trend is established. The independent variable is then applied to the target group and data collected following this introduction are then compared to the established baselines within the target group (pre-training baseline) and between the target group and the concurrent baselines of the remaining control groups. Accordingly, changes that occur following the implementation of the independent variable on the first subject, are graphically compared to that group’s baseline data and the subsequent groups’ concurrent baselines to rule out confounding variables (e.g., history effects). Further, the independent variable (TCIT) is not applied to control groups until 1) a stable change in mean, trend, level, and/or latency of the dependent variable (Labeled Praise) has been observed in the target training group and 2) stable baselines persist in the control group. The current study utilized a baseline probe prior to implementing the
independent variable to subsequent control groups. If a baseline probe was discrepant from previous baseline levels and/or trends, additional baseline data would have been collected to establish a new stable baseline. Of note, this criterion was not utilized when determining when a given classroom shifted from the CDI phase to the TDI phase of training; rather, the PCIT mastery criteria were utilized for this purpose (10 Labeled Praises, 10 Reflections, 10 Behavior Descriptions, and three or fewer Commands, Questions, and Criticisms in five-minutes). Classrooms within a given training group were assessed independently with regard to CDI and TDI mastery criteria and progressed to subsequent training phases independently. For additional explanation regarding multiple probe designs and visual inspection criteria, please refer to Baer (1968), Cooper et al., (2007), and Kazdin (1982).
RESULTS

*Interrater Agreement:* In-training DPICS interrater agreement occurred in 40 of 53, or 75%, of opportunities and resulted in a mean interrater agreement of 87% overall; 86% during CDI sessions; 87% during TDI sessions; 88% during CDI mastery sessions; and 93% during TDI mastery sessions (Table 1). In-classroom DPICS interrater agreement occurred in 23 of 52, or 44%, of opportunities and resulted in a mean interrater agreement of 86% overall. In-classroom REDSOCS interrater agreement occurred in 109 of 306, or 36%, of opportunities and resulted in a mean interrater agreement of 96% overall. Per interrater agreement guidelines set forth at onset of study, interrater agreement exceeded or met minimum expectations and therefore remedial training was not required at any point during this study.

Table 1

*Interrater Reliability*

<table>
<thead>
<tr>
<th>DTIC Coach In-Treatment</th>
<th># Sessions</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>13</td>
<td>95%</td>
</tr>
<tr>
<td>CDI</td>
<td>21</td>
<td>86%</td>
</tr>
<tr>
<td>CDI Mastery Session</td>
<td>1</td>
<td>88%</td>
</tr>
<tr>
<td>TDI</td>
<td>19*</td>
<td>87%</td>
</tr>
<tr>
<td>TDI Mastery Session</td>
<td>1</td>
<td>93%</td>
</tr>
<tr>
<td>Total # TCT Sessions</td>
<td>53**</td>
<td>NA</td>
</tr>
<tr>
<td>Total # Coded For Reliability</td>
<td>40 (75%)</td>
<td>87%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DTICS In-Class</th>
<th># Sessions</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sessions Coded</td>
<td>52</td>
<td>NA</td>
</tr>
<tr>
<td>Total Sessions Coded for Reliability</td>
<td>23 (44%)</td>
<td>86%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>REDSOCS In-Class***</th>
<th># Sessions</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sessions Coded</td>
<td>306</td>
<td>NA</td>
</tr>
<tr>
<td>Total Sessions Coded for Reliability</td>
<td>109 (36%)</td>
<td>96%</td>
</tr>
</tbody>
</table>

* Represents # TDI sessions less TDI Coach 1 (no coding).
** Represents total # sessions less CDI TDI Teach and TDI Coach 1 sessions.
*** Represents REDSOCS data collected plus disruptions and aggressions.
Table 2

*Intervention Time Requirements*

<table>
<thead>
<tr>
<th>Class</th>
<th>HL (5 conditions @ 30 Min each)</th>
<th>CDI Teach (30 Min)</th>
<th>CDI Coach (30 Min)</th>
<th>TDI Teach (30 Min)</th>
<th>TDI Coach (30 Min)</th>
<th>Total # Sessions</th>
<th>Total Time (hrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>18</td>
<td>8.5</td>
</tr>
<tr>
<td>Class 2</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>12</td>
<td>5.5</td>
</tr>
<tr>
<td>Class 3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>13</td>
<td>5.75</td>
</tr>
<tr>
<td>Class 4</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>14</td>
<td>6.75</td>
</tr>
<tr>
<td>Class 5</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td>4.75</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>5</td>
<td>21</td>
<td>5</td>
<td>24</td>
<td>68</td>
<td>30.75</td>
</tr>
</tbody>
</table>

