2021

A Comparison of Clinic-Based and Telehealth Parent-Child Interaction Therapy

Abigail Peskin

A COMPARISON OF CLINIC-BASED AND TELEHEALTH PARENT-CHILD INTERACTION THERAPY

by

Abigail Peskin

A Dissertation Proposal Presented to the College of Psychology
of Nova Southeastern University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

NOVA SOUTHEASTERN UNIVERSITY

2021

DISSERTATION APPROVAL SHEET

This Dissertation was submitted by Abigail Peskin under the direction of the Chairperson of the Dissertation committee listed below. It was submitted to the College of Psychology and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Clinical Psychology at Nova Southeastern University.

September 30, 2021
Date of Defense

Approved:

Diana Formoso
Diana Formoso, Ph.D., Chairperson

Christian DeLucia, Ph.D.

Lourdes Suarez-Morales, Ph.D.

Jason Jent, Ph.D.

October 14, 2021
Date of Final Approval

Diana Formoso
Diana Formoso, Ph.D., Chairperson

Acknowledgements

A huge thank you to my advisor and dissertation chair, Dr. Diana Formoso, for your mentorship and guidance during my time in graduate school. Thank you to my dissertation committee, Dr. Christian Delucia and Dr. Lourdes Suarez-Morales, and all the faculty members who provided me with support throughout graduate school. I want to offer my deepest gratitude to the Parent-Child Interaction Therapy team at the University of Miami, particularly Dr. Jason Jent and Dr. Meaghan Parlade, who have supported me through thick and thin throughout this grueling, frustrating, often exhausting dissertation process. Thank you for always having the faith that I would get through it and pushing me to leadership positions on our team and in the field of IPCIT overall over the last few years. Working with both of you has made me feel like I truly belong in this field and have something valuable to offer the families and children I work with every day. Thank you to all of my friends and family who have supported me throughout my graduate training. Thank you to the team I consider family at UM, and the hundreds of families who have pushed their way through PCIT and come out the other side changed. This project literally would not be possible without you, and I feel grateful every day I get to learn more from you.

TABLE OF CONTENTS

LIST OF TABLES

A COMPARISON OF CLINIC-BASED AND TELEHEALTH PARENT-CHILD INTERACTION THERAPY

by

Abigail Peskin

Nova Southeastern University

ABSTRACT

Despite the rich evidence for early intervention to prevent/treat emotional and behavioral disorders, a gap continues to exist between research and practice, and multitudinous barriers prevent families from accessing vital evidence-based services. Telehealth is an emerging area of research and clinical practice, often proposed as a solution for multiple barriers to service provision. However, despite scattered promising evidence translating in-person treatments to telehealth, many treatments still lack established effectiveness in a virtual format.

Parent-child interaction therapy (PCIT) is a traditionally in-person behavioral parenting intervention uniquely well-suited to telehealth, as during typical clinic practice the clinician is not in the clinic room, but rather coaching caregivers from behind a one-way mirror. Translating to the home setting via telehealth allows the clinician to increase the potential for generalization while simultaneously decreasing treatment barriers for both clinician and family. In this study, investigators examined change in child behaviors, caregiver stress, and caregiver child-directed statements for 58 families who completed internet-based PCIT and 140 families who completed clinic-based PCIT. Caregivers self-selected into either clinic or telehealth treatment, so propensity scores were computed to control for pre-existing variability between groups due to selection bias. Paired *t* tests confirmed that all caregiver and child outcome variables changed significantly from pre-

to post-treatment, including child disruptive, externalizing, adaptive behavior and compliance, caregiver stress, and caregiver positive and directive/corrective child-directed statements. Then, hierarchical linear regressions were conducted to examine the effect of treatment group (i.e., IPCIT or clinic) on the variability in each outcome. Outcomes did not significantly differ across treatment modalities for child disruptive, externalizing, adaptive behaviors and compliance and for caregiver stress. However, caregivers in the IPCIT modality demonstrated smaller increases in positive child-directed statements and smaller decreases in directive/corrective child-directed statements compared to the clinic-based group, exhibiting small effect sizes for both outcomes. These findings replicate previous research establishing IPCIT as effective at improving child behavior and caregiver stress and extend such findings to a wider group of children using a larger, more diverse sample who received a shorter course of treatment than previous IPCIT studies. Implications and directions for future research are discussed.

