A Randomized Controlled Trial Comparing PCIT-Toddler, Circle of Security, and Waitlist Controls to Improve Child and Caregiver Emotion Regulation

Robin Choi Han
West Virginia University, rc0115@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Child Psychology Commons

Recommended Citation
Han, Robin Choi, "A Randomized Controlled Trial Comparing PCIT-Toddler, Circle of Security, and Waitlist Controls to Improve Child and Caregiver Emotion Regulation" (2023). Graduate Theses, Dissertations, and Problem Reports. 12020.
https://researchrepository.wvu.edu/etd/12020

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
RCT COMPARING PCIT-T AND COS-P

A Randomized Controlled Trial Comparing PCIT-Toddler, Circle of Security, and Waitlist Controls to Improve Child and Caregiver Emotion Regulation

Robin C. Han, M.S.Ed., M.S.

Dissertation submitted
to the Eberly College of Arts and Sciences
at West Virginia University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in
Psychology

Cheryl McNeil, Ph.D., Chair
Amy Gentzler, Ph.D.
Elisa Krackow, Ph.D.
Carrie Rishel, Ph.D.

Department of Psychology
Morgantown, West Virginia
2022

Keywords: Parent-Child Interaction Therapy, Circle of Security, toddlers, emotion regulation

Copyright 2022 Robin Han
ABSTRACT

A Randomized Controlled Trial Comparing PCIT-Toddler, Circle of Security, and Waitlist Controls to Improve Child and Caregiver Emotion Regulation

Robin Han

Emotion regulation is an important developmental task during toddlerhood that is associated with positive psychosocial outcomes (Zeman et al., 2006). The development of adaptive emotion regulation during early childhood occurs largely within the context of a supportive caregiver-child relationship (Morris et al., 2007). Thus, parent-mediated interventions are a promising medium through which emotion regulation problems in toddlers can be treated. However, few interventions specifically designed to treat behavioral and emotion regulation difficulties are available for children in the toddler age range, and these interventions have yet to establish a solid evidence base supporting their efficacy. To fill this gap in the literature, the current study investigated the efficacy of two parenting interventions designed to improve emotion regulation in toddlers: Parent-Child Interaction Therapy-Toddler (PCIT-T; Girard et al., 2018) and Circle of Security-Parenting (COS-P; Cooper et al., 2009). Using a randomized controlled trial design, 76 parent-child dyads were randomly assigned to PCIT-T, COS-P, or waitlist control groups. Of the 76 dyads, 51 completed both Time 1 and Time 2 assessments. Observational coding and parent-report questionnaires were used to measure child and parent emotion regulation and related constructs at pre-treatment and post-treatment. Two-way mixed ANOVA were conducted to examine effects of group, time, and group-by-time interaction on parent and child emotion regulation. No significant interaction effects were found for any of the analyses. However, significant and large main effects of time were found for observed parent negative talk, observed maternal support-seeking, parent-reported child dysregulation, and parent-reported child externalizing behavior. Results are discussed in light of study limitations and future directions.
Acknowledgments

They say it takes a village to raise a child. Then what does it take to raise a child and complete a dissertation? I know firsthand that it takes a highly supportive and dedicated group of villagers: Thank you, Dr. Cheryl McNeil, for the gift of your lifelong mentorship and for nurturing me these past three years—not just as a scientist, clinician, and student but also as a wife and mother. This project would not have been possible without Dr. Jane Kohlhoff, Sara Cibralic, and the staff and clinicians at the Karitane Toddler Clinic; I am so grateful for your collaboration and partnership. I would like to thank my committee members, Drs. Elisa Krackow, Amy Gentzler, and Carrie Rishel, for their invaluable feedback during the development of this study. Thank you to my lab siblings—Chris, Lindsay, Sharon, and Erinn—and the team of undergraduate research assistants in the WVU Parent-Child Interaction Lab for their tireless work on this project and for their moral support.

I could not have made it through my long and circuitous graduate school journey without the incredible sacrifice of my husband and parents. Andrew, thank you for uprooting your life and moving to West Virginia so that I could learn from the best. Thank you for doing the heavy-lifting around the house and as a father so that I could focus on school and maintain some semblance of work-life balance. I am eternally grateful to my mom, who cared for Julian during the height of the pandemic and flew halfway across the world to help care for him during the completion of this document. Thank you to my dad, who always encouraged me to dream big and strive toward my goals, no matter how lofty. Lastly, I would like to thank Julian for bringing me immense joy each and every day and inspiring me to be a better parent and clinician. Thank you all for being my village and for getting me to the finish line.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iv</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>28</td>
</tr>
<tr>
<td>Study Design</td>
<td>28</td>
</tr>
<tr>
<td>Participants</td>
<td>29</td>
</tr>
<tr>
<td>Procedure</td>
<td>31</td>
</tr>
<tr>
<td>Measures</td>
<td>35</td>
</tr>
<tr>
<td>Data Analysis Plan and Hypotheses</td>
<td>44</td>
</tr>
<tr>
<td>Results</td>
<td>47</td>
</tr>
<tr>
<td>Preliminary Analyses</td>
<td>47</td>
</tr>
<tr>
<td>Main Analyses</td>
<td>50</td>
</tr>
<tr>
<td>Attrition Analyses</td>
<td>61</td>
</tr>
<tr>
<td>Discussion</td>
<td>62</td>
</tr>
<tr>
<td>References</td>
<td>80</td>
</tr>
<tr>
<td>Appendix</td>
<td>107</td>
</tr>
</tbody>
</table>
The toddler years, defined as the time between a child’s first and third birthday, are a critical period in the emotional development of a child. During this stage, toddlers make the transition from predominantly relying on caregivers to manage their emotions to increasingly relying on their own emotion regulation skills (Gross & Thompson, 2007). Caregivers play an integral role in scaffolding effective regulatory strategies for toddlers and helping them navigate this shift from interpersonal to intrapersonal emotion regulation (Zaki & Williams, 2013). Given the influence of caregiver characteristics and behaviors on children’s emotion regulation (Morris et al., 2007), parent-mediated interventions may be an effective way to address emotion regulation difficulties in toddlerhood. Unfortunately, parenting interventions targeting emotion dysregulation in toddlers are scarce in number (England-Mason & Gonzalez, 2020), and further study is needed to establish their effectiveness. The current study sought to contribute to the limited research on this topic by examining the efficacy of a novel parenting interventions, Parent-Child Interaction Therapy for Toddlers (PCIT-T), on the emotion regulation of toddlers identified as having behavioral difficulties and their primary caregivers.

Defining Emotion Regulation

Several definitions of emotion regulation have been proposed in the developmental literature. Broadly, emotion regulation has been defined as the processes that influence “which emotions we have, when we have them, and how we experience or express them” (Gross, 1998; Gross et al., 2011, p. 767). More specifically, Gross (2014) delineates three core features of emotion regulation: the goal, the strategy, and the outcome. Although emotion regulation is commonly perceived as the act of downregulating negative emotions, the goal of emotion regulation is...
regulation can be to increase or decrease the intensity or duration of either positive or negative affective responses (Gross et al., 2006). Once a regulatory goal has been activated, there are a number of emotion regulation strategies that an individual can employ to achieve this goal: situation selection, situation modification, attentional deployment, cognitive change, and response modulation (Gross, 1998). These strategies vary in the point of the regulatory process at which the emotion is targeted, with situation selection occurring at the earliest phase of the process and response modulation occurring after an emotional response has already been initiated (Gross, 2014). The use of these strategies can have various outcomes on one or more aspects of emotion (e.g., changes in emotion latency, intensity, duration) that can manifest affectively, behaviorally, cognitively, and socially (Gross, 2014).

Another widely accepted definition of emotion regulation is the monitoring, evaluation, and modification of emotional responses to accomplish one’s goals, which can involve changes to the valence, intensity, or duration of the activated emotion (Thompson, 1990; 1994). Eisenberg and colleagues (2000; 2007) expanded this definition of emotion regulation to include both physiological processes and behavioral expressions of emotions in addition to emotion-related internal states. To add to the complexity of these definitions, some researchers have differentiated between dimensions or processes of emotion regulation: intrinsic versus extrinsic (Gross, 2015), interpersonal versus intrapersonal (Zaki & Williams, 2013), automatic versus controlled (Gross & Thompson, 2007), and conscious versus unconscious (Gyurak et al., 2011). For the purpose and scope of the current study, however, the broader definition proposed by Gross et al. (2011; 2014) was used, thereby conceptualizing emotion regulation as a dynamic process by which an individual manages, modulates, or changes the experience or expression of an activated emotion in service of a goal.
Due to the multifaceted nature of the construct, the measurement of emotion regulation can present researchers with some challenges. Because a gold-standard, global assessment of emotion regulation does not exist, measurement of emotion regulation requires a careful selection of tools that evaluate the particular aspect of interest. Since the umbrella term “emotion regulation” can be used to refer to different strategies, forms, or processes, Rottenberg and Gross (2007) cautioned against using emotion regulation measures without close examination of the specific construct being assessed and recommend using differentiated assessment of emotion regulation. Taking these points into consideration, the current study employed a multi-method approach to identify and assess emotion regulation and related constructs in both toddlers and their caregivers.

**Emotion Regulation in Young Children**

As the bridge between infancy and the preschool years, toddlerhood serves as a critical time for children to develop the ability to regulate their own emotions. Thus, effective emotion regulation is an important developmental goal that impacts children’s psychosocial functioning (Zeman et al., 2006). Adaptive emotion regulation in young children has been linked to a myriad of positive social and cognitive outcomes, including social competence (Cohen & Mendez, 2009; Monopoli & Kingston, 2012), sustained attention (Graziano et al., 2011), and school readiness (Brophy-Herb et al., 2013). In contrast, emotion regulation difficulties in early childhood have been associated with externalizing problems (Halligan et al., 2013; Sidor et al., 2017), poor social skills (Spritz et al., 2010), and development of psychopathology (Keenan, 2000). Thus, poor emotion regulation in toddlerhood is a worthy target of intervention given its negative impact on other developmental domains.
The early development of emotion regulation is best examined in the context of the family unit. Morris and colleagues (2007) proposed a tripartite model of contextual factors that influence children’s emotion regulation development and adjustment. According to this model, (1) parent modeling of emotion regulation, (2) parenting practices related to emotion management, and (3) the family emotional climate (e.g., parenting style, attachment relationship, family expressiveness) collectively shape a child’s emotion regulation development and adjustment. Previous studies provide support for familial influences on emotion regulation development. Children were more likely to use adaptive emotion regulation behaviors if their parents used supportive emotion parenting (i.e., parenting behaviors that showed awareness and validation of their child’s emotional experiences) than unsupportive emotion parenting (Morelen & Suveg, 2012). When mothers engaged their children in adaptive emotion regulation strategies (e.g., attentional deployment, cognitive change) during tasks designed to elicit negative affect, children displayed less anger and sadness (Morris et al., 2011). Parent coaching of emotions was linked to more successful emotion regulation in children than those whose parents did not use emotion coaching (Morris et al., 2017). These studies highlight the importance of the caregiver and family context in the development of adaptive emotion regulation.

One approach to studying emotion regulation in young children is to examine the behavioral strategies used during emotion-inducing situations. According to Kopp (1989), emotion regulation strategies become more socially interactive during the transition from infancy to toddlerhood, as children cue and recruit their caregivers to assist in their emotion regulation. For example, children use less oral self-soothing at 13.5 months than at 3 months of age (Rothbart et al., 1992) and more directing, information seeking, and social referencing at 18 months than at 12 months of age (Parritz, 1996). Research suggests that the use of emotion
regulation strategies employed may change as a function of context and caregiver availability (Diener & Mangelsdorf, 1999). In a study of early childhood emotion regulation strategies, toddlers used more behavioral strategies during situations that elicited positive emotions or frustration than in fear-inducing situations (Roque & Veríssimo, 2011). Caregiver availability and involvement influenced the types of strategies used, with more distraction used during a delayed gratification task (Grolnick et al., 1996) and less fussing and passive disengagement used during a frustration task (Roque & Veríssimo, 2011) when mothers were engaged versus disengaged. As primary agents of emotion socialization in childhood, caregivers facilitate the development of adaptive emotion regulation and effective emotion regulation strategies in their children. Therefore, parent-mediated interventions may be ideal for targeting emotion regulation problems in toddlerhood.

**Self-Regulation**

Emotion regulation is often placed under the broader framework of self-regulation (Holodynski et al., 2013). Self-regulation is a concept that encompasses emotion regulation and refers to the process of managing and modulating affect, behavior, and cognition to cope with the demands of the environment (Calkins & Williford, 2009; Posner & Rothbart, 2000). The development of self-regulation in toddlerhood is tied to maturing attentional processes and inhibitory control characteristic of this stage (Eisenberg et al., 2004; Kochanska et al., 2001). Additionally, parenting behavior is a predictor of self-regulation (Bradley, 2000; Calkins & Fox, 2002). Problems with self-regulation, or dysregulation, in childhood is associated with maladjustment and psychosocial impairment in later life (Ayer et al., 2009; Meyer et al., 2009). Calkins (2007) posited that failure to meet these early regulatory processes have trickle-down effects on a child’s adaptive and social-emotional skills. Because affective, behavioral, and
RCT COMPARING PCIT-T AND COS-P

cognitive regulation are interconnected, there is evidence that difficulties in one domain
correspond with impairments in other domains (Calkins & Fox, 2002) and that improvements
observed in one area of self-regulation are typically observed in the other areas (Tottenham et al.,
2011). For this reason, Eisenberg and colleagues (2017) recommended broadening the types of
measures used when studying emotion regulation to include assessments of constructs that
contribute to self-regulation (e.g., attentional control, inhibitory control). For the current study,
this recommendation was taken into account when selecting measures to assess child emotion
regulation.

*Externalizing Behavior*

Numerous studies have demonstrated an association between poor emotion regulation in
young children and externalizing symptoms (Campbell et al., 2000; Cole et al., 1996; Eisenberg
et al., 2001; Zeman et al., 2002). Additionally, toddlers’ use of ineffective emotion regulation
strategies has been linked to higher externalizing behavior problems (Calkins & Dedmon, 2000).
In a longitudinal study of toddler externalizing behavior, poor emotion regulation and inattention
at age 2 years predicted clinically elevated externalizing behavior at ages 4 and 5 years in girls
but only inattention was a predictor for boys (Hill et al., 2006), highlighting potential gender
differences in the association between emotion regulation and disruptive behavior. When
parenting behaviors are incorporated into analyses, emotion regulation appears to mediate the
association between parenting and disruptive behavior in toddlers (Eisenberg et al., 2005;
Spinrad et al., 2007). Relatedly, emotion regulation has been found to mediate the relation
between harsh parenting practices and child aggressive behavior in the school environment
(Chang et al., 2003). Thus, there is consensus in the literature that problems with emotion
RCT COMPARING PCIT-T AND COS-P

regulation (and, more broadly, self-regulation) contribute to toddlers’ externalizing behaviors and mediate the relation between parenting practices and externalizing problems.

**Caregiver Emotion Regulation and Related Factors**

Given the central role that caregivers play in the emotional development of toddlers, there is great value in examining caregiver factors that may directly or indirectly contribute to children’s ability to regulate their emotions. In a comprehensive review of the literature pertaining to caregiver and child emotional functioning, Bariola and colleagues (2011) outlined three key caregiver factors that negatively impact children’s emotion regulation: (1) poor caregiver emotion regulation, (2) parental unsupportive responses to children’s displays of emotions, and (3) harsh parenting styles. Thus, facets of these three factors were included in the current investigation in order to examine their influences on the emotion regulation of toddlers with disruptive behaviors.

**Caregiver Emotion Regulation**

In order for caregivers to serve as primary emotion socialization agents and model effective emotion regulation for their children, it stands to reason that they must possess adequate emotion regulation abilities themselves. However, caregiver emotion regulation as it relates to child emotion regulation has been understudied in the literature, especially among the toddler age group, with the majority of research focusing on other caregiver factors (e.g., parenting styles, parenting behaviors, caregiver emotion expressivity). In a longitudinal study of 9- to 13-year-old boys, parental emotion dysregulation was predictive of child emotion dysregulation 6 years later (Kim et al., 2009). Bariola et al. (2012) found that maternal emotion suppression, an emotion regulation strategy, was related to child emotion suppression among a sample of 9- to 19-year-old children, though this was not the case for paternal use of this
strategy. Another study by Quetsch and colleagues (2018) found child emotion regulation to be a mediator for the relation between parent emotion regulation and child externalizing behaviors. Further study is warranted to better understand the connection between toddler and caregiver emotion regulation and the processes by which they are related.

**Caregiver Emotion Socialization**

Parental responsiveness to children’s displays of emotions is a method of emotion socialization that falls into two categories: supportive or unsupportive (Morelen et al., 2014). Responses that are supportive of children’s emotional experiences are those that validate expressed emotions, aid in the child’s understanding of these emotions, and model appropriate emotional reactions for the child (Garner et al., 2008). On the other hand, unsupportive responses seek to minimize, punish, or criticize the child’s emotions (Fabes et al., 2002). Supportive parental responses to children’s emotions have been linked to adaptive emotion regulation in middle childhood (McDowell et al., 2002), whereas unsupportive parental responses to child emotions have been linked to child emotion regulation difficulties (Eisenberg et al., 1999). Emotion dismissal, which falls into the category of unsupportive parental response to children’s emotions, has been associated with poor emotion regulation and behavioral difficulties (Lunkenheimer et al., 2007), as well as poor social competence and heightened anger responses (Snyder et al., 2003). These findings suggest that interventions that increase supportive emotion-related parenting behaviors and simultaneously decrease unsupportive behaviors may be successful in improving emotion regulation difficulties in young children.

**Harsh Parenting Practices**

Harsh parenting practices are those that are negative, controlling, or punishing in nature (Fabes et al., 2001) and include behaviors such as yelling, criticism, negative commands, and
physical aggression (Chang et al., 2003). In the literature, harsh parenting is often conceptualized as an indicator of poor caregiver emotion regulation (Eisenberg et al., 1999) or as a form of parental socialization of negative emotions (Gottman et al., 1997). Thus, observations of negative, controlling, or dismissive parenting behaviors may provide insight into caregiver emotion regulation above and beyond what is garnered through self-report measures. Harsh parenting practices have long been linked to unfavorable child outcomes, including reduced capacity for emotion regulation. For example, caregivers who used more controlling behaviors with their toddlers during emotion-provoking tasks had toddlers who were more emotionally dysregulated (Calkins et al., 1998) and distressed when met with frustration (Calkins & Johnson, 1998). Additionally, observed intrusive parenting (marked by overly directive and controlling parenting behaviors) at age 2 years was associated with more behavior problems and less emotion regulation at age 3 years (Diemer et al., 2021). There is evidence to suggest that this association is bidirectional in nature (Patterson et al., 1982; Snyder & Stoolmiller, 2002), with child disruptive behaviors also contributing to the use of harsh parenting practices.

**Interventions Targeting Emotion Regulation in Toddlers**

Although there is an extensive evidence base for the use of parenting interventions to treat child externalizing behavior (Eyberg et al., 2008), little research has been conducted on the use of parenting interventions to improve child emotion regulation. England-Mason and Gonzalez (2020) conducted a review of the literature on emotion socialization parenting programs for young children between birth to six years of age and identified three interventions: Tuning in to Kids (TIK; Havighurst et al., 2009), Parent-Child Interaction Therapy-Emotion Development (PCIT-ED; Lenze et al., 2011), and Emotion-Enhanced Triple P (EETP; Salmon et al., 2014). Despite differences in theoretical foundation, the three programs consisted of similar
treatment components (e.g., psychoeducation on children’s emotional development, training caregivers in emotion coaching). Of the articles included in this review, only half (n = 6) included a measure of child emotion regulation, and upon closer examination, the majority of these assessed a related construct (e.g., emotional competence, emotion expression). Further, only one program (Tuning in to Toddlers, an adaptation of TIK) was designed specifically for toddlers; a pilot study of TOTS demonstrated increased emotion coaching and decreased emotion dismissing behaviors in caregivers and reduced behavior problems in children (Lauw et al., 2014). Interventions targeting the toddler age group are worth developing and studying, as they may provide insight into tackling emotion regulation problems from a preventive framework (Morawska et al., 2019).