*Primary Dependent Variable – Labeled Praise:* Teachers’ in-training rates of Labeled Praise were assessed by examining mean, trend, level, and latency across training phases within and between each of the three training groups (Figure 1). All three groups demonstrated low mean levels of Labeled Praise with either stable or decreasing trends (overall baseline mean LP = 1.56). Group three initially demonstrated a higher level of Labeled Praise (6 LPs) during the initial baseline assessment; however, this level was followed by a decreasing and ultimately stable trend, matching the other two training groups’ level and trend of Labeled Praise prior to the CDI condition.

During the CDI phase of training, mean rate of Labeled Praise increased in all three groups from an overall mean of 1.56 to 9.32. Of note, all five classrooms in the three training groups were coached to CDI mastery criteria. Regarding trend, all three
groups demonstrated changes from horizontal or decreasing trends in Labeled Praise at baseline to increasing trends in Labeled Praise during the CDI phase of training. Each of the three training groups demonstrated a minimum rate of 10 Labeled Praises prior to the implementation of CDI in the subsequent training group. Regarding latency of behavior change, all three groups experienced an increase in level of Labeled Praise immediately following CDI Teach that was greater than prior baseline values. These changes were not observed during the concurrent baseline phases between groups. Thus, a causal relationship is inferred between the implementation of the CDI phase of TCIT and the subsequent increasing mean, horizontal trend, positive rise in level, and the brief latency of teachers’ use of Labeled Praise.

During the TDI phase of training, mean rate of Labeled Praise increased further in all three groups from an overall CDI mean of 9.32 to 13.75. Regarding trend, all three groups demonstrated stable or increasing trends in rate of Labeled Praise following the onset of the TDI phase of training. Regarding level, all three training groups demonstrated levels of Labeled Praise equal to, or greater than, levels of Labeled Praise demonstrated during CDI mastery. This is particularly notable, given that teachers were no longer being coached specifically toward CDI mastery, yet continued to demonstrate training gains post-CDI. In fact, overall mean rate of Labeled Praise demonstrated by teachers during TDI mastery sessions increased to 15.4, with one classroom increasing in level to 22.0 Labeled Praises. Regarding latency, two out of three groups demonstrated greater latency of Labeled Praise responding during the TDI phase. As teachers’ rates of effective command sequences improved, their use of Labeled Praises increased accordingly.
Secondary Dependent Variable- CDI Do Skills and Don’t Behaviors: Teachers’ in-training rate of the CDI “Do Skills” and “Don’t Behaviors” were assessed by examining mean, trend, level, and latency across training phases within and between groups (Figure 2). All three groups demonstrated low baseline mean rates of CDI “Do Skills” (overall baseline mean “Do Skills” = 8.6). Group Three demonstrated an initially higher level of CDI “Do Skills” than “Don’t Behaviors” during the first baseline condition but leveled out to an equal, low level, at the second baseline condition. All three groups demonstrated either stable or decreasing trends of “Do Skills” at baseline.

During the CDI phase of training, mean rates of “Do Skills” increased in all three groups from an overall mean of 8.6 to 38.0; mean rates of “Don’t Behaviors” decreased
in all three groups from an overall mean of 11.7 to 5.0. Regarding trends in “Do Skills,”
all three groups demonstrated a shift from stable or decreasing trends at baseline to
increasing trends during the CDI phase of training. Trends in “Don’t Behaviors” shifted
from horizontal to decreasing trends from baseline to the CDI phase, respectively.
Marked shifts in levels of “Do Skills” was observed in two out of three groups, with the
third group demonstrating a more gradual increase in “Do Skills” during the CDI phase.
Gradual decreases in “Don’t Behaviors” were observed from Baseline to CDI mastery.
Class three and five experienced the shortest latency in behavior change, as mastery was
reached in both classes in three coaching sessions. Classes one, two, and four required
five CDI coaching sessions before CDI mastery criteria were obtained.