A Comparison of Clinic-Based and Telehealth Parent-Child Interaction Therapy

Chapter I: Statement of the Problem

Although the numbers fluctuate annually, each year approximately 20% of children experience emotional and/or behavioral symptoms severe enough to be considered eligible to receive treatment in the United States (NIMH, 2017). Despite the staggering number of children experiencing emotional and/or behavioral symptoms, about 80% of those children in need of services never receive them (Bringewatt & Gershoff, 2010), resulting in high cost over time both to society and families. By some estimates, disruptive behavior has been found to be the costliest of all mental health problems in the United States (Kazdin, 1995). In the United States, it is approximated that over 2 million children ages 3-17 currently qualify for a disruptive behavior disorder, such as Oppositional Defiant Disorder (ODD) or Conduct Disorder (CD), a number which does not capture those whose behaviors are severe and impairing but do not qualify for a diagnosis, or those who are undiagnosed (Perou et al., 2013). Further, longitudinal studies indicate that even early in childhood the presence of disruptive behaviors can predict their persistence over time (Moffitt, 1993; Patterson, Capaldi, & Bank, 1991). Clinically elevated child disruptive behaviors have even been described as a chronic illness due to the clear course of symptom development beginning in childhood, and the poor long-term prognosis indicated in the absence of effective, evidence-based treatment (Kazdin, 1990).

Untreated children with clinically elevated disruptive behavior are at heightened risk for the development of multiple adverse life events, including incarceration, alcohol and substance use, teenage pregnancy, and inter-partner violence (Fergusson, Horwood,

& Ridder, 2005). Indeed, many of these outcomes occur even after treatment due to the proliferation of treatments that are not evidence-based (Weisz et al., 2005, Comer et al., 2015). However, early access to high-quality evidence-based treatment can have significant implications for child behavioral and developmental trajectories (Scott, Augimeri, & Fifield, 2017). For children who require such services, early intervention is key. The earlier children receive services, the less the services will cost the family and society, and the less intensive the treatment will need to be over time (Scott, Augimeri, & Fifield, 2017). Therefore, early access to effective, evidence-based treatments is crucial for children with clinically elevated disruptive behaviors, to prevent the development of adverse life events in the future, and to decrease family and societal costs.

Telehealth as a Potential Solution to Barriers to Evidence-Based Treatment

Despite the importance of early, effective intervention for disruptive behaviors, many children do not receive psychological services when they need them, if they receive them at all (Scott, Augimeri, & Fifield, 2017). Traditional behavioral interventions for families often include caregivers traveling to a clinic and speaking with a therapist, learning strategies in a group or individual setting, and then taking those behavioral strategies home to practice them with their children. Sometimes the child is involved in the session, either primarily with the caregiver who is coached by the clinician, or with the clinician while the caregiver waits or watches, with the goal of learning from observation. This traditional format for behavioral interventions usually requires families to take time out of their schedules not just for the treatment session, but also for the commute to and from the clinic on a weekly basis. It often requires that caregivers wait with their children in a waiting room with other families. The limited access to treatment

experienced by so many families can be often attributed to multitudinous barriers that are either too much for caregivers to surmount or do not present with obvious or attainable solutions within the confines of the traditional format of clinical behavioral treatment. However, many of the barriers that exist for traditional psychological services can find their solutions in the world of telehealth. Below several prominent barriers to mental health care access are described, as well as the means by which telehealth can be harnessed as a mechanism for decreasing or overcoming those barriers.

One of the drawbacks of clinic-based services for parenting is that, although caregivers find that child behavior improves in the clinic, sometimes they find it more difficult to generalize behavior outside of the clinic because in the clinic, the therapist can standardize many factors that cannot be controlled in other settings (e.g., at home; Benoit, Edwards, Olmi, Wilcynski, & Mandal, 2001; Forehand & Atkeson, 1977, Swan, Carper, & Kendall, 2016). Behaviorally, when children present to treatment, they have a history of noncompliance in the home setting, but no history of noncompliance/disruptive behavior in the clinic. Therefore, the child is conditioned to misbehave more quickly and more consistently at home, where behavioral patterns and contingencies are pre-established. In the clinic, disrupting these patterns occurs more quickly, and the clinic becomes a conditioned setting for appropriate behavior and compliance. To generalize to the home setting, the caregiver must practice the same strategies at home, and practice them more persistently to undo the behavioral patterns that are conditioned on the cues in the home setting. Telehealth therefore increases the ecological validity of the treatment overall (Comer et al., 2015).