One attachment-based parenting program designed for caregivers of infants and toddlers is the Attachment and Biobehavioral Catch-up (ABC; Dozier et al., 2006). The intervention is delivered by parent coaches who provide live coaching to caregivers in the family’s home. ABC has a strong evidence base for improving a variety of child outcomes, including internalizing problems, externalizing problems, and emotion regulation via saliva sampling and cortisol testing (Grube & Liming, 2018). An adaptation for toddlers aged 14-71 months has been developed (ABC-Toddler; Lind et al., 2017), and preliminary research suggests that this adaptation may be effective in improving toddlers’ executive functioning during a set-shifting task (Lind et al., 2017) and caregiver responsivity (Imrisek et al., 2018). Due to the setting of service delivery, however, this home-based intervention may not always be feasible for outpatient clinics to administer.

The current study focused on a relatively novel intervention, Parent-Child Interaction Therapy-Toddler (PCIT-T; Girard et al., 2018) and compared it to a widely used parenting
RCT COMPARING PCIT-T AND COS-P

program, Circle of Security-Parenting (COS-P; Cooper et al., 2009), both of which are appropriate for the toddler age range and can be administered in an outpatient setting. A detailed description of these interventions as well as relevant research on outcomes related to child and caregiver emotion regulation are provided.

**Parent-Child Interaction Therapy-Toddler**

PCIT-Toddler (Girard et al., 2018) is an adaptation of Parent-Child Interaction Therapy (PCIT) for children between the ages of 12-24 months. PCIT is a behavioral parent training program that was originally developed to treat disruptive behavior disorders in children ages 2-7 years (McNeil & Hembree-Kigin, 2010). Research on the efficacy and effectiveness of PCIT for a variety of clinical populations spans more than four decades and has established a strong evidence base (see Lieneman et al., 2019, for a comprehensive review). Unlike many other parenting interventions, PCIT involves both the caregiver and the child and utilizes in vivo coaching by a therapist from behind a one-way mirror. PCIT is built on the premise that challenging behaviors in preschool-aged children are often attention-seeking in function or are products of a coercive parent-child cycle (Patterson, 2002). Thus, treatment is focused on first improving the parent-child relationship and then teaching parents how to respond predictably and consistently to child noncompliance and negative attention-seeking behavior.

Although many of the basic treatment components of PCIT are maintained in PCIT-T, a number of changes have been made to meet the unique emotional and physical needs of toddlers (Girard et al., 2018). In PCIT-T, disruptive behaviors in toddlerhood are conceptualized differently in that they are not considered to be attention-seeking in nature but rather behavioral manifestations of emotional dysregulation. Further, parent-child attachment is emphasized in PCIT-T, as caregivers become more attuned to their child’s needs and learn how to respond with
RCT COMPARING PCIT-T AND COS-P

sensitivity and nurturance. Although withdrawal of parental attention and discipline procedures are included in this model, they are adapted to be more developmentally appropriate for the toddler age range and used more sparingly than in standard PCIT. Thus, the focus of the adapted treatment is to improve caregiver-child attachment by increasing parental responsiveness to the child’s needs and their ability to support their child when they are emotionally dysregulated.

**Child-Directed Interaction-Toddler.** There are two treatment phases in PCIT-T: (1) Child-Directed Interaction-Toddler (CDI-T) and (2) Parent-Directed Interaction-Toddler (PDI-T). In CDI-T, caregivers are instructed to follow their toddler’s lead and use positive parenting skills (i.e., “PRIDE” skills; praises, reflections, imitation, behavior descriptions, and enjoyment) during a dyadic free play interaction. Caregivers are taught to use labeled praises (e.g., “Thank you for using a quiet voice”) rather than unlabeled praises (e.g., “Thank you”) to indicate positive evaluation of a specific action or attribute of the child. They are also instructed to “reflect” (or restate) their toddler’s words or word approximations and describe their toddler’s appropriate behaviors. Additionally, caregivers are encouraged to imitate their child’s appropriate play and express enjoyment of the dyadic interaction. Caregivers are taught to avoid using questions, commands, and criticism during the free-play session, as these take the lead away from the child and may add negativity to the interaction. In order to progress to the next phase of treatment (i.e., PDI-T), caregivers must reach the goal criteria of 10 labeled praises, 10 reflections, and 10 behavioral descriptions with fewer than 4 questions, commands, or criticisms in a 5-minute play session.

In addition to the PRIDE skills, caregivers are coached to use the CARES model as a way to help toddlers regulate their emotions. The acronym “CARES” spells out a set of prescribed behaviors used in response to toddler emotion dysregulation: *come in, assist child, reassure child,*
emotional validation, and soothe. Specifically, parents are encouraged to draw physically close to the child, provide scaffolded guidance to support the child, give verbal reassurance to the child (e.g., “It’s okay. Dad is here.”), label the child’s emotions, and use a soothing tone of voice and physical touch to comfort the child. Though there are no goal criteria associated with the CARES steps, caregivers are coached by the therapist to use these skills flexibly whenever their child is experiencing emotional distress.

As caregivers help their children regulate using CARES, PCIT-T therapists follow a parallel process and guide the caregiver through a similar set of steps to help regulate their own emotions. The CARES steps for caregivers are as follows: check cognitions and clue into oneself; assist self; reassure self; become emotionally aware; and use sensitive and soothing tone. Caregivers may experience intense emotions while engaging with their toddler, especially in response to their child’s misbehavior or emotion dysregulation. In such instances, therapists assist caregivers in using a variety of emotion regulation strategies: recognizing their own cognitions, emotions, and corresponding physiological responses; using relaxation techniques (e.g., deep breathing, progressive muscle relaxation); engaging in positive self-talk; validating and accepting their emotions; and speaking to themselves in a calm, soothing voice. The premise of the caregiver CARES steps is that caregiver emotion regulation is integral to maintaining a positive caregiver-child relationship and promoting successful child emotion regulation.

In response to misbehavior during the dyadic play interaction, caregivers are instructed to use “underreaction and redirection” (Girard et al., 2018, p. 25), that is, to refrain from acknowledging the child’s misbehavior and instead describe the caregiver’s own play. Caregivers use a labeled praise for a positive opposite of the misbehavior as a way to further encourage their child to engage in appropriate behavior. For child behavior that is destructive or
dangerous in nature (e.g., hitting, biting, throwing a toy), caregivers are taught to use a brief, developmentally appropriate discipline procedure. Caregivers are coached to hold the child’s hands, say “No hurting,” in a firm voice, and look away from the child for three seconds. Then the caregiver lets go of the child’s hands, says “No hurting. Gentle hands,” and physically moves the child toward another toy. Therapists then coach caregivers to use PRIDE skills and CARES steps as needed to redirect or help calm the child.

**Parent-Directed Interaction-Toddler.** In the second phase of treatment, PDI-T, caregivers are taught to use simple, direct commands that are developmentally appropriate for the child, and children are taught to listen using a guided compliance technique referred to as the “Tell-Show-Try Again-Guide” procedure. This procedure is based on the presumption that compliance for toddlers is an emerging developmental skill that must be taught with a positive and constructive approach. In the Tell-Show-Try Again-Guide procedure, caregivers issue a simple command (e.g., “Please hand me the car.”) paired with a physical prompt (e.g., pointing to the toy car). The caregiver gives the child five seconds to complete the instruction, remaining quiet yet repeating the physical prompt as many times as needed. If the child completes the command at this stage, the caregiver praises the child and provides a warm, positive touch. If the child does not complete the command at this initial step, the caregiver moves onto the “Show” step, in which the verbal command is repeated followed by the words “like this” (e.g., “Please hand me the car like this.”). The caregiver demonstrates the completion of the command and then puts the object back in its original location for the child to have another try. Again, compliance is met with enthusiastic praise and positive touch. If the child does not complete the command within five seconds, the caregiver moves onto the “Try Again” step, in which the caregiver says, “Your turn” before repeating the command with physical prompts. If the child
successfully completes the command within five seconds, the child receives a praise paired with a positive touch. If not, the caregiver moves onto the final “Guide” step in which the caregiver says, “Mommy will help you (original command),” and places their hand over their toddler’s hand to complete the command. The caregiver then concludes the sequence by saying, “That’s (original command).”

Caregivers reach PDI-T criteria when they are able to demonstrate competency in issuing commands, 75% of which must be effective (e.g., direct, simply worded, developmentally appropriate) and followed through correctly (i.e., enthusiastic praise and positive touch if the child completes the command, progression through the Tell-Show-Try Again-Guide sequence if the child does not). For further details regarding theoretical rationale and treatment components, please refer to the PCIT-Toddler treatment protocol (Girard et al., 2018).

**PCIT-T as an Emotion Regulation Intervention.** In Girard et al. (2018), the authors make a compelling case for the conceptualization of PCIT-T as an emotion regulation treatment for toddlers. PCIT-T contains a number of treatment components that directly target both child and caregiver emotion regulation (e.g., therapist coaching of caregiver through the CARES steps, caregiver modeling of emotion regulation skills, caregiver use of CARES in response to toddler distress). Because of the relative novelty of PCIT-T, research on the effects of this intervention on child and caregiver emotion regulation is lacking. In fact, only a handful of outcome studies have been published on PCIT-T to date, with the majority of these studies examining the first phase of treatment (CDI-T). In a pilot study of PCIT-T, Kohlhoff and Morgan (2014) found significant improvements with large effect sizes ($d = 1.02-1.03$) in the frequency and problematic nature of child disruptive behavior and in maternal depression for 28 caregiver-toddler dyads. Using a randomized controlled trial design with 66 mother-toddler dyads, Kohlhoff et al. (2021)
RCT COMPARING PCIT-T AND COS-P

found significantly greater improvements in parent-reported maternal sensitivity and non-intrusiveness and greater reductions in child internalizing and externalizing behaviors for families in the CDI-T group compared to those in the waitlist control group. These gains were maintained at a 4-month follow-up and were medium to large in effect size (Kohlhoff et al., 2021). The entire model (i.e., both CDI-T and PDI-T phases) was examined in a case study of a toddler with autism spectrum disorder (Cibralic et al., 2021); overall improvements in parent-reported child externalizing behavior and self-regulation were observed from pre-treatment to the 6-month follow-up after completion of PCIT-T. Despite these improvements, the child remained in the borderline clinical range and “area of need” range even after treatment. Thus, further research employing robust methodologies (e.g., use of control group, multiple baseline) and direct measurement of emotion regulation is needed to establish it as an evidence-based treatment for toddlers with emotional regulation difficulties.

**Standard PCIT as an Emotion Regulation Intervention.** Despite the dearth of research examining the efficacy of PCIT-T in improving emotion regulation, there is evidence to support the use of standard PCIT in treating emotion regulation problems. Due to the considerable overlap in treatment components between the two models, the evidence surrounding standard PCIT as an emotion regulation intervention suggests that PCIT-T has the potential to be similarly efficacious. The literature on standard PCIT and variations of the standard model are reviewed, with particular attention given to treatment effects on child and caregiver emotion regulation and related constructs (e.g., internalizing symptoms, externalizing behavior).

In a study of 86 children referred to a university-based clinic for PCIT, Rothenberg et al. (2019) found significant pre-post improvements in child disruptive behavior as reported by parents on the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999) Intensity scale
RCT COMPARING PCIT-T AND COS-P

and the Behavior Assessment System for Children, 2nd Edition (BASC-2; Reynolds & Kamphaus, 2004) Externalizing Problems scale. Significant reductions were found in child emotion dysregulation on the BASC-2 Anger Control and Emotional Control scales and on the Behavior Rating Inventory of Executive Function (BRIEF; Gioia et al., 2000) Emotional Control scale upon completion of PCIT. These improvements in child behavior uniquely predicted emotional dysregulation at post-treatment even after controlling for baseline emotion dysregulation. Further, pre-post child behavior improvements mediated the association between pre- and post-treatment emotion dysregulation, such that greater pre-treatment emotion dysregulation predicted more behavioral improvement thereby predicting lower post-treatment emotion dysregulation scores. In this model, the robust behavioral improvements typically observed after PCIT are conceptualized as a mechanism through which changes in child emotion regulation occur over the course of treatment. Of note, positive parenting skills, considered an active ingredient of PCIT, did not mediate the association between pre- and post-treatment emotion regulation.

As part of a larger study examining the effects of incentives on treatment outcomes in PCIT, Lieneman and colleagues (2020) found significant improvements in child lability/negativity ($d = 1.93$) as reported by parents on the Emotion Regulation Checklist (ERC; Shields & Cicchetti, 1997) across the course of treatment for a group of 66 caregivers. As expected, child disruptive behavior per caregiver report on the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999) also improved significantly with a huge effect size ($d = 2.10$; Sawilowsky, 2009). Although further research is needed, these studies suggest that standard PCIT may be considered a treatment for child emotion dysregulation in addition to child disruptive behaviors.
In the same study, Lieneman et al. (2020) also demonstrated improvements in caregiver emotion regulation with a medium effect size \( (d = 0.78) \) as reported on the Difficulties in Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). Post-hoc analyses of DERS scores indicated significant improvements in caregiver emotion regulation from pre- to mid-treatment (i.e., successful completion of CDI criteria) and mid- to post-treatment (i.e., successful completion of PDI criteria). Further, no significant differences were found between pre-treatment caregiver or child emotion regulation and attrition from treatment, which suggests that baseline emotion regulation difficulties had little to do with treatment completion. Similarly, Zimmer-Gembeck and colleagues (2019) found significant improvements in caregiver emotion regulation on the DERS with a small effect size \( (d = 0.24) \) and greater use of cognitive reappraisal as an emotion regulation strategy with a small effect size \( (d = -0.39) \) among 139 caregivers receiving PCIT services in Australia. Findings were similar, though with slightly larger effect sizes, when only treatment completers \( (n = 90) \) were examined. Child externalizing behavior on the BASC-2 improved significantly with medium effect sizes (i.e., \( d = .59 \) for treatment completers and \( d = .57 \) for the total sample). Additionally, this study found that greater improvement in caregivers’ use of cognitive appraisal was significantly associated with a greater reduction in child externalizing symptoms.

Both Lieneman et al. (2020) and Zimmer-Gembeck et al. (2019) posit that specific features of PCIT contribute to the pre- to post-treatment improvements in child and caregiver emotion regulation. In CDI, caregivers are coached to ignore their child’s negative attention-seeking behavior and redirect the child to another toy or activity. In PDI, caregivers are instructed to respond to child noncompliance with a structured discipline procedure and a script recited in a calm, neutral tone of voice. Thus, caregivers are required to regulate their emotions
in PCIT rather than respond negatively to their child’s emotional dysregulation and continue the coercive caregiver-child cycle. Although these two studies provide preliminary support that standard PCIT may be effective in improving caregiver emotion regulation, neither included a control group to rule out other plausible causes unrelated to the treatment (e.g., maturation, statistical regression). Further study is needed to establish stronger causal inferences and to better understand the ways in which child and caregiver emotion regulation relate to one another in the context of PCIT. Additionally, use of self-report measures alone may be problematic in that caregiver report of their own or their child’s emotion regulation abilities may be subject to bias; thus, the use of multi-informant report or other methods of data collection is needed to bolster these findings.

A number of studies on PCIT have incorporated behavioral (Rodriguez et al., 2014) and physiological measures (Bagner et al., 2012; Graziano et al., 2012) to investigate changes in child emotion regulation. In a randomized controlled trial of 28 mothers and their children born preterm with disruptive behavior problems (Bagner et al., 2012), baseline respiratory sinus arrhythmia (RSA), a physiological correlate of emotion regulation difficulties, moderated the effect of PCIT on child disruptive behavior. Specifically, children with lower baseline RSA (often associated with lower capacity for self-regulation) saw greater improvements in parent-reported disruptive behavior on the ECBI Intensity scale after completion of PCIT than did children with higher baseline RSA. Another study (Rodriguez et al., 2014) of the same sample replicated these findings by using a behavioral coding scheme during free play and clean-up tasks to measure child emotion regulation. Researchers coded children’s abilities to regulate their emotional responses; a code of 0 indicated dysregulation or lack of control over their emotional distress, and a code of 4 represented good regulatory abilities. Children who had poorer observed
emotion regulation at baseline had greater improvements in parent-reported disruptive behavior than did children with better observed emotion regulation at baseline. These results suggest that the degree to which child disruptive behavior improves after PCIT may vary depending on children’s capacity for emotion regulation at pre-treatment, though the factors contributing to these differential effects have yet to be studied. Graziano and colleagues (2012) examined changes in child RSA suppression (calculated by subtracting RSA collected during a clean-up task from resting RSA) and its relation to parent skill acquisition over the course of PCIT. Meaningful pre-post improvements in RSA suppression with a medium effect size ($d = .68$) were found for 45% of the sample, and maternal improvements in positive parenting skills predicted improvements in RSA suppression. This finding was inconsistent with the results from Rothenberg et al., (2019), which did not find caregiver use of positive parenting skills to significantly account for the association between pre- and post-treatment child emotion regulation.

Several studies have adapted PCIT to more specifically address emotion dysregulation in both children and caregivers. PCIT Emotion Development (PCIT-ED; Lenze et al., 2011) is an adaptation of PCIT that targets depression in preschool-aged children. In this adaptation, caregiver-child dyads receive 6 sessions each of CDI and PDI followed by an 8-session Emotion Development (ED) module focused on teaching caregivers how to recognize, label, and respond appropriately to their child’s negative emotions using live coaching through dyadic interactions involving emotion-inducing stressors. In a randomized controlled trial comparing a PCIT-ED group ($n = 100$) to a waitlist control group ($n = 91$), PCIT-ED was effective in significantly reducing depression symptom severity and functional impairment as well as improving emotional functioning for those in the treatment group (Luby et al., 2018). These findings were
consistent with results from a previous pilot study of PCIT-ED (Lenze et al., 2011). Further
study of PCIT-ED demonstrated significantly greater reduction in depression severity during the
CDI and ED modules for children in the treatment versus the control group (Luby et al., 2020).
Analysis of EEG recording during a guessing task (in which children could either win or lose
points based on their guesses) showed a significant increase in neural reward response to wins
only after completion of the ED module compared to waitlist controls (Luby et al., 2020). PCIT-
ED has also been found to improve child emotion lability/negativity (as reported by both parents
and teachers on the Emotion Regulation Checklist) for children with a diagnosis of ADHD
(Chronis-Tuscano et al., 2016). Additionally, there is preliminary evidence to support the effect
of PCIT-ED on caregiver responses to their child’s emotions. During the ED module
specifically, caregivers in the treatment group reported significantly greater improvement over
time in their use of expressive encouragement and reductions in their use of distress, punitive,
emotion-focused, and minimization reactions in response to their child’s negative emotions than
those in the control group (Luby et al., 2020). The use of more expressive encouragement (i.e.,
emotion coaching) and less distress responding suggest that caregivers were better able to
regulate their emotions, at least in the context of their child’s emotional distress, upon
completion of the ED phase of treatment. These results highlight the unique contributions of the
ED module in improving emotional functioning for both caregivers and their children.

PCIT with parent emotion coaching (PCIT-ECo; Chronis-Tuscano et al., 2016) is an
adaptation of PCIT-ED for children with ADHD that systematically incorporates emotion
coaching into PDI, removes elements of PCIT-ED that are less relevant or feasible for young
children with ADHD (e.g., joy coaching, relaxation techniques), and adds treatment components
targeting ADHD symptoms (e.g., school-home daily report card, psychoeducation regarding
school supports). In a pilot study investigating the use of PCIT-ECo for three children with ADHD-combined type (Chronis-Tuscano et al., 2016), clinically significant improvements (as evidenced by the reliable change index) in parent-reported emotional lability/negativity on the ERC were observed at post-treatment and maintained at a 6-month follow-up assessment. Per teacher report on the ERC, however, only one child demonstrated reliable change in emotional lability/negativity at post-treatment. Parent and teacher report of child externalizing behavior on the Child Behavior Checklist/Teacher Report Form (Achenbach & Rescorla, 2000) demonstrated improvements at post-treatment and at follow-up, though only two produced clinically significant change. Despite these favorable treatment outcomes, it is difficult to assess the degree to which emotion coaching contributed to the improvements in child emotion regulation and if it contributed above and beyond the treatment components found in standard PCIT.