During the TDI phase of training, overall mean rates of “Do Skills” decreased
from the CDI phase from 38.0 to 28.0. Rates of “Do Skills” decreased in level in all three
groups immediately following introduction of TDI, and subsequently maintained stable
horizontal trends greater than levels observed during baseline. It should be noted, “Don’t
Behaviors” were not coded during TDI sessions, as coaching in TDI specifically taught
teachers to use Direct Commands whereas during CDI coaching taught teachers to avoid
using Direct Commands. Thus, data on “Don’t Behaviors” would not be comparable
across the two training phases.
Secondary Dependent Variable- Effective and Ineffective Command Sequences:

Teachers’ rates of Effective and Ineffective Command Sequences were assessed by examining mean, trend, level, and latency across training phases within and between each classroom individually (Figure 3). Note, data on Effective and Ineffective Command Sequences were assessed in each classroom individually (rather than training groups), as the TDI phase of training was introduced based on individual classroom mastery criteria and not dependent on trends of Labeled Praise. All five classrooms demonstrated stable (low to zero level) rates of Effective Command Sequences at baseline; only one Effective Command Sequence was observed during all baseline observations.
During the CDI phase of training, rates of Effective Command Sequences remained at a horizontal trend and zero level; no Effective Command Sequences were observed during CDI. It should be noted, teachers were being coached to avoid all commands during this phase of training and could not meet mastery until three or fewer total “Don’t Behaviors” were observed. Rates of Ineffective Command Sequences decreased from an overall mean of 8.6 at baseline to 2.2 during CDI. Class four demonstrated the most immediate decrease in level of Ineffective Command Sequences, whereas the other classes demonstrated a more gradual decrease over two to five sessions.

During the TDI phase of training, Effective Command Sequences demonstrated increasing trends and improved from 0.08 at baseline and 0.0 at CDI to 4.0 during TDI. Class five demonstrated a marked increase in level of Effective Command Sequences from 0.0 at CDI to 5.0 during the first coded TDI session. Regarding latency to mastery, teachers required an average number of 4.8 TDI coaching sessions and ranged from three to nine sessions. Overall mean rates of Ineffective Command Sequences remained below baseline rates of 8.6 but increased from 2.2 at CDI to 6.5 at TDI.
Secondary Dependent Variable - Generalization of Labeled Praise to the Classroom

Setting: Generalization of Labeled Praise to the classroom was assessed by examining in-training and in-classroom rates of Labeled Praise within each group (Figure 4).

All three groups demonstrated greater baseline mean rates of Labeled Praise in the classroom setting (mean in-classroom LP = 8.0, 5.0, and 7.0, respectively) than in the training setting (mean in-training LP = 1.0, 0.5, and 3.0, respectively). From baseline to CDI, all three groups demonstrated an increase in overall mean rates of Labeled Praise in the classroom setting, from 8.0 to 9.6 in Group One; 5.0 to 8.0 in Group Two; and 7.0 to 12.0 in Group Three. Groups Two and Three demonstrated a decrease in overall mean rates of Labeled Praise from CDI to TDI (9.6 to 5.3 in Group One, and 8.0 to 6.6 in
Group Two). In-classroom coding was not collected during TDI for Group Three due to high rate of teacher absences.

Regarding trend, Group One demonstrated a decreasing trend in rates of Labeled Praise in the classroom setting overall, as compared to an increasing trend in the rates of concurrent Labeled Praise in the training setting. Group Two demonstrated an increasing trend in the rate of Labeled Praise observed in both settings, with a marked increase in level of Labeled Praise following the CDI phase in the training setting and a marked increase in level of Labeled Praise following the TDI phase in the classroom setting. Group Three demonstrated the most notable generalization effects. Rates of Labeled Praise were low in both the classroom and training setting at baseline. Following the CDI phase of training, rates of Labeled Praise demonstrated a steadily increasing trend in the classroom and the training setting. These gains were maintained through the TDI phase in the training setting, however, in-classroom data collection was discontinued prior to initiating the TDI phase in Group Three.
Secondary Dependent Variable – Teacher Report SESBI-R Intensity T-Scores:

Teacher reports of classroom problem behaviors were assessed through comparison of mean SESBI-R Intensity T-Scores within each classroom at baseline, following CDI, and following TDI. SESBI-R intensity T-Scores of 60.0 and higher are considered to be in the “clinical” range (Figure 5). Results were variable and documented inconsistent findings.