Home-based behavioral therapy (i.e., when the therapist goes to the home) has the potential to increase generalization of behavioral gains and decrease barriers to treatment access (e.g., thus decreasing attrition; Fowles et al., 2017), increasing the settings in which behavior is managed and producing long-term, lasting gains. However, often conducting in-home treatment is costly and logistically difficult due to the clinician needing to travel between many homes over the course of a day and the greater costliness of cancellations to productivity, given that clinicians are not in the office to be able to complete other tasks. Therefore, telehealth treatment conducted in the home may be a viable alternative to the therapist physically being in the home.

Some caregivers/patients experience self-conscious feelings about bringing their children to treatment or having others see their child's behaviors in public. For these patients, receiving treatment at home increases their likelihood of seeking and participating in sessions because they are not exposed to the judgement of others (Comer et al., 2015). Some families also find it uncomfortable to wait in a waiting room where it is clear to any others present that they are seeking health or mental health services. Waiting in the privacy of their homes negates this concern and increases the likelihood that families who are embarrassed about seeking mental health services will receive them (Luxton, Pruitt, & Osenbach, 2014).

Lack of motivation to attend sessions has also been found to be a significant reason why many people do not seek out mental health services. Indeed, caregivers who experience their own depression are much less likely to attend their child's treatment sessions than other caregivers (Ofonedu, Belcher, Budhathoki, & Gross, 2016). Committing to attending a session in a clinic requires activating the behaviors for finding

transportation, getting to the clinic, waiting in the waiting room (in the case of caregiver training, waiting with a child with disruptive behavior, which elicits judgement from others in public), and then getting home. In a telehealth format, on the other hand, patients are only required to sign into the session. They do not need to leave their homes, and thus much less behavioral activation is required for them to access services.

There are additionally several structural/physical barriers that impede client access to services. For example, one barrier that many clinics find difficult to address is a patient's lack of access to transportation to regularly attend sessions (Reardon et al., 2017). Many areas where patients need mental health services do not have the option of public transportation, limiting access to families who have the means to drive to session. Additionally, even if public transportation is available, it is often unreliable and takes a prohibitively long time to get from one place to another. These hurdles frequently make it too difficult to attend a regularly scheduled session at the same time each week. Lack of access to reliable transportation is therefore a frequent reason why many people are unable to access mental health services.

Relatedly, some patients simply do not live close enough to evidence-based, affordable mental health services, or their living situations are so fluid that attending sessions in a fixed location is not a realistic option. Particularly for families living in poverty, living too far away from the clinic is a common reason for dropping out of treatment prematurely (Bornheimer, Acri & Gopalan, 2018). Distance from evidence-based mental health care is a barrier that affects families in cities, but disproportionately affects those in rural, or nonmetro communities. Indeed, in several states throughout the United States, over 50% of rural counties lack any access to a psychologist at all, let

alone sufficient access to serve the needs of the entire population in that county (Summers-Gabr, 2020).

Transportation is sometimes a barrier due to access and time, as described above. The cost of transportation is also sometimes an additional barrier. When caregivers must leave the house to access treatment, they must finance transportation to session, parking when they arrive, and childcare for other children. Indeed, low-income caregivers specifically list access to transportation as one of the most interfering barriers to help-seeking (Keller & McDade, 2000). Telehealth removes the need for any of these expenses, thus opening the opportunity for treatment to a wider variety of low-income families (Smit, Cuijpers, Oostenbrink, Batelaan, de Graaf, & Beekman, 2006). Comparison of the cost of a treatment delivered via telehealth vs. in-person has demonstrated that telehealth is vastly more cost-effective for both families and providers than treatment delivered in person (Little, Wallisch, Pope, & Dunn, 2018). Specifically, Little and colleagues (2018) found that for a parent-coaching intervention with children with Autism Spectrum Disorder (ASD), switching intervention from the clinic to telehealth produced similar treatment outcomes and also saved the families and clinicians involved an average of $100 per session.

Beginning December 2019, psychologists worldwide encountered a new barrier to in-person services in the form of the novel coronavirus (i.e., COVID-19). To decrease the spread of this highly infectious virus, many countries enacted lockdown procedures which restricted clinic-based services and decreased family mobility. Worldwide surveys of mental health practitioners conducted by the World Health Organization discovered that with the onset of stay-at-home orders and clinic closures due to COVID-19, 93% of

countries reported that their services had been disrupted (WHO, 2020). At the same time as access to in-person services decreased, the need for mental health services increased exponentially, thus causing many agencies to attempt to quickly adapt to telehealth delivery of services (Pierce, Perrin, Tyler, McKee, & Watson, 2020). Mental health experts around the world have described a mental health crisis catalyzed by COVID-19's onset, including increases in both caregiver and child mental illness (Fontanesi et al., 2020). Psychologists attribute this wave of mental health needs to a multitude of factors, including the social isolation caused by social distancing and remote work and school, the stress of many caregivers losing their jobs/incomes, the worry about contracting COVID-19 or infecting vulnerable family or friends, increased screen time use by children, and the loss of many child activities that normally were held outside of school hours but were canceled or postponed to prevent the spread of COVID-19 (Fitzpatrick, Carson, & Weisz, 2020). As these services are more widely necessary due to worldwide barriers to in-person services, it is necessary to ascertain the efficacy of each treatment delivered virtually.