In another adaptation of PCIT, Dialectical Behavior Therapy (DBT) skills were incorporated into a time-limited delivery of PCIT for six parent-child dyads at an outpatient mental health clinic (Rohrig, 2019). Parents were taught and coached to use select DBT techniques (i.e., mindfulness of parents’ current emotions, validation of parents’ and child’s emotions, and distress tolerance) in addition to the standard PCIT skills (i.e., PRIDE skills, effective command delivery, time-out procedure). Using a repeated-measures, single-subject design, child and caregiver emotion regulation skills were measured across sessions using parent-report measures (ERC and DERS, respectively). Of the six parents, four reported a drop in DERS Total scores from pre- to post-treatment, but only two of these parents were below the clinical cutoff by the end of treatment. With regard to child emotion regulation, four of the six children demonstrated improvements (two on the Emotion Regulation subscale and two on the Lability/Negativity scale) with large effect sizes (i.e., standardized mean difference > .8).
RCT COMPARING PCIT-T AND COS-P

Additionally, caregiver responses to children’s negative emotions were evaluated at pre- and post-treatment using the Coping with Children’s Negative Emotions Scale (CCNES; Fabes et al., 1990). Visual inspection of the data indicated an increase in “Supportive” parental responses to negative child emotion for three of the six participants following completion of PCIT.

Taken together, these studies provide evidence for the effectiveness of PCIT in improving child and caregiver emotion regulation and thus suggest the potential for PCIT-T to have similar treatment effects, especially considering the components of PCIT-T that are specifically designed to target these constructs. However, the lack of studies examining PCIT-T specifically highlights the need for continued research on this novel adaptation as an intervention to treat child and caregiver emotion regulation.

Circle of Security

Circle of Security-Intensive. Circle of Security-Intensive (COS-I; Cooper et al., 2005; Powell et al., 2009) is a 20-week, manualized parenting intervention for children birth to 5 years of age that is rooted in attachment theory and object relations theory. A primary goal of COS-I is to improve attachment patterns by increasing caregiver awareness of problematic caregiver-child interactions and the internal processes underlying them (Powell et al., 2014). In COS-I, caregivers are encouraged to reflect on their own upbringing and attachment history in order to understand the psychological defenses that may preclude a more secure attachment with their child. Over the course of treatment, caregivers learn to monitor their own cognitive and affective responses to their child’s distress, use observational and inferential skills to interpret their child’s actions, and respond appropriately to their child’s physical and emotional needs. A skill that is emphasized and developed throughout treatment is parental reflective functioning, or the ability to understand their own psychological state and experiences as well as those of their child’s
RCT COMPARING PCIT-T AND COS-P

(Powell et al., 2014). COS-I aims to increase parental capacity for reflective functioning, thereby allowing them to recognize how their own emotional and cognitive states impact their parenting and eventually overcome their insecure or disorganized attachment styles.

A hallmark of COS-I is the use of diagrams, metaphors, and videos to highlight key concepts and facilitate therapeutic change. One such diagram that is used throughout treatment is that of the Circle, which depicts three components integral to caregiver-child attachment: the attachment system, the exploratory system, and the caregiving system (Powell et al., 2014). The attachment system refers to the caregiver’s role as a safe haven that the child can return to for comfort. The exploratory system refers to the caregiver as a secure base from which the child can explore and gain competency in their environment. The caregiving system refers to the caregiver’s role as someone who is “bigger, stronger, wiser, and kind” (Powell et al., 2014, p. 31) in relation to the child, follows the child’s need whenever possible, and takes charge when necessary. The COS-I intervention is built on the premise that caregivers who recognize their child’s needs and can move adeptly from one role to another are better able to facilitate secure caregiver-child attachments.

The concept of “shark music” is a metaphor that is used frequently during the COS-I intervention and refers to the discomfort that caregivers feel with certain needs expressed by their child (Powell et al., 2009). The term makes reference to the iconic theme music in the movie Jaws that has become universally known as a signal for imminent danger. In COS-I, caregivers learn to recognize shark music, the situations that evoke shark music, and the defensive reactions that it elicits. Over the course of treatment, caregivers become increasingly aware of the ways in which shark music can prevent them from noticing and supporting their
child’s needs and the inadvertent messages they may be communicating to their child through these unmet needs (i.e., that the child’s needs are unacceptable or unsafe).

COS-I begins with an assessment procedure (a modified version of the Strange Situation Protocol; Ainsworth et al., 1978) that involves videotaping a dyadic interaction between the caregiver, child, and a stranger. This recorded interaction is then analyzed by the therapist to determine the targets of intervention. Video footage is selected by the therapist to highlight episodes of interactional struggles and strengths. The purpose of video review, which occurs on three separate occasions over the course of the program, is to help caregivers develop their observational and inferential skills and highlight the areas that require additional attention and work.

**Circle of Security-Parenting.** Due to the time-intensive nature of the 20-week intervention, an 8-module adaptation, Circle of Security-Parenting (COS-P; Cooper et al., 2009), was developed in order to increase feasibility and acceptability among caregivers and treatment providers alike. It is important to note that this abbreviated adaptation was used for the current study in lieu of the intensive standard protocol. COS-P is a DVD-based parent education program that can be delivered in a variety of clinical contexts (e.g., individual therapy, group therapy, home visitation). A significant change made to the standard protocol is the use of stock footage of secure and problematic caregiver-child interactions in lieu of video footage of the pre-treatment interactional assessment. However, many of the treatment components of the original model are maintained, including observational and inferential skill training and engagement in reflective dialogue.

Facilitators of COS-P use a DVD, handouts, and manual to illustrate key concepts, guide sessions, and engage participants in discussion and self-reflection. The first two modules provide
RCT COMPARING PCIT-T AND COS-P

psychoeducation on attachment and introduce the Circle as a roadmap for navigating caregiver-child interactions (Cassidy et al., 2017). In modules 3 and 4, caregivers learn about their role as a safe haven for their child and responding to their child’s emotional needs. The next module focuses on caregivers’ attachment histories and introduces the concept of shark music. In this module, caregivers learn to recognize and anticipate their defensive responses, engage in self-calming strategies, and respond appropriately to their child’s needs. Modules 6 and 7 review the factors that contribute to disorganized attachments and the process of repairing ruptures in the caregiver-child relationship. The final module involves a summary of content learned, a reflection of caregiver experiences in the program, and celebration of caregivers’ successes.

**COS-P and Emotion Regulation.** Research on the effects of COS-P on caregiver and child emotion regulation is lacking, and the limited number of studies examining this topic have been with children outside of the toddler age range. To date, only one randomized controlled trial of COS-P has been conducted (Cassidy et al., 2017). In a Head Start program, 148 children (ages 3-5 years) were randomly assigned to receive group-based COS-P or remain on the waitlist to receive the intervention. An examination of treatment completers found that those in the intervention group saw a significant reduction with a small effect size ($d = 0.37$) in mothers’ unsupportive responses to child’s negative emotions as reported on the Coping with Toddlers’ Negative Emotions Scale (CTNES; Spinard et al., 2004). Further, children in the COS-P group demonstrated significantly greater improvements ($d = 0.40$) in inhibitory control on a “Simon Says” task than children in the control group. However, no significant between-group differences were found in child externalizing behavior, internalizing behavior, or caregivers’ supportive responses to child distress.
Krishnamoorthy et al. (2020) conducted a quasi-experimental study of 54 foster parents and their children (ages ranging from 6 to 12 years) in Australia and found significant improvements on all scales of the Parenting Stress Index-4 Short Form (PSI-4-SF; Abidin, 2012) and a small but significant decrease in scores on the Emotional Symptoms subscale on the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). However, the lack of control or comparison group made it difficult to rule out other confounding factors that may have contributed to these improvements.

The remaining two studies of COS-P and emotion regulation employ a single-subject design. Pazzagli and colleagues (2014) conducted a case study of COS-P in the context of individual psychotherapy with a 43-year-old father and 5-year-old daughter in Italy. In addition to measures of attachment and parenting alliance, variables associated with caregiver and child emotion regulation were examined at pre- and post-treatment. Italian versions of the Parenting Stress Index-Short Form (PSI-SF; Abidin, 1995; Guarino et al., 2008) and the Strengths and Difficulties Questionnaire (Goodman, 1997; Di Riso et al., 2010) were administered and the reliable change index (RCI) was used to determine clinically significant changes upon treatment completion. Per parent report, significant improvements on the PSI-SF Total Stress, Parental Distress, and Parent-Child Dysfunctional Interaction scales (but not the Difficult Child scale) were found. Additionally, clinically significant improvement on the SDQ Total Difficulties Score was found but not on the Emotional Symptoms or Conduct Problems subscales. These findings suggest that COS-P may be effective in reducing caregiver stress around parenting and the parent-child interaction but not necessarily in improving child emotional and behavioral problems. Similarly, Kim et al., (2018) conducted a case study of 26-year-old mother and her 4-year-old son receiving COS-P in a group-based format. Pre-post decreases in scores on the
RCT COMPARING PCIT-T AND COS-P

Internalizing, Externalizing, and Total Problems scales on the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001), though scores were within normal limits at pre-treatment. No changes were reported in parenting stress or parental responses to child’s distress. Further study with more robust methodology (e.g., multiple baselines, control group) is needed to draw conclusions about the effectiveness of COS-P on caregiver and child outcomes.

**Purpose**

The ability to effectively regulate emotions is an important developmental task during toddlerhood and failure to develop adaptive emotion regulation during this stage can have deleterious consequences later in life (Keenan, 2000). Given that caregivers play a pivotal role in the development of toddlers’ regulatory abilities, parent-mediated interventions are the preferred treatment approach for treating emotion regulation difficulties in young children. However, the parenting interventions specifically targeting toddlers’ emotion regulation are scarce in number, and there is limited outcome research on the interventions that are available. Thus, the main purpose of the current study was to determine the efficacy of PCIT-T on child and caregiver emotion regulation and related variables (e.g., parental emotion dismissing, child disruptive behaviors) in comparison to a popular parenting program (i.e., COS-P) and a waitlist control group. Additionally, the association between pre-treatment variables were explored, as were differences in attrition between groups.

**Method**

**Study Design**

The current study used data collected as part of a larger randomized controlled trial that randomly assigned families to one of three conditions (PCIT-Toddler, COS-Parenting, or waitlist control groups) to examine a number of caregiver and child outcome variables, including
attachment security, positive parenting skills, negative parenting behaviors, and child social-emotional functioning. The project was funded by the Karitane Toddler Clinic and the University of New South Wales. For more details about the parent study, see the study protocol published by Kohlhoff et al. (2020).

**Participants**

Participants were recruited from families referred to an early childhood outpatient behavioral health clinic in New South Wales, Australia. Families were informed of the study after referral to the clinic and could choose to participate in the study if they met the following inclusion criteria: (1) had a child aged 14-24 months, (2) were referred to the clinic by a health professional, and (3) had endorsed one or both of the two screening questions (“Do you have concerns about your child’s behavior?” or “Do you have difficulties managing your child’s behavior?”). Families were excluded from the study if there was evidence of severe mental health conditions (including but not limited to severe depression, suicidality, and psychosis) from the referral or intake interview or if they had limited English proficiency, as the study measures were written in English (Kohlhoff et al., 2020). A total of 91 parent-child dyads enrolled in the parent study. For the current study, an additional exclusion criterion was included; children with a diagnosis of autism spectrum disorder were excluded from the study due to the neural basis for emotion regulation difficulties found among individuals with autism (Mazefesky et al., 2020; Richey et al., 2015). As a result, a subset of 76 dyads from the parent study were included in the current investigation: $n = 24$ in the PCIT-T group, $n = 27$ in the COS-P group, and $n = 25$ in the waitlist control group. Only treatment completers (i.e., those who completed measures at the first and second timepoints) were included in the main analyses; these subsamples differed depending
on the variable being examined but ranged from 43 to 50 dyads. Thus, sample sizes for each outcome variable are reported in the results section.

Demographic characteristics of the sample were collected at baseline. The children averaged 19.13 months of age (Range: 14-24 months; $SD = 2.98$) and were 50% female. The mean age of the mothers in the sample was 32.94 years (Range: 19-45 years; $SD = 5.51$), and the majority of mothers were married 60.5% ($n = 46$). Of the mothers in the sample, 26.3% ($n = 20$) were stay-at-home mothers, 25% ($n = 19$) were professionals, 13.2% ($n = 10$) were community and personal service workers, 6.6% ($n = 5$) were clerical and administrative workers, 5.3% ($n = 4$) were managers, 3.9% ($n = 3$) were students, 3.9% were technicians and trade workers ($n = 3$), 1.3% ($n = 1$) were sales workers, and 14.5% ($n = 11$) did not report an occupation. In terms of maternal education level, 5.3% ($n = 4$) had completed Year 10 (last year of compulsory education or junior high school in Australia); 7.9% ($n = 6$) had completed Year 12 (last year of senior high school); 30.3% ($n = 23$), had completed vocational/technical school; 34.2% ($n = 26$) had an undergraduate degree, 7.9% ($n = 6$) had a graduate degree; and 14.5% ($n = 11$) did not report their educational level. There was a range of income levels (reported in Australian dollars, or AUD) represented in the sample: 18.4% ($n = 14$) estimated family income to be less than $50,000 per year, 10.5% ($n = 8$) estimated between $50-75,000 per year, 6.6% ($n = 5$) estimated $76-100,000 per year, 34.2% ($n = 26$) estimated $101-150,000 per year, 18.4% ($n = 14$) estimated more than $150,000 per year, and 11.8% ($n = 9$) did not report their income level. (At the time this document was written, 1 Australian dollar was equivalent to 0.74 US dollars.) The sample comprised 39.5% Caucasian ($n = 30$), 11.8% Asian ($n = 9$), 9.2% Middle Eastern ($n = 7$), 6.6% European ($n = 5$), 5.3% Other ($n = 4$), and 3.9% Hispanic ($n = 3$) mothers. Eighteen mothers (23.7% of the sample) did not report their ethnicity. The majority of mothers spoke
RCT COMPARING PCIT-T AND COS-P

English as their primary language (69.7%; n = 53), though 15 mothers spoke a different primary language (e.g., Arabic, Spanish, Australian, Farsi, Bengali, Cantonese, Korean). The average number of individuals living in the household was 3.72 (Range: 2-6; SD = 1.08).

Procedure

Families who met eligibility criteria provided written, informed consent to participate in the parent RCT study. Block randomization (block size of $n = 6$) was used to randomly allocate participants to the three study conditions: PCIT-Toddler, Circle of Security-Parenting, or waitlist control. Data were collected across two timepoints: pre-treatment (T1) and post-treatment (T2).

Time 1 Assessment

A combination of observational and parent-report measures was administered at T1. All observational tasks took place at the clinic in a playroom with a one-way mirror and a standardized set of developmentally appropriate toys. The first observational task was a series of 5 consecutive 5-minute scenarios involving the parent-child dyad, each beginning with standardized instructions delivered to the parent through a wireless headset. In the first scenario, Free Play, parents were instructed to play with their child in an unstructured manner. In the second scenario, Child-Led Play, parents were instructed to allow their child to take the lead and follow along with their child’s play. After a brief warm-up period and five minutes of Child-Led Play, a researcher introduced the third scenario, Toy Removal, informing the parent through the wireless headset that she would be entering the playroom to remove all of the toys. Parents were instructed to tell their child that the therapist would be coming to take away the toys and to manage their child as they typically would after the therapist left the room with the toys. After 5 minutes, the therapist returned the toys to the playroom for the fourth scenario, Toy Reunion. For the fifth scenario, Clean-Up, parents were instructed to tell their child to clean up the toys and to
RCT COMPARING PCIT-T AND COS-P

refrain from assisting their child with this task. All dyadic interactions were videorecorded, transcribed, and coded for parent and child verbalizations and child vocalizations.

The second observational task was a 5-minute frustration task in which children were presented with a locked box containing an attractive toy (see Measures section). Through the wireless headset, parents were instructed to complete questionnaires and attend minimally to their child’s overtures for the duration of this task. This observational task was also videorecorded, transcribed, and coded for child emotion regulation.

Additionally, parents were given a packet of questionnaires (see Measures section) to complete at home and return to the research team at the T2 assessment. Families were not informed of their group allocation until after the completion of the T1 assessment battery.

Time 2 Assessment

For those assigned to a treatment condition (i.e., PCIT-T or COS-P), the intervention began as soon as possible following the completion of the T1 assessment. The T2 assessment was administered 8 weeks later, immediately following the completion of the intervention. At this time, the waitlist control group completed their T2 assessment as their post-waitlist assessment. As compensation for their time, participants were given a $25 gift card and a small gift for the child after the completion of each assessment.

Interventions

Families assigned to the PCIT-T condition received the intervention at the clinic as outlined in the PCIT-T protocol (Girard et al., 2018). Families received a Child-Directed Interaction for Toddlers (CDI-T) didactic session followed by 6-8 CDI-T coaching sessions. Families then received a Parent-Directed Interaction for Toddlers (PDI-T) didactic session followed by 2-4 PDI-T coaching sessions. Treatment took place over a period of 8 weeks with 1-
RCT COMPARING PCIT-T AND COS-P

2 parent-only or parent-child 45-minute sessions per week, totaling 16 intervention hours.

Treatment fidelity was monitored in two ways: fidelity checklists completed by therapists at the end of every session and regular supervision meetings led by one of the treatment developers.

Families assigned to the COS-P condition received the intervention at the clinic as outlined in the COS-P protocol (Cooper et al., 2009). The intervention was delivered in a group format with two facilitators trained and certified in COS-P. Treatment consisted of 8 parent-only, 2-hour group sessions, totaling 16 intervention hours. Treatment fidelity was monitored in two ways: fidelity checklists completed by group facilitators after every session and regular supervision meetings led by a seasoned COS-P facilitator.

**Therapists**

A total of ten therapists delivered the interventions for the current study. Therapists were employed at the clinic and working in the area of perinatal, infant, and early childhood mental health. Therapists were selected based on clinical availability at the time of the COS-P groups and PCIT-T cases. Three of the therapists were designated to administer PCIT-T to participating families, and the remaining seven therapists administered COS-P. No therapist administered or assisted in administering both interventions to study participants. All three PCIT-T therapists were certified by PCIT International as either Within-Agency Trainers or Regional Trainers. Two of the PCIT-T therapists were clinical psychologists and the third was a clinical nurse consultant. COS-P groups were co-led by two trained facilitators, one of whom was not trained in PCIT or PCIT-T. Of the seven COS-P facilitators, four were clinical psychologists, one was a psychiatrist, and 2 were nurses.

**Waitlist**
RCT COMPARING PCIT-T AND COS-P

Families assigned to the waitlist condition completed the T1 assessment and did not receive treatment for a period of 8 weeks. Following the completion of the T2 assessment, they were given the option to receive the treatment of their choice (i.e., PCIT-T or COS-P).

**Coding Procedure**

For DPICS and DECS coding, video footage of observational tasks were transcribed by undergraduate research assistants. After transcriptions are completed, trained undergraduate and graduate coders used the transcripts and video footage to code caregiver statements using both the Dyadic Parent-Child Interaction Coding System (DPICS-IV; Eyberg et al., 2014) and the Dyadic Emotion Coding System (DECS; Owen et al., in preparation). All coders first received didactic training on the DPICS led by graduate students proficient in the coding system. Training was supplemented with coding exercises and quizzes, as well as ongoing feedback and support from the graduate and advanced undergraduate coders. Coders who received a score of 80% or higher on assigned quizzes were permitted to double-code three training videos. Coders who obtained a kappa of .80 or above were permitted to code videos used in the current study. Coders who did not meet the quiz score or interrater reliability requirement underwent additional training with either graduate or advanced undergraduate coders before they were provided with another quiz to complete or training video to code. Once DPICS coding was completed for all videos, the same training and coding procedure was employed for the DECS.

For Child Distress and Emotion Regulation Maturity coding, two independent coders viewed video recordings of the task and coded first for toddler distress and then for emotion regulation maturity and maternal support seeking. The coders used a coding system developed by Johnson (2015) that was based on Gross’ (2013) process model of emotion. The first coder coded
RCT COMPARING PCIT-T AND COS-P

70% of the videos, whereas the second coder coded 30% of the videos. Thirty videos were
double coded and intra-class coefficients were calculated.