At baseline, classrooms one and four demonstrated mean SESBI-R Intensity T-scores in the “clinical” range (T-scores of 61.0). Classrooms two, three, and five demonstrated baseline T-scores below the clinical range (59.0, 55.0, and 51.0, respectively). Following the CDI phase of training, results indicated a slight increase in mean T-Scores from baseline to CDI in three out of four classrooms; classroom three demonstrated a slight
decrease from baseline to CDI. Following the TDI phase of training, classrooms one and five demonstrated slight decreases in T-scores, while classroom four remained unchanged and classrooms two and three demonstrated a slight increase.

Figure 5. Mean teacher report SESBI-R Intensity T-Scores.

Secondary Dependent Variable- Observation of Student Classroom Behavior (REDSOCS): Student classroom behavior was assessed through comparison of mean percentages of Off-Task, Inappropriate, Disruptive, and Aggressive Behaviors during observations conducted by blind observers utilizing the REDSOCS at baseline, CDI, and TDI phases (Figure 6). Overall data were variable and occurred at low percentages throughout all three conditions, ranging from 0% to 22% of intervals observed.
Results indicated a decrease in Off-Task Behavior from baseline to CDI in classrooms one, two, and five; classrooms three and four demonstrated an increase in Off-Task Behavior from baseline to CDI. From CDI to TDI, classrooms two and five demonstrated a decrease in Off-Task Behavior while classes one, three, and four demonstrated increases during this time.

Regarding Inappropriate Behavior, results were variable and indicated an increase in mean number of Inappropriate Behavior from baseline to CDI in four out of five classes; classroom five demonstrated a slight decrease. From CDI to TDI, four out of five classes demonstrated improvements in Inappropriate Behavior, while classroom three demonstrated an increase during this time.

Aggressive and Disruptive Behaviors were assessed, although it should be noted that these behaviors were tracked as distinct behavioral categories rather than included in the category of Inappropriate Behaviors as operationalized per the REDSOCS. No instances of Aggression were observed during in-classroom observation probes across all three phases. Regarding Disruptive Behaviors, results were also variable and documented overall low base rates at baseline in all five classrooms. From baseline to CDI, classrooms one and three remained unchanged; classroom five demonstrated a decrease; and classrooms two and four demonstrated an increase in disruptions. From CDI to TDI, classes one, two, and five demonstrated improvements from baseline and CDI phases; classroom four demonstrated a slight decrease from CDI but remained above baseline rates; classes three and four demonstrated an increase in Disruptive Behaviors.
Figure 6. Mean in-classroom percentage of adaptive & maladaptive student behavior.
DISCUSSION

Disruptive behavior in young children poses significant challenges to families, schools, and mental health professionals, and results in considerable social and academic impairment (APA, 2000). Among preschool children, DBDs represent the most common reason for mental health referrals (Campbell, 2002) and occur with considerable frequency in the general preschool population due to developmental changes at this age (Wakschlag et al., 2005). Preschool teachers represent an important intervention point for preventing and targeting DBDs, as they have consistent access to preschoolers, DBDs interfere with learning, and contribute to teacher anxiety and burnout; teachers report not only being unprepared to manage challenging behaviors, but also a lack of manageable classroom behavioral interventions (Hart, 1987; Marrett & Wheldall, 1993; Greenlee & Ogletree, 1993; Campbell, 2002). Various adaptations of PCIT have shown promising results for targeting challenging behaviors in the preschool classroom setting (Fernandez & Kurtz, 2009; Lyon et al., 2006; Filcheck et al., 2004; McIntosh et al., 2000; Tiano & McNeil, 2006). The current study sought to expand on prior studies by evaluating the efficacy of TCIT with increased fidelity to the PCIT protocol, delivered to five day-treatment classrooms in an urban, ethnically diverse, and socioeconomically disadvantaged sample, utilizing a concurrent multiple probe design across settings. Interrater reliability collected for in-training DPICS coding and in-classroom DPICS and REDSOCS coding demonstrated minimum or greater reliability standards. Therefore, results should be considered a valid assessment of behavior change, per the DPICS and REDSOCS coding criteria. Results provided support for the primary hypothesis that TCIT would increase teachers’ rates of Labeled Praise during pull-out training sessions. Results also provided support for the first two secondary hypotheses, that TCIT would
increase teachers’ overall Do Skills and Effective Command Sequences while decreasing their Don’t Behaviors and Ineffective Command Sequences in the training setting. Results pertaining to the hypothesis that the effects of TCIT in-training would generalize to teacher and student in-classroom behavior proved variable and largely inconclusive; data demonstrated the generalization of Labeled Praise to three out of five classrooms. The results of the intervention on child in-classroom behavior were inconclusive, as classroom observations indicated an absence of teacher behavior change and, not surprisingly, a lack of secondary behavior change in child in-classroom behavior.