Interrater reliability for each coding scheme or rating scale was determined using
Cohen’s kappa for nominal data and intraclass coefficient (ICC) for ordinal and interval data.
Qualitative descriptors are included in the reporting of kappa statistics (Landis & Koch, 1977)
and ICC (Koo & Li, 2016). For DPICS Parent Negative Talk, overall agreement between two
coders was almost perfect, $\kappa = .846$, 95% CI [.831 to .860]. For DPICS Child Negative
Verbalizations, overall agreement between two coders was substantial, $\kappa = .723$, 95% CI
[.662, .808]. For DECS codes, reliability was almost perfect between two coders, $\kappa = .905$, 95%
CI [.884, .927]. The ICCs for attentional deployment, situation modification, cognitive change
and maternal support seeking were .80, .88, .98 and .93, respectively, indicating good to
excellent reliability.

Measures

A combination of parent-report and observational measures was used to measure
caregiver and child emotion regulation and related constructs.

Demographic Information

A demographic information form was administered to all caregivers at pre-treatment.
Caregivers reported on their date of birth, occupation, education level, family income, ethnicity,
marital status, primary language, other language spoken in the home, and number of people
within the household, as well as their child’s date of birth and sex.

Caregiver Emotion Regulation

Difficulties in Emotion Regulation Scale. The Difficulties in Emotion Regulation Scale
(DERS; Gratz & Roemer, 2004) is a 36-item questionnaire that provides a comprehensive
measure of emotion dysregulation in adults. For each item, respondents provide a rating on a 5-point Likert-type scale to indicate the degree to which a statement applies to them, such that a rating of 1 is “almost never” and 5 is “almost always.” The measure yields a total score as well as six subscales. The six subscales are as follows: (1) Nonaccept (i.e., difficulty accepting one’s negative emotions), (2) Goals (i.e., difficulty concentrating on and completing tasks when experiencing negative emotions), (3) Impulse (i.e., difficulty controlling one’s behavior when experiencing negative emotions), (4) Awareness (i.e., lack of emotional awareness or difficulty attending to emotions), (5) Strategies (i.e., difficulty accessing strategies to regulate negative emotions effectively), and (6) Clarity (i.e., lack of emotional clarity). Higher scores indicate greater difficulty with emotion regulation. The DERS has demonstrated sound psychometric properties, including high internal consistency (α = .93), incremental validity in relation to other measures of ER, predictive validity, and adequate test-retest reliability across a period of four to eight weeks (Gratz & Roemer, 2004). The DERS has demonstrated reductions in scores following the completion of treatments targeting emotion regulation (Ben-Porath et al., 2014; Gratz et al., 2015; Gratz et al., 2014). This suggests that the DERS is sensitive to change and is appropriate for use in the current study and in other treatment outcome research. For the purposes of the current study, only the DERS Total score was used in the analyses.

Variables Associated with Caregiver Emotion Regulation.

DPICS Parent Negative Talk. The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV; Eyberg et al., 2014) is an observational coding system of caregiver-child interactions during three structured tasks: Child-Led Play, Parent-Led Play, and Clean-Up. The DPICS is used in PCIT to determine whether or not caregivers have reached CDI and PDI goal criteria and to establish behavioral targets for each coaching session. One particular code
from the DPICS that was studied in this investigation was Parent Negative Talk (NTA). Parent NTA is defined as verbal disapproval of the child or the child’s characteristics, activities, or products; it also includes sarcastic or rude remarks directed toward the child (Eyberg et al., 2013). Some examples of Negative Talk are “Stop yelling,” “That’s not how we play with the blocks,” or “Smooth move,” after a child accidentally knocks over a cup. The psychometric properties of the DPICS have been well-documented (see *DPICS-IV Comprehensive Manual for Research and Training* for psychometric data; Eyberg et al., 2013). Psychometric characteristics of the Parent NTA code specifically have been cited in a number of studies. Parent NTA has demonstrated adequate intercoder reliability (Bagner & Eyberg, 2007; Callahan & Eyberg, 2010) and sensitivity to change after PCIT (Chaffin et al., 2004; Eyberg et al., 2001; McCabe & Yeh, 2009). More maternal Negative Talk at pre-treatment has also been found to predict a higher likelihood of treatment dropout (Fernandez & Eyberg, 2009) and child noncompliance to parental commands during Clean-Up (Richerson, 2008), providing support for its predictive validity.

Quetsch and colleagues (2018) found that significantly more Negative Talk was used by parents of children with externalizing behavior problems than parents of children without behavior problems and that NTA was significantly and positively correlated with self-reported difficulty in emotion regulation on the DERS. Thus, it was proposed that NTA could be conceptualized as a measure of parental emotion regulation, namely for families with greater emotion regulation challenges (Quetsch et al., 2018). For the purpose of the current study, the number of NTA codes was summed across all observational tasks (i.e., Free-Play, Child-Led Play, Toy Removal, Toy Reunion, and Clean-Up) and divided by the total number of parent verbalizations to yield a negative talk percentage.
DECS Emotion Dismissing. The Dyadic Emotion Coding System (DECS; Owen et al., in preparation) is a novel observational coding system initially developed to capture caregiver labeling of child emotions during caregiver-child dyadic interactions. The DECS codes emotion-related caregiver verbalizations on three dimensions: (1) the emotion label, (2) valence of emotion, and (3) intensity of emotion. For the current study, however, only the Emotion Dismissing (ED) code from the first dimension was examined. Emotion Dismissing is defined as a verbal expression of negative evaluation of the child’s emotion (see Figure 1 for an excerpt with examples of the ED code from the DECS Coding Manual). The validity and reliability of the DECS are currently being evaluated for caregiver-toddler dyadic interactions (Owen et al., in preparation). However, emotion dismissal has been linked to negative child outcomes, such as poorer child emotion regulation and more externalizing and internalizing problems (Lunkenheimer et al., 2007). Children with parents characterized as “emotion dismissing” (Gottman et al., 1997) are more likely to view their negative emotions as inappropriate or invalid (Gottman & DeClaire, 1997). More emotion dismissal has also been significantly associated with greater concern in achieving parent-centered goals as opposed to empathetic goals centered around fostering a positive dyadic relationship (Lagacé-Séguin & Coplan, 2005). For the purpose of the current study, the number of ED codes was summed across all observational tasks (i.e., Free-Play, Child-Led Play, Toy Removal, Toy Reunion, and Clean-Up) and divided by the total number of parent verbalizations to yield an emotion dismissing percentage.

Child Emotion Regulation

Attractive Toy in Transparent Box (ATTB) Task. Child emotion regulation was measured using a toy retrieval task that was designed to elicit momentary frustration in the children. This paradigm was based on a task from the Laboratory Temperament Assessment
Battery (Lab-TAB; Goldsmith et al., 1999), a standardized instrument for the assessment of early temperament. Children were given a large Plexiglas box (120 x 60 x 60 cm) containing a colorful toy with flashing lights. To elicit frustration in the child, the box was made impossible for a toddler to open without assistance; additionally, there was a hole at one end of the box that was large enough for the toddler to reach inside but too small to extract the toy from the box. Children were left with the transparent box for five minutes while their parents were completing questionnaires in the same room. Parents were instructed to provide minimal attention to their child’s overtures for the duration of the task.

**Child Distress.** Intensity of child distress was measured on a 5-point rating scale, with 0 indicating no distress and 4 indicating extreme distress (see Figure 2 for rating form). Videos of the toy retrieval task were divided into 10-second intervals, and each interval was given a rating of child distress by research assistants blinded to group allocation. Total Distress scores were calculated by summing the 30 scores across each 5-minute video.

**Maturity of Emotion Regulation Strategies.** Maturity of the child’s emotion regulation strategies was measured using a coding system based on the process model of emotion regulation (Gross, 1988). Ten-second intervals with a child distress rating of 0 (i.e., no distress) were given one of four codes: Situation Modification, Attention Deployment, Cognitive Change, or Maternal Support Seeking (see Figure 3 for coding form). Situation Modification was coded when a child actively modifies the distressing situation to manage their emotional state. Attentional Deployment was coded when a child’s behavior indicates an attentional shift away from the distressing situation to manage their emotional state. Cognitive Change was coded when a child’s behavior indicates an acceptance or reinterpretation of the distressing situation. These three codes represented intrinsic emotion regulation strategies. Lastly, Maternal Support Seeking
RCT COMPARING PCIT-T AND COS-P

was coded when the child was observed seeking assistance, acknowledgement, or comfort from the mother in the absence of distress. This code represented an external emotion regulation strategy. These codes are hierarchical in nature with Situation Modification reflecting the least mature emotion regulation strategy and Cognitive Change reflecting the most mature strategy. A total Emotion Regulation Maturity score was calculated by multiplying the totals codes for Situation Modification, Attentional Deployment, and Cognitive Change by a factor of 1, 2, or 3, respectively (i.e., Emotion Regulation Maturity score = Situation Modification score * 1 + Attentional Deployment score * 2 + Cognitive Change score * 3). This total score was then divided by 30 (i.e., the number of intervals in the coding period) to yield an average; this score is referred to in the document as the Emotion Regulation Maturity composite score. Higher scores were indicative of greater maturity of emotion regulation strategies.

Variables Associated with Child Emotion Regulation

**DPICS Child Negative Talk, Whine, and Yell.** Child Negative Talk (NTA) is a code taken from the *DPICS-IV Comprehensive Manual for Research and Training* (Eyberg et al., 2013). Similar to the Parent Negative Talk code, the Child Negative Talk code captures verbal disapproval of the caregiver or the caregiver’s characteristics, activities, or products; it also includes sarcastic or rude remarks directed toward the caregiver (Eyberg et al., 2013). Some examples of Child Negative Talk are “Don’t touch my tower,” “You can’t use markers on this,” or “No,” after a parent tells a child to pick up the toys. Child Negative Talk has not been given as much scholarly attention as Parent Negative Talk; thus, psychometric data is relatively limited. Eyberg et al. (1995) found that approximately 10% of verbalizations at pre-treatment were categorized as NTA for clinic-referred children compared to 1% of verbalizations for non-referred children, though this difference was not statistically significant. Cotter and Brestan-
Knight (2020) provided evidence of the code’s convergent validity in that children who used more Negative Talk and were less compliant with parental commands had higher parent ratings on the ECBI Intensity and Behavior Assessment System for Children-Second Edition, Parent Rating Scale (BASC-2 PRS) Externalizing Problems scales.

Child Whine (WH) and Yell (YE) are two distinct codes from the Appendix of the DPICS-IV (Eyberg et al., 2013) that fall under the category of child vocalizations. In the DPICS-IV, a vocalization code is distinguished from a verbalization code in that the focus is on the tone of voice used as opposed to the content of the speech uttered. A Whine is defined as an utterance delivered in a “slurring, moaning, high-pitched, or falsetto voice” (Eyberg et al., 2013, p. 223). A Yell is defined as a “screech, scream, or shout” (Eyberg et al., 2013, p. 221) or any utterance that is unpleasantly loud. If a vocalization is both a Yell and a Whine (e.g., whining that is loud in volume), the YE code takes priority over the WH code. Each episode of a Yell or Whine is characterized by termination of speech or a pause that is 2 seconds or longer. Unlike a verbalization code, a vocalization can be coded with or without intelligible speech. Thus, vocalization codes can be used independently or in conjunction with a verbalization code. For example, a child screaming “Ahh!” would be given a YE code, as would a child who yells, “You’re stupid!” (YE-NTA). Robinson and Eyberg (1981) found that clinically referred children displayed significantly more whining and yelling than did children in a normative group in CLP and PLP scenarios, providing support for the discriminant validity and clinical utility of these vocalization codes. The Yell and Whine codes have demonstrated good interrater reliability based on both kappa and intraclass correlations (Bessmer, 1996; Foote, 1999). When a composite variable combining the YE, WH, and NTA codes was used, significant group differences were found between clinic-referred and non-clinic-referred children (Foote, 1999). This composite
RCT COMPARING PCIT-T AND COS-P

score was significantly and positively correlated with parent-reported child disruptive behavior on the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999); Child Domain and Total Stress scores on the PSI (Abidin, 1995); and the Parental Locus of Control-Short Form (PLOC-SF; Rayfield et al., 1995). Additionally, frequency of Child Yell and Whine during the DPICS observation was found to be positively and significantly correlated to parent ratings on the ECBI Intensity and BASC-2 PRS Externalizing Problems and Adaptive Skills scales (Cotter & Brestan-Knight, 2020), providing support for the convergent validity of these codes. For the purpose of the current study, Child NTA, YE, and WH codes were combined across observational tasks to form a single Child Negative Verbalizations composite variable.

Child Dysregulation. The Child Behavior Checklist/1.5-5 (CBCL; Achenbach & Rescorla, 2001) is a 99-item questionnaire for caregivers of young children that assesses for a variety of internalizing and externalizing problems. It is important to note that the CBCL/1.5-5 has only been validated for use with children between the ages of 18 months and 5 years. Although a portion of the current sample is below the 18-month age cut-off for the CBCL, the CBCL was chosen for the proposed study for several reasons. First, the CBCL is one of the few general assessment of emotional and behavioral functioning available to children in the toddler age range. Second, the CBCL was interpreted based on comparison groups that also include children under the 1.5-year age cut-off. Third, those younger than the cut-off aged into the acceptable age range over the course of treatment. Nevertheless, main analyses involving this variable were conducted twice: once with the entire sample of treatment completers and once with only the treatment completers who were in the valid age range for the CBCL.

For each item on the CBCL, respondents provide a rating for each child problem behavior on a 3-point scale (0 = not true, 1 = somewhat or sometimes true, and 2 = very true or
often true). The CBCL consists of three broadband scales: Externalizing Problems, Internalizing Problems, and Total Problems. It also yields scores on seven syndrome scales (Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Attention Problems, Aggressive Behavior, and Sleep Problems) and five DSM-oriented scales (Depressive Problems, Anxiety Problems, Autism Spectrum Problems, Attention-Deficit/Hyperactivity Problems, and Oppositional Defiant Problems). The CBCL/1.5-5 has demonstrated good test-retest reliability ($r = .68$ to $.92$) and discriminative validity, with elevated scores on the Externalizing and Internalizing Problems scales accurately identifying $74\%$ of clinically referred children (Achenbach & Rescorla, 2001).

For the current study, a Dysregulation Profile (CBCL-DP) score was calculated by summing the scores from the Anxious/Depressed, Attention Problems, and Aggressive Behavior syndrome scales. Although originally created to identify children with bipolar disorder (Biederman et al., 1995), the CBCL-DP has become more widely accepted and used as a measure of dysregulation and functional impairment independent of a specific DSM diagnosis (Althoff et al., 2010). The CBCL-DP can thus be conceptualized as a measure of self-regulation that spans three domains: affective dysregulation (Anxious/Depressed); behavioral dysregulation (Aggressive Behavior), and cognitive dysregulation (Attention Problems; Althoff et al, 2010).

The CBCL-DP has demonstrated good diagnostic validity (Jucksch et al., 2011) and convergent validity, with elevated CBCL-DP scores strongly associated with increased suicidality (Althoff et al., 2006; Ayer et al., 2009; Holtmann et al., 2008) and greater psychosocial impairment (Hazell et al., 2003). Longitudinal studies have found that elevated CBCL-DP scores in childhood predicted substance use disorders, suicidal thoughts and behavior,
and greater functional impairment in adulthood (Biederman et al., 2009; Holtmann et al., 2010; Meyer et al., 2009), demonstrating good predictive validity of the Dysregulation Profile.

**Child Externalizing Behavior.** The Externalizing Problems scale from the CBCL/1.5-5 was used as a measure of child externalizing behavior. This broadband scale is formed by combining the Aggressive and Attention Problems subscales. Main analyses involving this variable were conducted twice: once with the entire sample of treatment completers and once with only the treatment completers who were in the valid age range for the CBCL.

**Data Analysis Plan and Hypotheses**

**Caregiver Variables**

**Hypothesis 1: Parent Emotion Regulation.** It was hypothesized that caregivers in the two treatment groups would demonstrate significantly greater pre-post improvements in caregiver emotion regulation than those in the control group as reported on the DERS. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for DERS Total scores. Effect sizes were reported as partial eta squared ($\eta^2_p$).

**Hypothesis 2: Parent Negative Talk.** It was hypothesized that caregivers in the PCIT-T group would have a significantly greater pre-post reduction in the number of negative verbalizations than those in the COS-P and waitlist control groups as coded using the DPICS. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for the DPICS Parent Negative Talk code. Effect sizes were reported as partial eta squared ($\eta^2_p$).

**Hypothesis 3: Parent Emotion Dismissing.** It was hypothesized that caregivers in both treatment groups would demonstrate significantly greater pre-post reductions in emotion
dismissing statements than caregivers in the control group as coded using the DECS. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for the DECS Emotion Dismissing code. Effect sizes were reported as partial eta squared ($\eta^2_p$).

**Child Variables**

**Hypothesis 4: Distress Rating.** It was hypothesized that children in both treatment groups would demonstrate significantly greater pre-post improvements in child distress than children in the control group. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for Child Distress ratings. Effect sizes were reported as partial eta squared ($\eta^2_p$).

**Hypothesis 5: Emotion Regulation Maturity.** It was hypothesized that children in both treatment groups would demonstrate significantly greater pre-post improvements in child distress than children in the control group. To test this hypothesis, repeated measures ANOVA were conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for both the Emotion Regulation Maturity composite and Maternal Support Seeking code. Effect sizes were reported as partial eta squared ($\eta^2_p$).

**Hypothesis 6: Child Dysregulation.** It was hypothesized that children in both treatment groups would demonstrate significantly greater pre-post improvements in child dysregulation as reported by caregivers on the CBCL. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for the CBCL Dysregulation Profile score. Effect sizes were reported as partial eta squared ($\eta^2_p$).
Hypothesis 7: Child Externalizing Problems. It was hypothesized that children in the PCIT-T group would demonstrate significantly greater pre-post improvements in externalizing problems than those in the COS-P and control groups as reported by caregivers on the CBCL. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for the CBCL Externalizing Problems score. Effect sizes were reported as partial eta squared ($\eta^2_p$).

Hypothesis 8: Child Negative Verbalizations. It was hypothesized that children in the PCIT-T group would demonstrate significantly greater pre-post improvements in negative verbalizations than those in either the COS-P or control group as coded using the DPICS. To test this hypothesis, a repeated measures ANOVA was conducted with group allocation as the between-subjects factor and assessment time point as the within-subjects factor for the composite code comprising Child Negative Talk, Child Yell, and Child Whine codes. Effect sizes were reported as partial eta squared ($\eta^2_p$).

Exploratory Questions

Bivariate correlations were used to explore associations between child and caregiver emotion regulation and related variables at baseline. Attrition analyses were conducted to explore group differences in attrition, as well as differences between treatment completers and non-completers on demographic characteristics and baseline variables.

Power Analysis

G*Power 3.1.9.4 (Faul et al., 2007) was used to conduct an a priori power analysis. In order to achieve statistical power of .80 and an estimated effect size of partial $\eta^2 = .04$, a total sample size of 51 dyads is needed for a repeated-measures, within-between ANOVA with 3 independent groups and outcome variables measured at 2 timepoints. Because the ANOVA were
run with only treatment completers, the study sample ranged from 43 to 50 dyads, which was lower than the number of participants needed to achieve adequate statistical power.

**Results**

**Preliminary Analyses**

**Missing Data**

IBM SPSS Statistics 28 was used to conduct all analyses. Data were examined for missing values. Many of the demographic variables were identified as having more than 5% of data missing: 23.7% of data missing for mothers’ ethnicity, 21.1% for number of people in the household, 14.5% for mothers’ education level, 14.5% for mothers’ occupation, 13.2% for marital status, and 11.8% for family income level.

Missingness was deemed problematic for outcome variables with over 5% of data missing (Schafer, 1999). For the DERS and CBCL, profiles with more than 7 or 8 items (respectively) missing were excluded per scoring instructions (Achenbach & Rescorla, 2000; Gratz & Roemer, 2004). Though no DERS profile fell into this category, two CBCL profiles were missing more than 8 items and were thus excluded from analyses. Of the valid DERS and CBCL profiles, missing items within a subscale were replaced by the mean of non-missing items in that subscale.

Little’s MCAR test was run to determine if the data used for each analysis were missing completely at random. Results of Little’s MCAR test for each pre-post outcome variable can be found in Table 1. Little’s MCAR test indicated that data were missing completely at random. Due to the data being MCAR and the small number of missing values, listwise deletion was deemed appropriate (Little, 1988) and used to handle missingness.

**Randomization Checks**
One-way ANOVA and chi-square tests were conducted to check for successful randomization. Eta squared ($\eta^2$) was used as an estimate of effect size for one-way ANOVA and interpreted as $\eta^2 = .01$ (small), $\eta^2 = .06$ (medium), and $\eta^2 = .138$ (large; Pallant, 2020). To check test assumptions for one-way ANOVA, data were examined for significant outliers, normality, and homogeneity of variances. The following pre-treatment variables had significant outliers (|Z| > 3.3): DERS, Parent Negative Talk, Parent Emotion Dismissing, and Child Emotional Distress. Analyses were run with and without outliers to determine if the presence of outliers altered the results. Outliers did not change the results and were thus included in the sample for these analyses. Skewness and kurtosis were examined to assess for problems with normality. The following pre-treatment variables had leptokurtic distributions (i.e., |κ| > 3.2): Parent Negative Talk, Parent Emotion Dismissing, and Child Emotional Distress. Logarithmic transformations of these variables did not correct their non-normal distributions; thus, the Kruskal-Wallis H test was conducted for these variables. Levene’s test was used to assess the equality of variances. Levene’s test was significant for pre-treatment Child Negative Verbalizations, which suggested that the homogeneity of variances assumption was violated. Therefore, Welch ANOVA was used instead of a one-way ANOVA for this variable. Means, standard deviations, results of the one-way ANOVA, and effect sizes for continuous demographic variables are reported in Table 2. The results of the chi-square tests for categorical demographic variables are reported in Table 3. The means, standard deviations, results of the one-way ANOVA or Kruskal-Wallis H tests, and effect sizes for pre-treatment outcome variables are reported in Table 4. There were no significant differences between the PCIT-T, COS-P, and waitlist control groups on any of the demographic and outcome variables measured at baseline, suggesting that random assignment was successful. Thus, no demographic or baseline variables were included in the main analyses as covariates.
Univariate and Bivariate Analyses

Descriptive statistics for pre-treatment and post-treatment variables stratified by group allocation are reported in Tables 5 and 6 for treatment completers only. For Parent Negative Talk and Parent Emotion Dismissing, a ratio (e.g., total number of Parent Negative Talk to total number of parent verbalizations) was reported in addition to a raw count in order to account for differences in verbalization frequency during the observational tasks. Correlations (i.e., Pearson’s, point-biserial, Cramer’s V) for select demographic and pre-treatment outcome variables were calculated and are reported in Table 7. To check test assumptions for Pearson’s and point-biserial correlations, data were examined for outliers, linearity, normality, and homogeneity of variance. Nonparametric alternatives (i.e., Spearman rank-order correlation and Kendall’s Tau) were utilized when data violated test assumptions. Pearson’s and point-biserial correlations between 0 to ±.2 were interpreted as negligible, ±.2 to ±.4 as low, ±.4 to ±.7 as substantial, and ±.7 to ±1.0 as high (Hemphill, 2003). Spearman’s rank-order correlations between .0 to ±.20 were interpreted as negligible, ±.21 to ±.40 as weak, ±.41 to ±.60 as moderate, ±.61 to ±.80 as strong, and ±.70 to ±1.0 as very strong (Prion & Haerling, 2014). Cramer’s V values between 0 and .1 were interpreted as very weak, .1 to .2 as weak, .2 to .3 as moderate, and greater than .3 as strong (Marchant-Shapiro, 2015).

Of note, maternal use of negative talk at baseline had a weak, significant association with maternal age, such that greater negative talk was related to younger maternal age, $r_s = -.251, p = .042$. Maternal education level was significantly related to child negative verbalizations at baseline, such that higher maternal education was associated with fewer negative child verbalizations, $\tau = -.239, p = .014$. Parent emotion regulation as reported on the DERS was significantly and positively associated with scores on the CBCL Dysregulation Profile, $r_s = .395$, ...
RCT COMPARING PCIT-T AND COS-P

$p < .001$, and CBCL Externalizing Problems, $r_s = .368, p = .002$. Observed parent negative talk at baseline had significantly positive associations with CBCL Dysregulation Profile, $r_s = .328, p = .006$, and CBCL Externalizing Problems scores, $r_s = .323, p = .007$. Surprisingly, observed parent negative talk was also significantly and positively associated with child emotion regulation maturity, such that greater parent negative talk was related to use of more mature emotion regulation strategies, $r_s = .309, p = .009$. Maternal use of emotion dismissing at baseline was positively and significantly related to the number of people in the household, such that greater emotion dismissing was associated with more family members residing in the home, $r_s = .282, p = .029$. Emotion dismissing also had significantly positive associations with child negative verbalizations, such that greater emotion dismissing during observational tasks was associated with more negative verbalizations used by children, $r_s = .331, p = .004$.

Caregiver difficulties with emotion regulation as reported on the DERS were not significantly associated with maternal use of negative talk and emotion-dismissing. Similarly, child emotion dysregulation as reported by parents on the CBCL Dysregulation Profile were not significantly associated with observed child distress, child emotion regulation maturity, and maternal support seeking.

Main Analyses

Repeated measures ANOVAs were conducted with group as the between-subjects factor and assessment time point as the within-subjects factor. Data points for which the absolute value of the studentized residual value was greater than 3 were considered outliers. ANOVAs were run with and without the presence of outliers to determine if outliers had undue influence on the test results. If results were not affected by the inclusion of outliers, outliers were included in order to increase statistical power. Due to the sample size of treatment completers ($n < 50$), the Shapiro-
RCT COMPARING PCIT-T AND COS-P

Wilk test of normality was used to determine if data within each cell of the two-way mixed ANOVA were normally distributed. The Levene’s test and Box’s M test were conducted to determine if the test assumptions of homogeneity of variance and covariance were met. The assumption of sphericity was met for all ANOVAs conducted, as only pre- and post-treatment time points were included in the analyses. Partial eta squared ($\eta^2_p$) was used as an estimate of effect size for all repeated measures ANOVAs, interpreted as $\eta^2_p = .01$ (small), $\eta^2_p = .06$ (medium), and $\eta^2_p = .14$ (large; Cohen, 1992). Bonferroni corrections were used to correct for multiple comparisons.

**Caregiver Emotion Regulation**

**Assumptions Testing.** There were no outliers in any of the three groups. The Shapiro-Wilk test of normality indicated that data were normally distributed for both pre- and post-treatment variables. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. Although the a priori power analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical power of .80, the sample size for this analysis was 48 ($n_{PCIT-T} = 18, n_{COS-P} = 11, n_{WLC} = 19$). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.

**Results.** A repeated measures ANOVA was conducted to compare the main effects of group and time and their interaction effect on DERS Total scores. There was no significant main effect of time, $F(1, 45) = 0.04, p = .842, \eta^2_p = .001$, nor was there a significant group-by-time interaction effect, $F(2, 45) = 1.71, p = .192, \eta^2_p = .071$. However, the effect size for the interaction effect was medium in size. There was a significant and large main effect of group, $F(2, 45) = 3.97, p = .026, \eta^2_p = .150$. Means, standard deviations, pre-post difference scores, and
RCT COMPARING PCIT-T AND COS-P

effect sizes for DERS Total scores are reported in Table 5. Post hoc analysis using Tukey's test revealed that the COS-P group ($M = 85.01, SE = 5.15$) was significantly higher in their DERS Total score than the PCIT-T group ($M = 68.60, SE = 4.03$), $p = .041$, 95% CI [0.57, 32.26]. DERS Total scores did not differ significantly between either treatment group and the waitlist control group.

**Parent Negative Talk**

*Assumptions Testing.* To account for differences in the total number of parent verbalizations during the observational tasks, a ratio of parent negative talk to the total number of parent verbalizations was created and used for the analysis. There was one outlier in the pre-treatment Parent Negative Talk Ratio variable and two outliers in the post-treatment Parent Negative Talk Ratio variable, as evidenced by studentized residual values greater than $\pm 3$. The Shapiro-Wilk test indicated that data were not normally distributed within each cell of the mixed ANOVA. Thus, logarithmic transformations were conducted on both pre- and post-treatment Parent Negative Talk Ratio. However, the two variables remained non-normal in distribution after transformation. Box’s M was significant for both the original and transformed data ($p = .006$), and Levene’s test was significant for the original and transformed post-treatment Parent Negative Talk Ratio ($p = .031$). The sample size for this analysis was 50 ($n_{PCIT-T} = 18$, $n_{COS-P} = 10$, $n_{WLC} = 22$). Although mixed ANOVA are generally robust against violations of normality (Field, 2018), the assumptions of equality of variance and covariance were also violated. The violations of test assumption along with the small sample size should be considered when interpreting the results of the ANOVA for parent negative talk.

**Results.** A repeated measures ANOVA was conducted to compare the main effects of group and time and their interaction effect on Parent Negative Talk Ratio. The original variables
RCT COMPARING PCIT-T AND COS-P

were used for this analysis, as the results did not change when the transformed variables were used. The repeated measures ANOVA was run with and without outliers; the inclusion of outliers did not influence the results, so outliers were included in this analysis to increase statistical power. There was no significant main effect of group, \( F(2, 47) = 0.40, p = .671, \eta^2_p = .017 \), nor was there a significant group-by-time interaction effect, \( F(2, 47) = 1.94, p = .156, \eta^2_p = .076 \). However, the effect size for the interaction effect was medium in size. There was a significant and large main effect of time, \( F(1, 45) = 19.23, p < .001, \eta^2_p = .290 \). Post-treatment Parent Negative Talk Ratio was significantly lower than pre-treatment Parent Negative Talk Ratio \((p < .001, 95\% \text{ CI } [-.031, -.012])\). Means, standard deviations, pre-post difference scores, and effect sizes for Parent Negative Talk are reported in Table 5.

**Parent Emotion Dismissing**

**Assumptions Testing.** To account for differences in the total number of parent verbalizations during the observational tasks, a ratio of parent emotion dismissing to the total number of parent verbalizations was created and used for the analysis. There were two outliers in the pre-treatment Parent Emotion Dismissing Ratio variable and one outlier in the post-treatment Parent Emotion Dismissing Ratio variable, as evidenced by studentized residual values greater than \( \pm 3 \). The Shapiro-Wilk test indicated that data were not normally distributed within each cell of the mixed ANOVA. Thus, logarithmic transformations were conducted on both pre- and post-treatment Parent Emotion Dismissing Ratio. However, the two variables remained non-normal in distribution after transformation. Box’s M was significant for both the original and transformed data \((p < .001)\), and Levene’s test was significant for the original and transformed pre-treatment Parent Emotion Dismissing Ratio. Although the a priori power analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical
RCT COMPARING PCIT-T AND COS-P

power of .80, the sample size for this analysis was 49 ($n_{PCIT-T} = 18$, $n_{COS-P} = 10$, $n_{WLC} = 21$).

Although mixed ANOVA are generally robust against violations of normality (Field, 2018), the assumptions of equality of variance and covariance were also violated. The violations of test assumption along with the small sample size should be considered when interpreting the results of the ANOVA for parent emotion dismissing.

**Results.** A repeated measures ANOVA was conducted to compare the main effects of group and time and their interaction effect on Parent Emotion Dismissing Ratio. The original variables were used for this analysis, as the results did not change when the transformed variables were used. The repeated measures ANOVA was run with and without outliers; the inclusion of outliers did not influence the results, so outliers were included in this analysis to increase statistical power. There was a medium, but non-significant group-by-time interaction effect, $F(2, 47) = 2.19, p = .124, \eta^2_p = .087$. Additionally, there was no significant main effect of time, $F(1, 46) = 0.43, p = .516, \eta^2_p = .009$, or group, $F(2, 46) = 0.52, p = .599, \eta^2_p = .022$. Means, standard deviations, pre-post difference scores, and effect sizes for Emotion Dismissing are reported in Table 5.

**Child Distress Rating**

**Assumptions Testing.** There was one outlier in the pre-treatment Child Distress variable, as indicated by a studentized residual value of 3.49. The Shapiro-Wilk test indicated that data were not normally distributed within each cell of the mixed ANOVA. Thus, logarithmic transformations were conducted on both pre- and post-treatment Child Distress variables, and the Shapiro-Wilk test indicated that the transformed variables were approximately normally distributed. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. Although the a priori power
RCT COMPARING PCIT-T AND COS-P

analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical power of .80, the sample size for this analysis was 43 ($n_{PCIT-T} = 13$, $n_{COS-P} = 11$, $n_{WLC} = 19$). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.

Results. A repeated measures ANOVA was conducted using the transformed variables to compare the main effects of group and time and their interaction effect on Child Distress. The group-by-time interaction effect was small but not significant, $F(2, 40) = 0.28, p = .756, \eta^2 = .014$. The main effect of time was small but not significant, $F(1, 40) = 0.40, p = .533, \eta^2 = .010$. The main effect of group was also not significant, $F(2, 40) = 0.13, p = .882, \eta^2 = .006$. Means, standard deviations, pre-post difference scores, and effect sizes for Child Distress Ratings are reported in Table 6.

Child Emotion Regulation

Emotion Regulation Maturity. There were no outliers in the pre- and post-treatment Emotion Regulation Maturity Composite variables. The Shapiro-Wilk test of normality indicated that data were normally distributed within each cell of the mixed ANOVA. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. Although the a priori power analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical power of .80, the sample size for this analysis was 44 ($n_{PCIT-T} = 14$, $n_{COS-P} = 11$, $n_{WLC} = 19$). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.

A repeated measures ANOVA was conducted to compare the main effects of group and time and their interaction effect on Child Emotion Regulation Maturity. The group-by-time
RCT COMPARING PCIT-T AND COS-P

interaction effect was not significant, \( F(2, 41) = 1.36, p = .267, \eta^2_p = .062 \). The main effects of
time, \( F(1, 41) = 0.25, p = .620, \eta^2_p = .006 \), and group, \( F(2, 41) = 0.59, p = .557, \eta^2_p = .028 \), were
not significant. The size of the main effect of group was small, and the size of the group-by-time
interaction effect was medium. Means, standard deviations, pre-post difference scores, and effect
sizes for Emotion Regulation Maturity composite scores are reported in Table 6.

**Maternal Support Seeking.** There were no outliers in the pre- and post-treatment
Maternal Support Seeking variables. The Shapiro-Wilk test indicated that data were not normally
distributed within each cell of the mixed ANOVA. Thus, logarithmic transformations were
conducted on both variables. The Shapiro-Wilk test indicated that the data were approximately
normally distributed after transformation. Levene’s test and Box’s M test were not significant,
indicating that the assumptions of homogeneity of variance and covariance, respectively, were
met. Although the a priori power analysis for this repeated measures ANOVA indicated that a
sample size of 51 would be required to achieve statistical power of .80, the sample size for this
analysis was 44 (\( n_{PCIT-T} = 14, n_{COS-P} = 11, n_{WLC} = 19 \)). Thus, the results of the ANOVA should be
interpreted with caution, as smaller sample size is associated with an increased risk of Type II
error.

A repeated measures ANOVA was conducted using the transformed variables to compare
the main effects of group and time and their interaction effect on Maternal Support Seeking. The
group-by-time interaction effect was small but not significant, \( F(2, 41) = 0.37, p = .691, \eta^2_p = .018 \).
The main effect of group was also small but not significant, \( F(2, 41) = 0.46, p = .637, \eta^2_p = .022 \).
However, the main effect of time was significant and large, \( F(1, 41) = 14.03, p < .001, \eta^2_p = .255 \), with significantly more maternal support seeking at post-treatment than at pre-treatment.
RCT COMPARING PCIT-T AND COS-P

Means, standard deviations, pre-post difference scores, and effect sizes for Maternal Support Seeking are reported in Table 6.

**Child Dysregulation**

**Assumptions Testing.** Examination of the studentized residual values revealed that there were no outliers in the pre- and post-treatment CBCL Dysregulation Profile variables. The Shapiro-Wilk test indicated that data were normally distributed within each cell of the mixed ANOVA. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. Although the a priori power analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical power of .80, the sample size for this analysis was 46 ($n_{PCIT-T} = 18$, $n_{COS-P} = 10$, $n_{WLC} = 18$). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.

**Results.** A repeated measures ANOVA was conducted to compare the main effects of group and time and their interaction effect on parent-reported child dysregulation. The group-by-time interaction effect was not significant, $F(2, 43) = 0.10, p = .905, \eta^2_p = .005$. However, both main effects of time, $F(1, 43) = 19.59, p < .001, \eta^2_p = .313$, and group, $F(2, 43) = 3.66, p = .034, \eta^2_p = .146$, were statistically significant and large in magnitude. CBCL Dysregulation Profile scores were significantly lower at post-treatment than at pre-treatment ($p < .001, 95\% CI [-8.42, -3.15]$). Tukey’s post-hoc test revealed that CBCL Dysregulation Profile scores were significantly lower for the PCIT-T group than for the COS-P group ($p = .042, 95\% CI [-15.78, -0.23]$), though neither treatment group significantly differed from the waitlist control group.

Means, standard deviations, pre-post difference scores, and effect sizes for CBCL Dysregulation Profile scores are reported in Table 6.
ANOVA with Children in Valid Age Range. Because the CBCL 1.5-5 has only been validated for children between 18-35 months, a repeated measures ANOVA was conducted for CBCL Dysregulation Profile with the subsample of children who were 18 months of age or older at the time of the baseline assessment. The Shapiro-Wilk test indicated that data were approximately normally distributed within each cell of the mixed ANOVA. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. The sample size for this analysis was 28 (\(n_{PCIT-T} = 11\), \(n_{COS-P} = 4\), \(n_{WLC} = 13\)). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error. For this subsample, only the main effect of time was statistically significant and large in magnitude, \(F(1, 25) = 8.37, p = .008, \eta^2 = .251\). CBCL Dysregulation Profile scores were significantly lower at post-treatment than at pre-treatment \((p = .008, 95\% \text{ CI} [-9.39, -1.58])\). The group-by-time interaction effect was not significant, \(F(2, 25) = 0.51, p = .607, \eta^2 = .039\), nor was the main effect of group, \(F(2, 25) = 1.57, p = .229, \eta^2 = .111\).

Child Externalizing Problems

Assumptions Testing. Examination of the studentized residual values revealed that there were no outliers in the pre- and post-treatment CBCL Externalizing Problems variables. The Shapiro-Wilk test indicated that data were approximately normally distributed within each cell of the mixed ANOVA. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. Although the a priori power analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical power of .80, the sample size for this analysis was 46
RCT COMPARING PCIT-T AND COS-P

\( (n_{PCIT-T} = 18, n_{COS-P} = 10, n_{WLC} = 18) \). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.

**Results.** A repeated measures ANOVA was conducted to compare the main effects of group and time and their interaction effect on parent-reported child externalizing problems. The group-by-time interaction effect was not significant, \( F(2, 43) = 0.07, p = .936, \eta^2_p = .003 \). However, both main effects of time, \( F(1, 43) = 17.53, p < .001, \eta^2 = .290 \), and group, \( F(2, 43) = 3.68, p = .033, \eta^2_p = .146 \), were statistically significant and large in magnitude. CBCL Externalizing Problems scores were significantly lower at post-treatment than at pre-treatment (\( p < .001, 95\% \text{ CI } [-7.23, -2.53] \)). Tukey’s post-hoc test revealed that CBCL Externalizing Problems scores were significantly lower for the PCIT-T group than for the COS-P group (\( p = .042, 95\% \text{ CI } [-14.33, -0.22] \)), though neither treatment group significantly differed from the waitlist control group. Means, standard deviations, pre-post difference scores, and effect sizes for CBCL Externalizing Problems scores are reported in Table 6.