Results pertaining to the primary hypothesis that TCIT would increase teachers’ use of Labeled Praise during pull-out sessions, demonstrated a causal relationship between the onset of TCIT and increases in Labeled Praise across all three groups. More specifically, trend and level analysis indicated immediate positive behavior change in all three groups, demonstrating that the CDI Teach session alone was sufficient to produce initial behavior change. However, latency analysis indicated that teachers required between three and five CDI coaching sessions before clinically significant changes were demonstrated (i.e., CDI mastery criteria were attained), indicating that in-vivo CDI coaching sessions were crucial for teachers to attain mastery criteria. Interestingly, gains in Labeled Praise were maintained or continued to rise through the TDI phase of training. This is notable, as coaching no longer explicitly emphasized the use of Labeled Praise, yet teachers continued to implement Labeled Praise while learning to implement Effective Command Sequences. While Labeled Praise is one required component of an Effective Command Sequence, the observed rates of Labeled Praise during TDI exceeded the rate commensurate with the number of Effective Command Sequences. Thus, teachers continued to deliver Labeled Praise between command sequences, as the ease of
use and efficacy appear to have maintained their use of Labeled Praise during TDI. This is important, as the TDI phase of training relies on students’ motivation for positive reinforcement. In sum, the use of Labeled Praise appears to be a promising and feasible skill-set to implement in the day-treatment setting.

Regarding the secondary hypotheses, it was not possible to draw conclusions as to the causal relationships between TCIT and teacher and child behavior change, as the current multiple probe design established experimental control utilizing data on Labeled Praise alone. Thus, conclusions around the secondary variables are based on correlational relationships relative to introduction of CDI and TDI interventions.

Rates of overall Do Skills and Don’t Behaviors improved following the implementation of CDI (i.e., Do Skills increased and Don’t Behaviors decreased). Trend and level analysis indicated a marked behavior change following the introduction of CDI from baseline. This pattern indicates a strong relationship between the introduction of CDI and increase in positive and decrease in negative teacher behaviors. Latency analysis followed a similar pattern to that observed in Labeled Praise, indicating a positive effect of the CDI Teach session alone in producing initial behavior change and a further enhanced effect following in-vivo coaching. In-vivo coaching seems especially important in reducing Don’t Behaviors, as a clinically significant decrease in such behaviors resulted from coaching over time. Regarding the TDI phase, Don’t Behaviors were no longer coded, as TDI procedures involved instruction in the use of Direct Commands (one of the CDI Don’t Behaviors). Thus, data would not have been comparable. Changes in Do Skills following the introduction of TDI followed a somewhat different pattern from Labeled Praise alone. While all groups maintained gains relative to their baselines, trend and level analysis indicated that rates of Do Skills dropped off following the
introduction in TDI. This finding may suggest that as instruction shifted to the use of Effective Command Sequences, teachers appeared to find Behavior Descriptions and Reflections less salient when confronted with the added verbal demands of the TDI command sequence. This supports the rationale put forth in PCIT regarding the over-learning of skills to combat drift. This may also indicate that teachers find the use of Labeled Praise to be more salient in attempting to gain compliance during limit-setting procedures.