**ANOVA with Children in Valid Age Range.** Because the CBCL 1.5-5 has only been validated for children between 18-35 months, a repeated measures ANOVA was conducted for CBCL Externalizing Problems with the subsample of children who were 18 months of age or older at the time of the baseline assessment. The Shapiro-Wilk test indicated that data were approximately normally distributed within each cell of the mixed ANOVA. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. The sample size for this analysis was 28 (\( n_{PCIT-T} = 11, n_{COS-P} = 4, n_{WLC} = 13 \)). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.
RCT COMPARING PCIT-T AND COS-P

For this subsample, only the main effect of time was statistically significant and large in magnitude, $F(1, 25) = 7.43, p = .012, \eta^2_p = .229$. CBCL Externalizing Problems scores were significantly lower at post-treatment than at pre-treatment ($p = .012, 95\% \text{ CI } [-7.99, -1.11]$). The group-by-time interaction effect was small but not significant, $F(2, 25) = 0.54, p = .590, \eta^2_p = .041$. The main effect of group was medium in size but not significant, $F(2, 25) = 1.46, p = .252, \eta^2_p = .105$.

Child Negative Verbalizations

Assumptions Testing. There was one outlier in the post-treatment Child Negative Verbalizations variable with a studentized residual value of 4.68. The Shapiro-Wilk test indicated that data were not normally distributed within each cell of the mixed ANOVA. Thus, logarithmic transformations were conducted on both variables, and the Shapiro-Wilk test indicated that data were approximately normally distributed after transformation. Levene’s test and Box’s M test were not significant, indicating that the assumptions of homogeneity of variance and covariance, respectively, were met. Although the a priori power analysis for this repeated measures ANOVA indicated that a sample size of 51 would be required to achieve statistical power of .80, the sample size for this analysis was 49 ($n_{PCIT-T} = 17, n_{COS-P} = 10, n_{WLC} = 21$). Thus, the results of the ANOVA should be interpreted with caution, as smaller sample size is associated with an increased risk of Type II error.

Results. A repeated measures ANOVA was conducted using transformed variables to compare the main effects of group and time and their interaction effect on Child Negative Verbalizations. The analysis was run with and without the outlier, and it was determined that the presence of the outlier had undue influence on the results. Thus, the case with the outlier was removed from the analysis. The group-by-time interaction effect was not significant but medium
RCT COMPARING PCIT-T AND COS-P

in magnitude, $F(2, 45) = 2.67, p = .080, \eta^2_p = .106$. Neither main effect of time, $F(1, 45) = 0.95, p = .336, \eta^2_p = .021$, nor group, $F(2, 45) = 0.61, p = .548, \eta^2_p = .026$, was statistically significant, and effect sizes for both were small. Means, standard deviations, pre-post difference scores, and effect sizes for Child Negative Verbalizations are reported in Table 6.

**Attrition Analyses**

Of the 76 mother-child dyads, 51 completed a Time 2 assessment. In other words, approximately 33% ($n = 25$) of the sample dropped out of the study between Time 1 and Time 2 assessments. Table 8 shows the percentage of attrition by group. The term “treatment completers” refers to participants who completed both Time 1 and Time 2 assessments, including those who completed treatment (as part of the PCIT-T and COS-P groups) and those who were waiting to receive treatment (as part of the waitlist control group). The COS-P group had the highest attrition percentage (59.26%), followed by the PCIT-T group (25%) and waitlist control group (12%). In order to assess for a significant association between group and attrition status (i.e., treatment completers and dropouts), a chi-square test of association was conducted between the categorical variables. All expected cell frequencies were greater than five. There was a statistically significant and moderately strong association between group and attrition status, $\chi^2(2) = 11.36, p = .003$, Cramer’s $V = .387, p = .003$. A clustered bar graph with number of participants on the y-axis and group membership on the x-axis can be found in Figure 4.

Further, paired samples $t$-tests and chi-square tests were conducted to determine if there were significant differences in baseline variables between treatment completers ($n = 51$) and dropouts ($n = 25$). Paired sample $t$-tests indicated that there were no significant differences between treatment completers and dropouts in child age, maternal age, number of people in the household, and all outcome variables at baseline. Results of the paired samples $t$-tests are
RCT COMPARING PCIT-T AND COS-P

reported in Table 9. Chi-square tests indicated that there were no significant associations between attrition status and the following categorical variables: child sex, maternal occupational category, maternal education, marital status, and family income level. Results of the chi-square tests are reported in Table 10.

Discussion

Summary of Findings

The present study was a randomized controlled trial investigating the efficacy of Parent-Child Interaction Therapy for Toddlers compared to Circle of Security-Parenting and waitlist controls on parent and child emotion regulation. Both observational coding and parent-report measures were used to examine changes in parent and child emotion regulation and related variables over an 8-week period. The hypotheses that both the PCIT-T and COS-P treatment groups would demonstrate significantly greater improvements than the waitlist control group in (1) parent-reported caregiver emotion regulation, (2) observed emotion dismissing, (3) observed child emotion regulation, and (4) parent-reported child dysregulation were not supported. Furthermore, the hypotheses that the PCIT-T group would demonstrate significantly greater reductions than the COS-P and WLC groups in (5) observed caregiver negative talk, (6) parent-reported child externalizing behavior, and (7) observed child negative talk were not supported.

Nevertheless, there were significant reductions in observed parent negative talk, parent-reported child dysregulation, and parent-reported child externalizing problems from pre-treatment to post-treatment across all groups, as well as significantly more maternal support seeking observed from pre- to post-treatment. Additionally, the PCIT-T group had significantly lower scores on the DERS, CBCL Dysregulation Profile, and CBCL Externalizing Problems than the COS-P group across the two timepoints. However, scores on these measures did not
RCT COMPARING PCIT-T AND COS-P

significantly differ between the treatment groups and the waitlist control group. These findings were inconsistent with previous studies that demonstrated significantly greater behavioral improvements for children receiving standard PCIT than those in a waitlist control group (e.g., Leung et al., 2009; McNeil et al., 1999; Mersky et al., 2016; Nixon et al., 2004; Thomas & Zimmer-Gembeck, 2012; Schuhmann et al., 1998). The null findings of the current study point to issues with power due to participant attrition that may have made it difficult to detect effects among a sample of 43-50 caregiver-child dyads; study limitations are addressed in subsequent sections.

Despite the lack of significant interaction effects, effect size calculations revealed medium to large standardized mean differences for parent negative talk (Hedges’ \( g = 0.87 \)), child externalizing behavior (Hedges’ \( g = 0.81 \)), child dysregulation (Hedges’ \( g = 0.78 \)), and maternal support seeking (Hedges’ \( g = 0.65 \)) for the PCIT-T group. These effect sizes are considerable for a treatment outcome study, but they are lower than what is typically found in PCIT effectiveness studies (see Thomas et al., 2017 for a meta-analysis of PCIT). Although the PCIT-T group did not significantly outperform the COS-P and WLC groups on these measures, the large effect sizes suggest that PCIT-T may show promise as an efficacious intervention for toddlers with emotional and behavioral issues. However, more research is needed to establish it as such.

**Caregiver Emotion Regulation**

The hypothesis that mothers in the two treatment groups would demonstrate significantly greater pre-post improvements in emotion regulation than those in the control group was not supported. Effect size calculations indicated that the magnitude of pre-post differences was small for both the waitlist control group (Hedges’ \( g = 0.46 \)) and the COS-P group (Hedges’ \( g = 0.23 \)). In comparison, the PCIT-T group had a much smaller effect size (Hedges’ \( g = 0.07 \)), which was
RCT COMPARING PCIT-T AND COS-P

inconsistent with previous findings of significantly lower DERS scores with small \((d = 0.25; \text{Zimmer-Gembeck et al., 2019})\) to medium effect sizes \((d = 0.78; \text{Lieneman et al., 2020})\) following the completion of standard PCIT.

The literature on emotion regulation in parenthood may point to possible explanations for these unexpected findings. There is evidence to suggest that emotion regulation abilities evolve across the lifespan (Bariola et al., 2012), particularly as adults assume the role of parents and are faced with the recurring challenge of regulating their own emotions in the midst of their child’s emotion dysregulation (Rutherford et al., 2015). Although the current study was similar to Lieneman et al. (2020) and Zimmer-Gembeck et al. (2019) in that it used a clinically referred sample, it differed in that the children were younger (i.e., \(M_{\text{age}} = 19.13\) months, \(SD = 2.98\) compared to \(M_{\text{age}} = 3.76\) years, \(SD = 1.14\) in Lieneman et al., 2020, and \(M_{\text{age}} = 53.3\) months, \(SD = 13.4\) in Zimmer-Gembeck et al., 2019). Thus, it is possible that emotion regulation difficulties found in parents of older children are closely related to their child’s disruptive behaviors and are markedly reduced with improvements in child behavior after PCIT. Caregivers of children referred for standard PCIT are typically entrenched in a coercive cycle with their child that has stabilized over time and guides future caregiver-child interactions (Granic & Patterson, 2006; Scaramella & Leve, 2004). In contrast, caregivers of toddler-aged children may perceive behavior problems as developmentally appropriate difficulties with emotion or behavior regulation rather than negative attention-seeking in nature. Therefore, emotion regulation difficulties in parents of toddlers may not be as closely related to child behavior problems as may be the case for parents of older children, making emotion regulation for parents of toddlers more impervious to interventions targeting attachment issues (i.e., COS-P) or behavior problems (PCIT-T).
In further distinguishing the current study from those of Lieneman et al. (2020) and Zimmer-Gembeck et al. (2019), it is worth noting that the pre-treatment mean for the PCIT-T group were lower in the current study ($M_{pre} = 69.03$) than in the other two studies ($M_{pre} = 88.56$ and $71.7$, respectively). Therefore, participants who received PCIT in these studies had greater room for improvement than did participants who received PCIT-T in the current study. Additionally, those in the study by Zimmer-Gembeck et al. (2019) had a longer course of treatment (i.e., approximately 14 weeks) than did those in the current study (i.e., 8 weeks), though the total number of intervention hours were comparable to one another. It is possible, then, that caregiver emotion regulation may require more time to demonstrate improvements and may benefit from follow-up assessment.

Because the focus of the DERS was on caregivers’ own emotional responses, awareness, and regulation strategies, future research in this area may benefit from incorporating measures that tap into emotion socialization practices and attitudes, attributions, or beliefs about their child’s emotions. For example, the DERS subscales measure a respondent’s difficulty in controlling impulses, engaging in goal-directed behavior, and accepting one’s own emotional responses. Because the DERS is intended for an audience wider than caregivers alone, the items do not specifically target emotion regulation difficulties that are unique to caregivers and caregiving situations. Neither PCIT-T nor COS-P yielded improvements in caregiver emotion regulation; however, it is possible that caregivers in these groups did acquire skills to regulate their behavior in the face of toddler emotion dysregulation. In PCIT-T, caregivers are coached to soothe and comfort their toddler when distressed and respond with consistency and calm when met with child noncompliance to parental directives. In COS-P, caregivers are taught to increase their reflective functioning so that they may be more attune to their children’s emotional needs.
RCT COMPARING PCIT-T AND COS-P

Improvements in these areas may not have been reflected in the DERS but may be observed in future studies with the use of caregiver-report measures of other emotion-related constructs.

**Caregiver Negative Talk**

Because decreasing parent negative talk is a specific target of PCIT-T, it was hypothesized that caregivers in the PCIT-T group would have a significantly greater pre-post reduction in negative talk than those in the COS-P and waitlist control groups. However, this hypothesis was not supported. Instead, parental negative talk was found to be significantly lower at post-treatment than at pre-treatment for the total sample of treatment completers, though these results must be interpreted with caution due to the violations of assumptions for the repeated measures ANOVA. Effect size calculations indicated that the magnitude of the pre-post treatment effect was large for the PCIT-T group (Hedges’ $g = 0.87$) but small for the COS-P group (Hedges’ $g = 0.23$). Unexpectedly, the waitlist control group had a pre-post mean difference and effect size (Hedges’ $g = 0.81$) that were comparable to that found in the PCIT-T group. Although this was likely spurious and not an indication of a meaningful change in negative talk for those in the waitlist control group, it highlights the importance of including a control group for treatment outcome studies in order to distinguish clinically significant, treatment-related improvements from changes unrelated to the intervention.

Correlation analyses revealed that observed parent negative talk at baseline was positively and significantly correlated with parent-reported child externalizing behavior ($r_s = .323, p < .001$), which was similar to what Quetsch and colleagues (2018) found in their investigation of differences in emotion regulation between clinical and nonclinical samples. Contrary to their findings, however, parent negative talk at baseline was not significantly associated with parent-reported emotion regulation on the DERS in the current study. It is
RCT COMPARING PCIT-T AND COS-P

possible that the age difference between the current sample ($M = 19.13$ months, $SD = 2.98$) and the sample in Quetsch et al. (2018; $M = 4.62$ years, $SD = 1.69$) accounted for the differential findings. Use of negative talk in caregivers of older children may be more directly related to child behavioral problems (e.g., oppositionality, noncompliance) and the caregiver-child coercive cycle that is often characterized by negative verbal interactions. On the other hand, parental negative talk for toddlers may not be used with the same intentions as it is for older age ranges and may focus more on teaching safe and appropriate behaviors than stopping disruptive, negative attention-seeking behaviors. Further analysis into the content of parental negative statements for toddlers may provide more context for these findings.

**Caregiver Emotion Dismissing**

The hypothesis that caregivers in both the PCIT-T and COS-P groups would demonstrate significantly greater pre-post reductions in emotion dismissing statements than caregivers in the control group was not supported. Effect sizes were calculated for pre-post differences, and only the waitlist control group had a positive pre-post difference with a small effect size (Hedges’ $g = 0.40$). These findings suggest that caregiver and child familiarity with the observational tasks intended to evoke negative affect (e.g., Toy Removal, Clean-Up) may have contributed to less parent negative talk (as reported above) and fewer emotion dismissing statements at Time 2 that were unrelated to the intervention. Further, emotion dismissing was a relatively low-incidence behavior at baseline (raw numbers: $M_{\text{PCIT-T}} = 0.72$, $M_{\text{COS-P}} = 1.10$, $M_{\text{WLC}} = 2.33$). Thus, there may have been a floor effect at play that made it difficult to observe any meaningful reductions across timepoints.

There is evidence to suggest that parent emotion dismissing is a risk factor for emotion regulation difficulties and behavior problems in children (Lunkenheimer et al., 2007). A glimpse
RCT COMPARING PCIT-T AND COS-P

of this was observed in the statistically significant, positive correlation between maternal emotion dismissing and child negative verbalizations (i.e., criticism, whining, and yelling) in the current sample ($r_s = .331, p < .001$). Thus, reducing parent emotion dismissing is certainly a worthy target of intervention. Nevertheless, the null findings of the current study indicate that interventions may only be effective in reducing emotion dismissing when it is included as a specific goal of the intervention. Because neither PCIT-T nor COS-P specifically aim to reduce emotion dismissing practices, it may have been unrealistic to expect significant improvements in this parental behavior. Parenting interventions that seek to reduce parent emotion dismissing may need to teach the use of emotion coaching statements and other constructive emotion socialization practices, which are associated with greater maturity of emotion regulation strategies in children (Thompson et al. 2013). Further, emotion dismissing statements may reflect parental beliefs about their child’s emotions as inappropriate or unacceptable (Gottman et al., 1997). Therefore, interventions may need to target parental attitudes and maladaptive beliefs about their child’s emotional negativity or lability in order to effect clinically significant behavioral change.

**Child Distress**

Contrary to the hypothesis, there were no significant interaction effect or main effects for observed child distress during a frustration task. Effect sizes for the pre-post difference were small for both the COS-P (Hedges’ $g = 0.21$) and WLC groups (Hedges’ $g = -0.25$) and even smaller for the PCIT-T group (Hedges’ $g = 0.11$). Additionally, observer ratings of child distress were not significantly correlated with parent-reported dysregulation on the CBCL. These findings call into question the content validity of this task as a measure of child distress. Time 1 and Time 2 means for all groups are low (i.e., < 22) considering the maximum rating a child
could have received during the 5-minute observation was 120. It is possible that toddlers in this sample did not find the toy attractive or find the situation particularly distressing, making the distress ratings during this task a poor indicator of the intensity and frequency of child distress. The current study could have benefited from checks to ensure that the toy was indeed of interest to the child (e.g., by giving the child a choice of novel toys to choose from); this may have increased the validity of the task as a measure of distress.

Other studies have successfully used the Attractive Toy in Transparent Box task, yet many have done so with older children, namely preschoolers and kindergarteners (e.g., Howse et al., 2003; Jahromi et al., 2012; Ntourou et al., 2013; Tan & Smith, 2018; Wu et al., 2020; Yan et al., 2021). Relatedly, the standardized ATTB task included in the Laboratory Temperament Assessment Battery (Lab-TAB; Goldsmith et al., 1999) is only found in the preschool and middle childhood versions. The toddler version of the Lab-TAB uses a different frustration-inducing episode, which involves the caregiver using a gentle physical restraint of the toddler after the child is shown a novel and interesting toy. This difference in task between the toddler and preschool versions of the Lab-TAB suggest that the ATTB task may not be developmentally appropriate for eliciting distress in the toddler age range.

**Child Emotion Regulation Strategies**

The hypothesis that children in both treatment groups would demonstrate significantly greater pre-post improvements in the maturity of emotion regulation strategies used during a frustration task was not supported. These null findings reflect the previously mentioned issues with the child distress variable, which are relevant to this variable in that emotion regulation strategies were only coded in the absence of child distress. The flaw with this method is that it assumes that children were distressed by the ATTB episode, which may not have been the case.
RCT COMPARING PCIT-T AND COS-P

as is evidenced by the low mean distress ratings across groups and timepoints. Therefore, a 10-second interval that lacked signs of toddler distress may not have been indicative of an emotion regulation strategy but rather of disinterest in the toy, curiosity about other objects in the room, or a desire to engage with the mother independent of the ATTB task.

Regarding extrinsic emotion regulation strategies, maternal support seeking did demonstrate large ($\eta_r = .255$) and significant increases from Time 1 to Time 2 across the entire sample of treatment completers. Effect size calculations revealed a medium pre-post mean difference for the PCIT-T group (Hedges’ $g = 0.65$) and a small pre-post mean difference for the COS-P group (Hedges’ $g = 0.35$). Unexpectedly, the effect size found in the WLC group mirrored that of the PCIT-T group (Hedges’ $g = 0.53$). One possible explanation for this result is that children were more inclined to seek maternal assistance during the ATTB task in their Time 2 encounter with the locked transparent box after confirming in the Time 1 encounter that the toy was inaccessible. Thus, the increase in maternal support seeking from Time 1 to Time 2 may be more reflective of learning that occurred during the baseline assessment than of an increase in extrinsic regulatory strategies.

**Parent-Reported Child Dysregulation**

Contrary to hypothesis, children in the PCIT-T and COS-P groups did not demonstrate significantly greater pre-post improvements in child dysregulation than did children in the waitlist control group. Nevertheless, there was a statistically significant improvement in child dysregulation from Time 1 to Time 2 across the entire sample. Effect sizes were larger in magnitude for the PCIT-T (Hedges’ $g = 0.78$) than the COS-P group (Hedges’ $g = 0.49$), suggesting that PCIT-T shows promise as an intervention that may help improve emotional and behavioral dysregulation in toddlerhood. An unexpected finding was the effect size for the pre-
RCT COMPARING PCIT-T AND COS-P

post difference in the waitlist control group (Hedges’ $g = 0.73$), which was higher than that of the COS-P group and similar to that of the PCIT-T group.

It is puzzling as to why the families in the control group showed a standardized mean difference for parent-reported dysregulation that was comparable to that of families whose caregivers were involved in PCIT-T, an intensive 8-week treatment. The results may point to the unique characteristics of the toddler age range that may distinguish it from older children. Toddlerhood represents a developmental shift from reliance on external sources of emotion regulation to greater use of intrinsic emotion regulation strategies (Kopp, 1989). Due to the rapid developmental changes in toddlerhood, it is possible that emotional and behavioral problems are normative and that fluctuations in emotion dysregulation are to be expected. Interventions targeting emotional and behavioral difficulties in toddlers, then, are largely preventative in nature. This is fundamentally different than children who are referred for standard PCIT whose behavior problems are outside of normal limits and independent of developmental issues. Interventions during this age range are not preventative but rather in direct response to clinically significant behavioral issues. Thus, studies assessing prevention efforts may not yield the drastic changes often seen in studies assessing intervention efforts, as toddler dysregulation may not be as indicative of clinically significant problems than it is for older children. Future studies examining the efficacy of PCIT-T may benefit from using measures that are intended for the toddler population and sensitive to clinically elevated emotional and behavioral issues.