All three groups demonstrated reductions in the use of commands to low or zero rates (three or fewer Don’t Behaviors at CDI mastery) during the CDI phase and an inverse relationship in trends of effective and ineffective command sequences following the introduction of the TDI phase of training. Thus, the introduction of TDI was related to increases in effective commands, decreases in ineffective commands, and overall improvements in teachers’ ability to set clear and consistent limits. Although clear improvements were eventually demonstrated, trend and latency analysis indicated little gain immediately following the TDI Teach session. In all five classrooms, teachers demonstrated increases in both Effective and Ineffective Command Sequences immediately following TDI Teach. Thus, subsequent in-vivo coaching was crucial in developing teachers’ ability to implement Effective Command Sequences while also reducing or eliminating the use of Ineffective Command Sequences. Thus, data suggest that without TDI coaching sessions, teachers may have regressed to their baseline rates of Ineffective Command Sequences and therefore continued to intermittently reinforce non-compliance.

Regarding generalization to the classroom setting, teachers’ rates of overall Do Skills and Don’t Behaviors were inconsistent and did not demonstrate clear changes in
mean, trend, level, or latency analysis following the introduction of TCIT in the training setting. Thus, teachers’ abilities to reach CDI and TDI mastery criteria in training sessions did not spontaneously generalize to the in-classroom settings. However, when rates of Labeled Praise were assessed alone, Groups Two and Three demonstrated corresponding increasing trends in rates of Labeled Praise in-training and in-classroom from baseline through TDI. Group one demonstrated an increasing trend in Labeled Praises in-training while in-classroom Labeled Praises displayed a decreasing trend. These results indicate inconsistent yet promising data to suggest the possible spontaneous generalization of Labeled Praise to the classroom setting. Overall, these data indicate the need for explicit coaching in the setting in which a teacher is expected to utilize the skills. One possible explanation is that teachers may become overwhelmed when managing more than one child, causing them to revert to previous behavior patterns. Alternatively, child behavior in the classroom setting may have shaped teachers’ behaviors (i.e., teachers were negatively reinforced by child behaviors, resulting in poor adherence to the TCIT skills). Also, assistant teachers were not included in the training process. As such, inconsistency of strategies implemented in the classroom may have interfered as well. Finally, the severity of problem behavior in a day-treatment setting may interfere in teachers’ ability to integrate a new skill set in a stressful environment.

Results were variable regarding teachers’ SESBI-R reports of students’ behaviors and REDSOCS data on in-classroom child behaviors. The SESBI-R norms were developed in a general education setting and therefore behavior norms may not be comparable to day-treatment settings. That is, teachers in a day-treatment setting may possess a higher threshold for problem behavior, which therefore limits the sensitivity of the measure. A day-treatment classroom consists of a more homogenous set of behavioral
norms compared to a general education setting in which disruptive behaviors stand out more. This paradigm may have also impacted our understanding of the REDSOCS data on students’ in-classroom behaviors. In addition, REDSOCS observations were collected during short intervals and may not have been conducted frequently enough to capture subtle or low frequency behavioral changes (e.g., disruptive and aggressive behaviors). Additionally, child behavior changes in the classroom setting were assessed to evaluate possible effects of teachers’ use of TCIT skills on child behaviors. However, teachers did not effectively generalize TCIT skills to the classroom overall. Therefore, child behavior change may not have been observed simply because the independent variable (TCIT) was not implemented in the classroom setting. Thus, data on the effects of TCIT on child behavior in the classroom were inconclusive.

**Limitations**

While threats to the internal validity of TCIT’s affect on the primary dependent variable, Labeled Praise, were well controlled through a concurrent multiple probe design, confounding variables affecting the secondary dependent variables cannot be completely ruled out. Due to time limitations, data from the five classes were collapsed into three groups where the sums of two classrooms formed the data of Groups One and Two; this may have reduced the study’s sensitivity in identifying changes across the various dependent variables. The same held true for data collected on SESBI-R teacher reports forms and REDSOCS student behavior. Analysis of individual student behavior change on both SESBI-R and REDSOCS data may have yielded more consistent data. Similarly, in-classroom coding by blind coders was conducted during brief time intervals and with students perhaps aware of coders in the room. Therefore, data may have been insensitive to subtle behavior changes and skewed by observation effects. The fact that
the student each teacher worked with in-training varied may have had an impact on teacher-child attachment. Also, teachers were required to participate in the current study as part of an overall program evaluation and intervention. Thus, results may have been adversely affected by teachers’ degree of motivation to implement the skills learned in the classroom setting. Additionally, results represent effects of TCIT as adapted and delivered in the current study. That is, while in-training DPICS teacher behaviors were coded for reliability, data were not collected on the fidelity of the TCIT intervention. Thus, consistency between replications would be difficult to measure and the implications of the current study on PCIT procedures should be interpreted with caution.