**Parent-Reported Child Externalizing Problems**

Although the hypothesis that the PCIT-T group would outperform the other two groups on pre-post improvements in child externalizing behavior was not supported, there was a significant improvement from Time 1 to Time 2 across the entire sample. The effect was large in
RCT COMPARING PCIT-T AND COS-P

magnitude for the PCIT-T (Hedges’ $g = 0.81$), whereas the size of the effects was small for the COS-P group (Hedges’ $g = 0.46$) and medium for the waitlist control group (Hedges’ $g = .65$). Although this pre-post change was not significantly greater than the change seen in the COS-P and WLC groups, the large effect size suggests a statistically significant difference may have been detected with a larger sample size and more power. Thus, these findings indicate that PCIT-T warrants further study to determine its efficacy as a treatment for behavior problems in toddlers.

The effect size for child externalizing behavior in the PCIT-T group was similar to those found in several studies of standard PCIT using the same CBCL subscale as an outcome measure (e.g., $d = 0.71$ in Danko et al., 2016 and $d = 0.79$ in McCabe & Yeh, 2009). A meta-analysis of 11 PCIT studies estimated an even larger mean effect size of 1.09 for CBCL Externalizing Problems (Cooley et al., 2014). It is worth noting here that the mean CBCL Externalizing Problems scores for all three groups were low at baseline, which may seem counterintuitive given that all participants had to endorse concerns about or difficulty managing their child’s behavior in order to be eligible for participation. Having more information about the reasons for referral would have been helpful, since it is possible that parents had concerns about their toddlers’ behavior that were unrelated to externalizing problems and instead related to developmental issues (e.g., eating, sleep, toileting). Thus, the low pre-treatment scores provide further evidence that the study sample may not have had clinically elevated problems and that the interventions administered were much more preventative in nature. Thus, the larger effect sizes observed in previous PCIT studies are to be expected given that they involve children with clinically significant externalizing problems and explicitly target these behaviors throughout treatment.
RCT COMPARING PCIT-T AND COS-P

Child Negative Verbalizations

Lastly, the hypothesis that children in the PCIT-T group would demonstrate significantly greater pre-post improvements in negative verbalizations than those in either the COS-P or control group was not supported. Very few studies of PCIT have included DPICS Child Negative Talk, Whine, and Yell as outcome variables since these are not direct targets of the intervention nor are they related to the goal criteria required for graduation. However, these null findings are in line with the results from a study by Eyberg et al. (1995) in which the frequency of child negative talk did not differ from pre- to post-treatment. Correlational analyses indicated that observed child negative verbalizations in the current sample were not significantly associated with parent-reported child dysregulation or externalizing problems on the CBCL. This was inconsistent with prior studies that found significant, positive correlations between child negative verbalizations and parent-reported child disruptive behavior on the ECBI (Foote, 1999) as well as between child whining/yelling and parent-reported externalizing problems on the BASC-2 PRS (Cotter & Brestan-Knight, 2020). Again, these discrepant findings point to the developmental differences between the toddler and older child age ranges. For toddlers, whose expressive language skills are emerging, whining and yelling may be a developmentally appropriate and adaptive way to communicate distress to their caregiver. However, older children with clinically significant behavior problems may engage in whining, yelling, and negative talk for the purposes of gaining negative attention from their caregivers. Thus, child negative verbalizations in the toddler age range may not be a valid measure of emotion regulation difficulties or disruptive behavior as it might be for older children.

Attrition
Attrition is a considerable problem in child treatment outcome research, with estimates of attrition rates in child mental health interventions ranging from 16 to 72% (de Haan et al., 2013). The current study was no exception; the attrition rate in the current study was 33%, or 25 out of 76 participants. Attrition analyses indicated that dropout rate significantly differed between the three groups. The COS-P group had the highest attrition rate, with more than half of the group dropping out before the Time 2 assessment. Further, no significant differences in demographic characteristics or baseline assessments were found between treatment completers and non-completers, which suggests that attrition had little to do with these factors. Because information regarding the reasons for and time of dropout were unavailable, it is impossible to ascertain if the reasons for leaving the study were related to the treatment itself. Attrition from parenting interventions can occur for a constellation of reasons, including high levels of parental stress (Werba et al., 2006), low socioeconomic status (Kazdin & Mazurick, 1994), and clinically elevated child behavior problems (Kazdin et al., 1993). A meta-analysis of attrition from child mental health interventions named parental perception of treatment barriers and treatment relevance as two of the main factors contributing to premature termination of services (de Haan et al., 2013). Because random assignment was used in the current study, it can be assumed that many of the child-, parent-, and family-related barriers contributing to attrition were randomized across groups. Therefore, this differential attrition found across groups hints at the possibility that many left the COS-P group because they were dissatisfied with the intervention or did not find the intervention useful.

**Limitations and Future Directions**

There were a number of limitations that are important to highlight, with the biggest limitation being the high attrition rate and lack of statistical power. The a priori power analysis
RCT COMPARING PCIT-T AND COS-P

revealed that a sample size of 51 participants was needed for a medium effect size. However, given the preventative nature of these interventions and the subclinical presentation of this toddler sample, it is likely that this effect size was an overestimation. Thus, the current sample size was insufficient to achieve adequate statistical power with a smaller effect size. Without sufficient power, some actual effects may have gone undetected.

Future studies examining the efficacy of PCIT-T against another treatment arm may benefit from oversampling to account for inevitable participant attrition and ensure adequate power. Overrecruiting is particularly important when the treatment is prevention-focused and the sample is subclinical in the domains of interest, since participants are less likely to have a desire to alleviate severe child problems as a motivating factor to remain in treatment. Additionally, there is evidence supporting the use of incentives (Quetsch et al., 2020) to reduce no-show rates in PCIT, though this was examined in a sample of low-income, Latinx families. Future studies of PCIT-T may benefit from including incentives to reduce attrition and increase session attendance, particularly among caregivers who may not perceive treatment as useful or relevant to their family.

*Unexpected Improvements in Waitlist Controls*

Although the PCIT-T group appeared to fare well against COS-P in the areas of parent negative talk, child externalizing behavior, child dysregulation, and maternal support seeking behavior as evidenced by larger pre-post differences and effect sizes, the PCIT-T group was comparable to the waitlist control group in these domains. This was a bizarre and unexpected occurrence. There are several possible explanations for these findings. First, improvements observed in the waitlist control group could have been spurious rather than the result of factors pertaining to their participation in the study. Second, improvements could have also been due to
RCT COMPARING PCIT-T AND COS-P

regression to the mean. Third, it is possible that these findings were the result of possible contamination of the waitlist control group (i.e., engagement in positive parenting practices during the waiting period). More information is needed to further understand these surprising findings. The improvements among participants in the waitlist control group point to another weakness of the current study: the inclusion criteria in this study may not have been stringent enough to recruit a sample of children with emotional and behavioral problems outside normal limits. Using a cut-off score from a measure of emotional or behavioral functioning that is normed on the toddler age range as an inclusion criterion would guarantee a sample that was somewhat elevated in the domains of interest.

It is worth noting that the families in the waitlist control group had the lowest attrition rate among the three groups (12% compared to 25% in the PCIT-T group and 59.3% in the COS-P group). This low attrition rate signifies that the majority of the families in the control group wanted to pursue treatment for their child despite behavioral and parent-reported improvements. This result highlights the need for improved or additional measurement to better capture the parenting challenges and behavioral difficulties experienced by these families resulting in their desire to seek treatment.

Regarding measurement, there were several limitations that likely contributed to the null findings of the current study. The CBCL was used as a measure of child externalizing behavior and dysregulation; however, this measure is only normed for children 18 months and older. Therefore, a measure normed for toddlers (i.e., ages 12-24 months) is recommended for future studies of PCIT-T. One such measure that may be fitting for this age group is the Infant Toddler Social Emotional Assessment (ITSEA; Carter & Briggs-Gowan, 2006), which is designed to assess social emotional functioning of children ages 12 to 36 months. The ITSEA produces
scores on four domains: externalizing problems, internalizing problems, dysregulation, and social-emotional competence.

Although the combination of parent-report measures and observational measures bolstered the study design, several of the observational measures demonstrated poor validity and did not align with expected parent-report measures. In particular, the ATTB task did not appear to elicit distress among toddlers, making it difficult to gauge the frequency and intensity of distress typically exhibited by the toddlers in distressing situations. Relatedly, the emotion regulation maturity composite was problematic in that it assumed a lack of distress was indicative of engagement in emotion regulation strategies. In other words, raters were required to select an emotion regulation strategy for every 10-second interval that was void of a distress reaction, even if the child appeared to have lost interest in the task. Thus, this rating system likely inflated the use and maturity level of emotion regulation strategies. Additionally, the five-minute frustration task took place in the clinic setting, which may have had poor ecological validity and dampened the intensity of distress that toddlers typically display in naturalistic settings.

The current study focused primarily on observations of toddler emotion regulation or lack thereof, yet it did not include observations of parental response to toddler distress. Emotion regulation in this age range is largely influenced by transactional processes between the toddler and caregiver (Sameroff, 2009). Although parent negative talk and emotion dismissing statements were coded across observational tasks, there was no coding of positive, emotion socialization practices. Maternal responsiveness and warmth have been linked to greater use of parent-oriented regulatory strategies in toddlers (Calkins & Hill, 2007), and caregiver support of and responsiveness to their toddlers’ affect contributes to the use of emotion regulation strategies.
RCT COMPARING PCIT-T AND COS-P

later in life (Spinrad et al., 2004). Thus, in addition to harsh parenting practices, future studies should examine parental warmth and responsiveness to toddlers’ emotions in order to better understand the reciprocal nature of emotion regulation development in toddlers. The Dyadic Emotion Coding System (Owen et al., in preparation) may be one such measure that would allow for this type of observational coding.

Another weakness of this study is that only pre- and post-treatment timepoints were examined. However, in investigating the efficacy and effectiveness of prevention programs, follow-up measurements are needed in order to determine the short- and longer-term effects of an intervention. Future studies of PCIT-T should seek to recruit participants who are at risk for developing clinically significant emotional and behavioral problems to provide evidence for the PCIT-T’s ability to prevent future behavior issues and child psychopathology.

Lastly, future research should seek to involve fathers in addition to mothers, if applicable, in order to better understand the role of the family context in the development of emotion regulation (Morris et al., 2007). The literature on parenting interventions is largely focused on mothers as the primary caregiver. However, the scant research examining the role of fathers in child emotion regulation development suggests that fathers play an integral part in early emotional development (Hazen et al., 2010). Paternal supportiveness and responsiveness have been shown to positively influence emotional development, particularly in younger children (Cabrera et al., 2007). Additionally, differences between maternal and paternal influences on emotional development are minimal in infancy but become more distinct as the child matures (Wilson & Durbin, 2013). For these reasons, subsequent studies evaluating parenting interventions for toddler emotion regulation should include fathers whenever applicable and
RCT COMPARING PCIT-T AND COS-P

consider the broader familial and cultural context in which emotion regulation development occurs.

**Conclusion**

In sum, PCIT-T shows some promise as an efficacious intervention for improving toddler emotion dysregulation, externalizing behavior, and maternal support seeking practices, as well as harsh parenting practices (i.e., parent negative talk) that are linked to child emotion regulation problems findings. However, no definitive conclusions can be drawn from these findings about the efficacy of PCIT-T as an intervention for toddler emotion regulation due to the lack of statistical power, several problematic measures, and no significant interaction effects. Additional research is needed, particularly studies that employ sensitive assessment tools to measure constructs that have greater relevance to the relational processes that influence emotion regulation development in young children.
RCT COMPARING PCIT-T AND COS-P

References


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P

https://doi.org/10.1037/0033-2909.112.1.155


RCT COMPARING PCIT-T AND COS-P


https://doi.org/10.1111/1467-8624.00037


Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Shepard, S. A.,
RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P

*Behavior Inventory-Revised (ECBI/SESBI-R) professional manual.* Psychological Assessment Resources.


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P
2099–2112.


RCT COMPARING PCIT-T AND COS-P


[https://doi.org/10.1002/imhj.21745](https://doi.org/10.1002/imhj.21745)


[https://doi.org/10.1080/02699931.2010.544160](https://doi.org/10.1080/02699931.2010.544160)

[https://doi.org/10.1017/S0954579412001137](https://doi.org/10.1017/S0954579412001137)


RCT COMPARING PCIT-T AND COS-P


https://doi.org/10.1037/0012-1649.42.5.913


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


[https://doi.org/10.1023/b:jacp.0000026140.60558.05](https://doi.org/10.1023/b:jacp.0000026140.60558.05)


[https://doi.org/10.1017/S0954579400003096](https://doi.org/10.1017/S0954579400003096)
RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


RCT COMPARING PCIT-T AND COS-P


Whiteside-Mansell, L., Ayoub, C., McKelvey, L., Faldowski, R. A., Hart, A., Shears,


https://doi.org/10.1207/S15374424JCCP3103_11

### Table 1

**Missingness of Outcome Variables and Results of Little’s MCAR Test**

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>n (out of 76)</th>
<th>Missingness</th>
<th>$\chi^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parent</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre DERS Total</td>
<td>71</td>
<td>6.58%</td>
<td>1.63</td>
<td>.444</td>
</tr>
<tr>
<td>Post DERS Total</td>
<td>49</td>
<td>35.53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre Negative Talk</td>
<td>75</td>
<td>1.32%</td>
<td>0.34</td>
<td>.559</td>
</tr>
<tr>
<td>Post Negative Talk</td>
<td>50</td>
<td>34.21%</td>
<td>0.02</td>
<td>.990</td>
</tr>
<tr>
<td>Pre Emotion Dismissing</td>
<td>74</td>
<td>2.60%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post Emotion Dismissing</td>
<td>50</td>
<td>34.21%</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Child</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre Distress Rating</td>
<td>71</td>
<td>6.58%</td>
<td>0.57</td>
<td>.452</td>
</tr>
<tr>
<td>Post Distress Rating</td>
<td>43</td>
<td>43.42%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre ERM Composite&lt;sup&gt;a&lt;/sup&gt;</td>
<td>71</td>
<td>6.58%</td>
<td>0.66</td>
<td>.418</td>
</tr>
<tr>
<td>Post ERM Composite&lt;sup&gt;a&lt;/sup&gt;</td>
<td>44</td>
<td>42.11%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre Maternal Support Seeking</td>
<td>71</td>
<td>6.58%</td>
<td>2.85</td>
<td>.091</td>
</tr>
<tr>
<td>Post Maternal Support Seeking</td>
<td>44</td>
<td>42.11%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre CBCL Dysregulation</td>
<td>69</td>
<td>9.21%</td>
<td>5.75</td>
<td>.056</td>
</tr>
<tr>
<td>Post CBCL Dysregulation</td>
<td>49</td>
<td>35.53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre CBCL Externalizing</td>
<td>69</td>
<td>9.21%</td>
<td>4.79</td>
<td>.091</td>
</tr>
<tr>
<td>Post CBCL Externalizing</td>
<td>49</td>
<td>35.53%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre Negative Verbalizations&lt;sup&gt;b&lt;/sup&gt;</td>
<td>74</td>
<td>2.63%</td>
<td>983</td>
<td>.612</td>
</tr>
<tr>
<td>Post Negative Verbalizations&lt;sup&gt;b&lt;/sup&gt;</td>
<td>50</td>
<td>34.21%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* DERS = Difficulties in Emotion Regulation Scale, ERM = Emotion Regulation Maturity, CBCL = Child Behavior Checklist

<sup>a</sup>Child ERM Composite is a composite variable comprising weighted averages of the Situation Modification, Attentional Deployment, and Cognitive Change variables.  
<sup>b</sup>Child Negative Verbalizations is a composite variable comprising Child Negative Talk, Child Whine, and Child Yell.
Table 2

Results of One-way ANOVA for Continuous Demographic Variables

<table>
<thead>
<tr>
<th>Demographic Variable</th>
<th>PCIT-T</th>
<th>COS-P</th>
<th>Waitlist</th>
<th>F</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Age (in months)</td>
<td>18.71</td>
<td>18.81</td>
<td>19.88</td>
<td>1.19</td>
<td>.310</td>
<td>.032</td>
</tr>
<tr>
<td>Maternal Age (in years)</td>
<td>33.86</td>
<td>33.04</td>
<td>31.90</td>
<td>0.66</td>
<td>.521</td>
<td>.020</td>
</tr>
<tr>
<td>People in Household</td>
<td>3.67</td>
<td>3.90</td>
<td>3.56</td>
<td>0.54</td>
<td>.587</td>
<td>.019</td>
</tr>
</tbody>
</table>

*Note. PCIT-T = Parent-Child Interaction Therapy for Toddlers, COS-P = Circle of Security-Parenting*
Table 3

Results of Chi-Square Tests for Categorical Demographic Variables

<table>
<thead>
<tr>
<th>Demographic Variable</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Sex</td>
<td>1.06</td>
<td>.588</td>
</tr>
<tr>
<td>Maternal Occupation</td>
<td>12.69</td>
<td>.551</td>
</tr>
<tr>
<td>Maternal Education</td>
<td>4.58</td>
<td>.801</td>
</tr>
<tr>
<td>Maternal Ethnicity</td>
<td>17.10</td>
<td>.072</td>
</tr>
<tr>
<td>Marital Status</td>
<td>2.03</td>
<td>.917</td>
</tr>
<tr>
<td>Maternal Primary Language</td>
<td>22.94</td>
<td>.523</td>
</tr>
<tr>
<td>Family Income Level</td>
<td>10.82</td>
<td>.212</td>
</tr>
</tbody>
</table>
Table 4

Results of One-way ANOVA or Kruskal-Wallis H Test for Pre-treatment Outcome Variables

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>PCIT-T M</th>
<th>PCIT-T SD</th>
<th>COS-P M</th>
<th>COS-P SD</th>
<th>Waitlist M</th>
<th>Waitlist SD</th>
<th>Test Statistic</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DERS</td>
<td>71.32</td>
<td>19.72</td>
<td>76.95</td>
<td>27.51</td>
<td>80.53</td>
<td>22.68</td>
<td>0.90</td>
<td>.413</td>
<td>.026</td>
</tr>
<tr>
<td>% NTA</td>
<td>4.15</td>
<td>3.29</td>
<td>4.04</td>
<td>3.29</td>
<td>4.14</td>
<td>3.33</td>
<td>0.03</td>
<td>.986</td>
<td>.000</td>
</tr>
<tr>
<td>% Emotion Dismissing</td>
<td>0.30</td>
<td>0.76</td>
<td>0.36</td>
<td>0.67</td>
<td>0.47</td>
<td>0.91</td>
<td>2.52</td>
<td>.283</td>
<td>.066</td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distress Rating</td>
<td>19.87</td>
<td>22.77</td>
<td>21.54</td>
<td>22.72</td>
<td>15.25</td>
<td>16.99</td>
<td>0.46</td>
<td>.794</td>
<td>.095</td>
</tr>
<tr>
<td>ERM Composite</td>
<td>1.83</td>
<td>0.95</td>
<td>2.01</td>
<td>1.05</td>
<td>2.33</td>
<td>0.97</td>
<td>1.53</td>
<td>.223</td>
<td>.149</td>
</tr>
<tr>
<td>Maternal Support</td>
<td>0.33</td>
<td>0.26</td>
<td>0.25</td>
<td>0.23</td>
<td>0.31</td>
<td>0.27</td>
<td>0.51</td>
<td>.601</td>
<td>.090</td>
</tr>
<tr>
<td>Dysregulation</td>
<td>T≥61e</td>
<td></td>
<td>T≈65</td>
<td></td>
<td>T≈63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL</td>
<td>18.66</td>
<td>9.42</td>
<td>24.16</td>
<td>8.50</td>
<td>20.06</td>
<td>9.56</td>
<td>2.33</td>
<td>.106</td>
<td>.186</td>
</tr>
<tr>
<td>Externalizing</td>
<td>T≥63</td>
<td></td>
<td>T≈66</td>
<td></td>
<td>T≈64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative Verbalizations</td>
<td>9.33</td>
<td>8.31</td>
<td>12.85</td>
<td>12.98</td>
<td>10.50</td>
<td>6.44</td>
<td>0.65d</td>
<td>.526</td>
<td>.109</td>
</tr>
</tbody>
</table>