Regarding threats to external validity, the current study consisted of five teacher participants in only one day-treatment pre-school program in a socioeconomically disadvantaged and ethnically diverse setting. Therefore, results may not be generalizable to additional academic settings and populations (e.g., general education preschool). The resources required to implement the current intervention also represent a limitation, as the time, cost, and materials required pose considerable demands. TCIT, as implemented in the current study, required an individualized training room free of extraneous or potentially hazardous stimuli. This is important, as elements of a classroom environment may pose safety risks in which ignoring disruptive behaviors (e.g., climbing, electrical outlets, throwing heavy objects, etc.) while differentially attending to alternative or adaptive behaviors may be risky, dangerous, or impractical. The current study also required a trainer with expertise in PCIT and coders with training in the DPICS and REDSOCS coding systems. While the study involved informal discussions with an in-house trainer to promote sustainability, developing a manualized protocol would reduce costs and increase sustainability over time. Generalization data were limited in that
teachers were not coached in-classroom to implement the skills, and consisted of a cohort that was mandated to receive the training. Teachers’ limited use of skills in the class may indicate the need for direct in-classroom coaching after CDI and TDI mastery. This limited use may also be due to a lack of motivation to utilize the skills in the absence of direct instruction and/or reinforcement. Another limitation is the absence of data on teacher acceptance of the intervention. Relationships among latency to mastery, generalization of behavior, and teachers’ acceptance are unknown. The teachers’ anecdotal reports also indicated that, despite attaining mastery criteria during pullout sessions, they did not feel equipped to implement the skills during the inherently more stressful in-classroom situations. Teachers explicitly asked for additional training in “how” to use the skills in specific classroom situations. Similarly, data on in-classroom student behavior were variable and inconclusive due to teachers’ limited use of the skills in the classroom. Also, low frequency yet severe behaviors may not have been captured during the brief coding observations. Further, the lack of follow-up data is limiting, as maintenance of teachers’ gains and their effect on student behaviors is unknown.

**Future Directions**

Although limitations in the current study exist, results support the use of TCIT in an urban day-treatment preschool setting to increase teachers’ use of positive behaviors and reduce negative behaviors in the training setting. Results also demonstrate the feasibility of implementing the core elements of PCIT, as previous studies assessed more limited aspects of the original PCIT protocol (i.e., Labeled Praise only, in-classroom coaching only, didactic instruction instead of in-vivo coaching, and absence of coaching-to-mastery criteria). Data also provide support for the use of TCIT in classroom settings to address the gap in services for children with disruptive behaviors and to provide
teachers with effective and efficient behavior interventions. TCIT provides teachers with a single skill set rather than requiring them to integrate and learn new strategies as new problems arise. Although initial resources may be costly, long-term consultation costs may be reduced considerably, as teachers would require fewer behavioral supports over time. Future research should examine the effects of TCIT with larger samples and additional groups of trainers, teachers, and students (e.g., students varying in age, gender, race, SES, and severity of psychopathology). Differences among teachers mandated to training versus those voluntarily involved would provide insight into the effects of motivation on the use of the TCIT skills in the classroom setting. Additional data should compare individual differences within specific students and teachers, as sums of student and teacher data may obscure more subtle individual behavior changes. Assistant teachers should also be trained alongside head teachers to promote consistency and systemic change throughout the classroom and school settings.

Future studies should also develop a manualized protocol to promote dissemination, sustainability, and assessment of training fidelity. Further, feedback from teachers in the current study should be incorporated into the manualized protocol to address their expressed concerns around “knowing what to do, but not how or when to do it.” That is, additional procedures for how to handle specific stressful classroom situations with multiple students should be incorporated into future protocols.