Note. PCIT-T = Parent-Child Interaction Therapy for Toddlers, COS-P = Circle of Security-Parenting, DERS = Difficulties in Emotion Regulation Scale, NTA = Negative Talk, ERM = Emotion Regulation Maturity, CBCL = Child Behavior Checklist

a Kruskal-Wallis H test conducted. b Child ERM Composite is a variable comprising weighted averages of the Situation Modification, Attentional Deployment, and Cognitive Change variables. c Child Negative Verbalizations is a composite variable comprising Child Negative Talk, Child Whine, and Child Yell. d Welch statistic reported. e Male and female T-scores were averaged to form the T-scores in this table.
### Table 5

*Descriptive Statistics and Effect Sizes for Pre- and Post-treatment Parent Variables*

*(Treatment Completers)*

<table>
<thead>
<tr>
<th></th>
<th>DERS</th>
<th>NTA Raw Count</th>
<th>NTA Ratio (%)</th>
<th>Emotion Dismissing</th>
<th>Emotion Dismissing Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PCIT-T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre $M (SD)$</td>
<td>69.03 (12.46)</td>
<td>16.89 (15.57)</td>
<td>4.30 (3.49)</td>
<td>0.72 (1.84)</td>
<td>0.18 (0.47)</td>
</tr>
<tr>
<td>Post $M (SD)$</td>
<td>68.17 (14.67)</td>
<td>3.89 (4.81)</td>
<td>0.98 (1.09)</td>
<td>0.83 (1.51)</td>
<td>0.23 (0.45)</td>
</tr>
<tr>
<td>Pre-Post $M$</td>
<td>0.87</td>
<td>13.00</td>
<td>3.31</td>
<td>-0.11</td>
<td>-0.04</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hedges’ $g$</td>
<td>.066</td>
<td>.796</td>
<td>.866</td>
<td>.051</td>
<td>.063</td>
</tr>
<tr>
<td><strong>COS-P</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre $M (SD)$</td>
<td>82.65 (20.38)</td>
<td>14.80 (8.00)</td>
<td>3.76 (2.27)</td>
<td>1.10 (1.20)</td>
<td>0.26 (0.25)</td>
</tr>
<tr>
<td>Post $M (SD)$</td>
<td>87.37 (20.98)</td>
<td>12.70 (14.35)</td>
<td>2.98 (2.99)</td>
<td>1.70 (2.36)</td>
<td>0.39 (0.50)</td>
</tr>
<tr>
<td>Pre-Post $M$</td>
<td>-4.72</td>
<td>2.10</td>
<td>0.78</td>
<td>-0.60</td>
<td>-0.13</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hedges’ $g$</td>
<td>.227</td>
<td>.127</td>
<td>.230</td>
<td>.195</td>
<td>.206</td>
</tr>
<tr>
<td><strong>Waitlist</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre $M (SD)$</td>
<td>83.82 (21.60)</td>
<td>18.21 (3.88)</td>
<td>4.16 (3.46)</td>
<td>2.33 (4.83)</td>
<td>0.53 (0.21)</td>
</tr>
<tr>
<td>Post $M (SD)$</td>
<td>78.71 (20.36)</td>
<td>7.77 (8.47)</td>
<td>1.84 (1.70)</td>
<td>0.50 (1.14)</td>
<td>0.12 (0.28)</td>
</tr>
<tr>
<td>Pre-Post $M$</td>
<td>5.11</td>
<td>8.82</td>
<td>2.32</td>
<td>1.86</td>
<td>0.41</td>
</tr>
<tr>
<td>Difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hedges’ $g$</td>
<td>.461</td>
<td>.673</td>
<td>.807</td>
<td>.360</td>
<td>.401</td>
</tr>
</tbody>
</table>

*Note.* PCIT-T = Parent-Child Interaction Therapy for Toddlers, COS-P = Circle of Security-Parenting, DERS = Difficulties in Emotion Regulation Scale, NTA = Negative Talk
### Table 6

**Descriptive Statistics and Effect Sizes for Pre- and Post-treatment Child Variables (Treatment Completers)**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Pre $M$ (SD)</th>
<th>Post $M$ (SD)</th>
<th>Pre-Post Difference</th>
<th>Hedges’ $g$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PCIT-T</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre $M$ (SD)</td>
<td>16.69 (8.08)</td>
<td>19.80 (9.64)</td>
<td>0.11</td>
<td>.806</td>
</tr>
<tr>
<td>$T$=61</td>
<td>$T$=60</td>
<td>$T$=59</td>
<td>$T$=56</td>
<td></td>
</tr>
<tr>
<td>Post $M$ (SD)</td>
<td>11.89 (7.26)</td>
<td>13.72 (8.56)</td>
<td>1.83</td>
<td>.777</td>
</tr>
<tr>
<td>$T$=59</td>
<td>$T$=56</td>
<td>$T$=59</td>
<td>$T$=56</td>
<td></td>
</tr>
<tr>
<td>Pre-Post $M$ Difference</td>
<td>4.80</td>
<td>6.08</td>
<td>3.28</td>
<td>.109</td>
</tr>
<tr>
<td>Hedges’ $g$</td>
<td>.806</td>
<td>.777</td>
<td>.109</td>
<td>.175</td>
</tr>
<tr>
<td><strong>COS-P</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre $M$ (SD)</td>
<td>24.30 (9.09)</td>
<td>27.90 (9.67)</td>
<td>3.60</td>
<td>.459</td>
</tr>
<tr>
<td>$T$=66</td>
<td>$T$=65</td>
<td>$T$=63</td>
<td>$T$=61</td>
<td></td>
</tr>
<tr>
<td>Post $M$ (SD)</td>
<td>18.84 (8.62)</td>
<td>21.64 (9.67)</td>
<td>2.80</td>
<td>.487</td>
</tr>
<tr>
<td>$T$=63</td>
<td>$T$=61</td>
<td>$T$=63</td>
<td>$T$=61</td>
<td></td>
</tr>
<tr>
<td>Pre-Post $M$ Difference</td>
<td>5.46</td>
<td>6.26</td>
<td>0.80</td>
<td>.214</td>
</tr>
<tr>
<td>Hedges’ $g$</td>
<td>.459</td>
<td>.487</td>
<td>.214</td>
<td>-.307</td>
</tr>
<tr>
<td><strong>Waitlist</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre $M$ (SD)</td>
<td>21.41 (9.19)</td>
<td>24.66 (9.39)</td>
<td>3.25</td>
<td>.646</td>
</tr>
<tr>
<td>$T$=64</td>
<td>$T$=63</td>
<td>$T$=64</td>
<td>$T$=63</td>
<td></td>
</tr>
<tr>
<td>Post $M$ (SD)</td>
<td>17.03 (7.90)</td>
<td>19.66 (8.49)</td>
<td>2.63</td>
<td>.729</td>
</tr>
<tr>
<td>$T$=64</td>
<td>$T$=63</td>
<td>$T$=64</td>
<td>$T$=63</td>
<td></td>
</tr>
<tr>
<td>Pre-Post $M$ Difference</td>
<td>4.38</td>
<td>5.00</td>
<td>-6.95</td>
<td>.249</td>
</tr>
<tr>
<td>Hedges’ $g$</td>
<td>.646</td>
<td>.729</td>
<td>.249</td>
<td>.250</td>
</tr>
</tbody>
</table>

- *Child ERM Composite is a variable comprising weighted averages of the Situation Modification, Attentional Deployment, and Cognitive Change variables. *Child Negative Verbalizations is a composite variable comprising Child Negative Talk, Child Whine, and Child Yell. *Male and female $T$-scores were averaged to form the $T$-scores in this table.
Table 7
Correlation Matrix Among Demographic and Pre-treatment Outcome Variables

<table>
<thead>
<tr>
<th></th>
<th>Correlation (Pearson’s $r$, $r_{pb}$, $r_s$, Kendall’s Tau, Cramer’s V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Sex</td>
<td>--</td>
</tr>
<tr>
<td>Child Age</td>
<td>-.068</td>
</tr>
<tr>
<td>Mo. Age</td>
<td>.057</td>
</tr>
<tr>
<td>Mo. Edu.</td>
<td>.215</td>
</tr>
<tr>
<td>Household No.</td>
<td>.090</td>
</tr>
<tr>
<td>Family Income</td>
<td>.155</td>
</tr>
<tr>
<td>DERS</td>
<td>.062</td>
</tr>
<tr>
<td>Mo. NTA</td>
<td>-.157</td>
</tr>
<tr>
<td>Mo. ED</td>
<td>-.048</td>
</tr>
<tr>
<td>Child Distress</td>
<td>.033</td>
</tr>
<tr>
<td>Child ERM</td>
<td>-.009</td>
</tr>
<tr>
<td>Maternal Support</td>
<td>-.107</td>
</tr>
<tr>
<td>CBCL Dys.</td>
<td>-.212</td>
</tr>
<tr>
<td>CBCL Ext.</td>
<td>-.237</td>
</tr>
<tr>
<td>Child Neg Verb</td>
<td>.036</td>
</tr>
</tbody>
</table>

* significant at the $p < .05$ level  
** significant at the $p < .01$ level
### Table 8

Attrition by Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Time 1 n</th>
<th>Time 2 n</th>
<th>Attrition Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCIT-T</td>
<td>24</td>
<td>18</td>
<td>25%</td>
</tr>
<tr>
<td>COS-P</td>
<td>27</td>
<td>11</td>
<td>59.26%</td>
</tr>
<tr>
<td>WLC</td>
<td>25</td>
<td>22</td>
<td>12%</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>51</td>
<td>32.89%</td>
</tr>
</tbody>
</table>

*Note. PCIT-T = Parent-Child Interaction Therapy for Toddlers, COS-P = Circle of Security-Parenting, WLC = Waitlist Control. Sample size was different depending on the variable; therefore, the largest sample size for Time 1 and Time 2 was reported in this table.*
### Table 9

Paired Samples *t*-Tests Comparing Treatment Completers and Dropouts on Baseline Variables

<table>
<thead>
<tr>
<th></th>
<th>Treatment Completers (n = 51)</th>
<th>Dropouts (n = 25)</th>
<th>t</th>
<th>p</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Age (in months)</td>
<td>18.94 2.30</td>
<td>19.45 2.97</td>
<td>-0.73</td>
<td>.479</td>
<td>-.171</td>
</tr>
<tr>
<td>Maternal Age (in years)</td>
<td>33.84 5.15</td>
<td>31.26 5.87</td>
<td>1.84</td>
<td>.070</td>
<td>.476</td>
</tr>
<tr>
<td>Number of People in Household</td>
<td>3.74 1.08</td>
<td>3.67 1.09</td>
<td>0.23</td>
<td>.816</td>
<td>.066</td>
</tr>
<tr>
<td>DERS Total</td>
<td>78.20 19.56</td>
<td>72.38 29.68</td>
<td>0.99</td>
<td>.324</td>
<td>.247</td>
</tr>
<tr>
<td>Parent Negative Talk Ratio (%)</td>
<td>4.21 3.29</td>
<td>3.93 3.25</td>
<td>0.36</td>
<td>.724</td>
<td>.085</td>
</tr>
<tr>
<td>Parent Emotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dismissing Ratio (%)</td>
<td>0.37 0.73</td>
<td>0.39 0.87</td>
<td>-0.13</td>
<td>.900</td>
<td>-0.030</td>
</tr>
<tr>
<td>Child Distress Rating</td>
<td>16.76 19.63</td>
<td>22.76 22.82</td>
<td>-1.16</td>
<td>.250</td>
<td>-0.288</td>
</tr>
<tr>
<td>Child Emotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regulation Maturity</td>
<td>2.15 0.93</td>
<td>1.90 1.11</td>
<td>1.01</td>
<td>.316</td>
<td>.251</td>
</tr>
<tr>
<td>Maternal Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seeking</td>
<td>0.26 0.27</td>
<td>0.35 0.26</td>
<td>-1.39</td>
<td>.169</td>
<td>-0.346</td>
</tr>
<tr>
<td>CBCL Dysregulation</td>
<td>23.25 26.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profile</td>
<td>T≈66a</td>
<td>9.87 T≈65</td>
<td>-1.31</td>
<td>.194</td>
<td>-0.332</td>
</tr>
<tr>
<td>CBCL Externalizing</td>
<td>20.00 22.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems</td>
<td>T≈64</td>
<td>9.08 T≈63</td>
<td>-1.23</td>
<td>.223</td>
<td>-0.311</td>
</tr>
<tr>
<td>Child Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbalizations</td>
<td>10.28 7.47</td>
<td>12.11 12.84</td>
<td>-0.78</td>
<td>.439</td>
<td>-0.188</td>
</tr>
</tbody>
</table>

*Note. CBCL = Child Behavior Checklist*

*aMale and female *T*-scores were averaged to form the *T*-scores in this table.*
Table 10

Results of Chi-Square Tests between Attrition Status and Select Categorical Demographic Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>$p$</th>
<th>Cramer’s V</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Sex</td>
<td>0.50</td>
<td>.479</td>
<td>.081</td>
<td>.479</td>
</tr>
<tr>
<td>Maternal Occupation</td>
<td>4.32</td>
<td>.743</td>
<td>.258</td>
<td>.743</td>
</tr>
<tr>
<td>Maternal Education</td>
<td>4.26</td>
<td>.372</td>
<td>.256</td>
<td>.372</td>
</tr>
<tr>
<td>Family Income</td>
<td>4.14</td>
<td>.387</td>
<td>.249</td>
<td>.387</td>
</tr>
<tr>
<td>Marital Status</td>
<td>2.64</td>
<td>.450</td>
<td>.200</td>
<td>.450</td>
</tr>
</tbody>
</table>
**Figure 1**

Excerpt from the DECS Coding Manual on Emotion Dismissing Code

<table>
<thead>
<tr>
<th>Emotion Dismissing (ED)</th>
<th>Negative evaluations of child emotions</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Lines</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ED-P-1</strong></td>
<td></td>
</tr>
<tr>
<td>P: Calm down</td>
<td>ED-P-1</td>
</tr>
<tr>
<td>P: Calm</td>
<td>ED-P-1</td>
</tr>
<tr>
<td>P: Ya gotta calm</td>
<td>ED-P-1</td>
</tr>
<tr>
<td><strong>ED-P-2</strong></td>
<td></td>
</tr>
<tr>
<td>P: You don’t care about Mummy at the moment, do you?</td>
<td>ED-P-2</td>
</tr>
<tr>
<td>P: So why are you being so silly?</td>
<td>ED-P-2</td>
</tr>
<tr>
<td>P: That’s not nice.</td>
<td>ED-P-2</td>
</tr>
<tr>
<td>P: What’s so funny?</td>
<td>ED-P-2</td>
</tr>
</tbody>
</table>
| P: It’s not nice  
*Following the “no hitting”* | ED-P-2 |
| P: Settle down.  
*Could this be similar to “calm down” or is ‘settle’ too different than ‘calm’?* | ED-P-2 |
| P: Ow gentle. | ED-P-2 |
| **ED-P-3** | |
| P: You have to give mama a kiss.  
*Child actively pulling away so is ED and EI and when both use priority rule ED > EI* | ED-P-3 |
| **ED-N-1** | |
| P: Over here so you don’t hurt yourself | ED-N-1 |
| P: No hurting *(Context: Child has hit the mother)* | ED-N-1 |
| P: Don’t be sad | ED-N-1 |
| P: **You need to be happy**, okay? (Following “But you’re upset”) | ED-N-1 |
| P: What are you **upset** for? | ED-N-1 |
| P: Okay, you be **upset** then | ED-N-1 |
| P: Don’t **worry** | ED-N-1 |
| P: Why are you getting **upset**? Why are you **upset**? | ED-N-1 |
| P: Don’t **hurt** her face  
- Talking about child hurting the baby | ED-N-1 |
| P: Don’t **stress**. | ED-N-1 |
| **ED-N-2** | |
| P: You can’t be **naughty** for mommy. | ED-N-2 |
| **ED-N-3** | |
| P: Oh don’t **cry**. | ED-N-3 |
| P: It doesn’t help that you keep **crying** (talking about child’s nose running) | ED-N-3 |
| P: Stop **crying**. | ED-N-3 |
| P: Hey, no **hitting** | ED-N-3 |
| P: You missed out if you’re too busy **crying** | ED-N-3 |
| P: Why are you **crying**? | ED-N-3 |
| P: Don’t **hit** me | ED-N-3 |
| P: You didn’t **cry**  
- (Followed “You did very good”) | ED-N-3 |
| P: Don’t **smack** me. | ED-N-3 |
| P: Why are you **fighting** me?  
*Negative tone | ED-N-3 |
Figure 2

Toddler Distress Rating Form

Participant ID: ______________________   Date Coded: _________________   Coder: _______
Rating: 0 (no distress), 1 (mild distress, e.g., first signs of frustration or distress such as an angry grunt or whinge, a stomp of the foot or small jumping action), 2 (moderate distress, e.g., persistent crying but with some variation in intensity, able to be placated by mother’s voice) or 3 (extreme distress, e.g., screaming, gasping for air, having a tantrum, crying persistently with consistent or increasing intensity, hitting the box or mother while crying). If a child demonstrated both mild and moderate distress in one interval, code only the highest level of distress (e.g., moderate). Distress scores were calculated by adding points across the four scales and had a possible range of 0 to 90.

<table>
<thead>
<tr>
<th>Minutes</th>
<th>Seconds</th>
<th>Code: 0 (no distress); 1 (mild distress); 2 (moderate distress); 3 (extreme distress)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>20 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>30 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>40 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>50 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>60 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>Second minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>20 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>30 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>40 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>50 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>60 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>Third minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>20 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>30 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>40 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>50 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>60 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>Fourth minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>20 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>30 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>40 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>50 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>60 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>Fifth minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>20 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>30 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>40 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>50 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
<tr>
<td>60 seconds</td>
<td>0 1 2 3</td>
<td></td>
</tr>
</tbody>
</table>

Total Score: /90

Additional comments:
Figure 3

*Toddler Emotion-Regulation Maturity Rating*

**Situation Modification:** Behaviours that altered the situation in order to reduce the likelihood of a negative reaction or increase the likelihood of a positive reaction (e.g., walking away from the box; leaving the room or taking on another task such as looking at a book or playing with another object).

**Attentional Deployment:** Toddlers diverted their attention away from the stressful situation without trying to change the situation in any way (e.g., gazing away from the box; turning away from the box without taking up an alternative task).

**Cognitive Change:** Behaviours that indicated cognitive change include trying different strategies to open the box, engaging with the box in a different way (e.g., pushing it around the room like a train engine; tapping on the box to make a noise), playing with an alternative toy on top of the box, or sitting on the box.

**Maternal Support-seeking:** Seeking assistance, acknowledgment or comfort from their mothers. Toddlers are only coded for maternal support seeking when they were not demonstrating distress.

<table>
<thead>
<tr>
<th>Minutes</th>
<th>Seconds</th>
<th>Behaviours observed</th>
<th>Y</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 secs</td>
<td></td>
<td>Attentional Depl.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Situation Mod.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maternal sup. Seek.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 secs</td>
<td></td>
<td>Attentional Depl.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Situation Mod.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maternal sup. Seek.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 secs</td>
<td></td>
<td>Attentional Depl.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Situation Mod.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maternal sup. Seek.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 secs</td>
<td></td>
<td>Attentional Depl.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Situation Mod.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maternal sup. Seek.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 secs</td>
<td></td>
<td>Attentional Depl.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Situation Mod.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maternal sup. Seek.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60 secs</td>
<td></td>
<td>Attentional Depl.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Situation Mod.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Ch.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maternal sup. Seek.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Attentional Depl Score: /6
Situation Mod. Score: /6
Cognitive Ch. Score: /6
Maternal sup. Seek Score: /6

Participant ID: ___________ Date coded: ___________ Coder: ___________
RCT COMPARING PCIT-T AND COS-P

Figure 4

Graph of Treatment Completers and Dropouts per Group