Once a manualized intervention is established as efficacious, further studies should examine the feasibility of disseminating TCIT, the preventive and/or cumulative implications of TCIT with non-identified students, and the possible adjunctive benefit of simultaneously implementing TCIT and PCIT together.
REFERENCES


APPENDIX A

PCIT MANUAL CONTENT: DESCRIPTION OF PRIDE SKILLS AND DON'T BEHAVIORS (Eyeberg et al., 1999)

The Do Rules

"Next we'll talk about what to DO -- the special skills to use during the CDI play sessions. We call these rules the PRIDE skills to help you remember them."

☐ 12. P is for Praise. Give your child Labeled Praises for positive behavior

- Praise compliments a child about his or her behavior
- There are two kinds of praise
  - Labeled praise is specific praise; "You choose such pretty colors!" "You're being so careful with that pen!", "I like it when you build quietly!"
  - Unlabeled praise is nonspecific: "Good!", "That's great!", "Nice job!", "I like that."
- Labeled praise is more effective because it lets the child know exactly what you like.
- Labeled praise increases the behavior that it describes.
- Labeled praise increases child's self-esteem.
- Labeled praise makes both parent and child feel good!

☐ 13. R is for Reflection. Reflect your child's appropriate talk

- Reflection is repeating/paraphrasing what your child says: "Yes, that is a blue crayon."
- Allows child to lead the conversation
- Shows child you're really listening
- Actually helps you learn to listen!
- Shows you accept/understand what child is saying
- Improves and increases child's speech and language
- May feel awkward at first, but becomes natural pretty quickly

☐ 14. I is for Imitation. Imitate your child's appropriate play

- Imitation means doing the same thing your child is doing, such as drawing a tree if your child is drawing a tree
- Helps you keep your attention/comments focused on what your child is doing
- Helps you play right at your child's developmental level
- Lets child lead the play
- Makes the play fun for your child
- Shows your approval of your child's activity
- Teaches child how to play well with others (for example, taking turns).

☐ 15. D is for Description. You describe the positive things your child is doing

- There are different kinds of descriptions
  - Describing the toys or what you are doing can make the play interesting and fun
  - Behavior Descriptions are especially important because they focus specifically on your child's good behavior
  - Behavioral Descriptions state exactly what your child is doing. "You're drawing a sun."
  - Like a sports announcer, a running commentary
  - Lets your child lead and helps you follow his/her play
  - Lets your child know you're interested and paying attention to him/her
  - Lets your child know you approve of what he or she is doing
  - Modulates speech and teaches vocabulary and concepts
  - Holds your child's attention to the task and teaches child how to hold his/her own attention to a task

☐ 16. E is for Enthusiasm. Be enthusiastic during special times with your child

- Enthusiasm means that you act happy and natural when you play with your child
- Enthusiasm includes positive touch, laughter, a tone of voice that expresses your interest
- Lets your child know that you are enjoying your time together
The Don't Rules

8. Avoid Commands
- Commands try to direct the play by suggesting what the child should do.
- There are two kinds of commands:
  - Direct: "Sit down." "Please hand me the car."
  - Indirect: "Would you like to sit down?" "Let's put the cars away."
- Commands take over the lead of the play.
- If the child doesn't obey, the play could stop being fun. CDI is a time when the child is to learn that it's fun to get along and play together nicely.

9. Avoid Questions
- A question asks for an answer from the child.
- There are different kinds of questions:
  - Questions that ask for information: who, what, where, when, how.
  - Unintentional questions: voice goes up at end of sentence; question tags.
  - Questions that are really hidden commands: "Would you like to clean up?"
- Questions take over the lead of the conversation.
- Questions sometimes suggest disapproval.
- Questions often suggest you aren't really listening to your child.

10. Avoid Criticism
- Criticism is a negative or contradictory statement about the child or his/her actions: "You're not nice," or "That doesn't go that way."
- Criticism points out mistakes rather than providing correction.
  - To correct without criticizing, could say, "It goes this way."
- Criticism tells the child what NOT to do: "Stop that." "Don't do that!" "Quit it."
- Criticism lowers a child's self-esteem.
- Criticism creates a negative interaction.
Appendix B

Teacher-Child Interaction Training Effective Command Sequence Protocol

(Madigan and Kurtz, In Preparation)