Barriers to Telehealth

Despite the advantages of delivering services via telehealth, there are several important drawbacks to this approach compared to traditional clinic-based treatment that clinicians will need to address for both families and therapists to reap the aforementioned benefits of treatment via telehealth. Cost of engaging in telehealth can be a barrier for patients accessing treatment. Unless families already have the technology for completing telehealth sessions, obtaining such tech would likely prevent many families from engaging in services. In the past, advanced hardware was required for families to access

telehealth in a manner that was encrypted and HIPAA compliant. However, in recent years software has been developed to allow for families to access HIPAA-compliant telehealth sessions on their tablets and even through their smartphones, thus greatly decreasing the cost to families and increasing access for a wider population (Brooks et al., 2013). Internet is often also required for telehealth services, and this requirement can additionally limit access for some patients, particularly low-income families who are much less likely to have access to high-speed internet (Zickuhr & Smith, 2012). In a similar vein, the COVID-19 pandemic has revealed how important it is that families obtain basic computer literacy skills to ensure that they engage in effective treatment (Garcia et al., 2021). Therefore, family computer literacy skills should be carefully assessed and addressed as relevant (e.g., by providing a "technology training" session before starting the intervention, as was done in the study described here and has been discussed in existing literature (Comer et al., 2017; Garcia et al., 2021).

Cost of treatment can additionally be a barrier on the therapist's side of treatment. First, most therapists do not have training in telehealth when they are trained initially, as it is a relatively new avenue for service delivery. Providing telehealth therefore requires clinicians to receive additional training (Glueckauf et al., 2018). Even among those providing telehealth services, about 40% of providers feel they do not have adequate training for the remote delivery of the services they are providing (Glueckauf et al., 2018), and many clinicians cite this lack of training as one of the main reasons that they prefer in-person treatment to telehealth, or why they have not attempted to provide telehealth treatment to their patients (Perry, Gold & Shearer, 2019). Adequate training is often cost-prohibitive for clinicians who operate independently outside of academic

institutions. Indeed, sometimes the technology required to complete telehealth services alone makes these services too expensive for individual providers (Kraetschmer, Deber, Dick, & Jennett, 2009; Kruse, Karem, Shifflett, Vegi, Ravi, & Brooks, 2018). Though notably, despite these initial training costs for clinicians, the expansion of services and lower dropout rate provided by technology-based intervention programs may eventually yield a net cost savings (Comer et al., 2017; Garcia et al., 2021). As technology allows the expansion of timely services to families in need, the costs of disruptive behavior to families and societal institutions should decrease (Scott, Augimeri, & Fifield, 2017).

One of the main goals of telehealth expansion is to decrease structural barriers to patient access to mental health treatment. However, licensure limitations significantly impede the growth of telehealth across state and country lines (Chou, Bry, & Comer, 2017). In the United States, each state has licensure limitations about who can practice telehealth, including where the clinician has to be located, and where the client needs to be located. As licensure is usually issued at a state level, providing treatment to someone in a different state requires the clinician to become licensed in every state where s/he wishes to practice, with few exceptions (Brooks et al., 2013). Although these licensure restrictions have been loosened in some ways in some places during the COVID-19 pandemic to allow for more social distancing and wider access to needed services, it remains unclear how many of these loosened restrictions will remain after the resolution of the pandemic (Shachar, Engel, & Elwyn, 2020).

Ethical Concerns for Telehealth

Confidentiality. The main barriers to the use of telehealth as a widespread practice involve the unique ethical concerns that arise in the use of treatment via

videoconferencing that do not occur in a clinic room. First, in a clinic room the clinician can, for the most part, control who can hear what is occurring in the session, improving confidentiality for the client. During a telehealth session, the clinician can control who hears what the client says on their side of the session but cannot control the environment on the client's end. Additionally, as the clinician can only see what is in the frame of the computer's camera, there is a chance during a telehealth session that there may be someone else present who does not have express consent to view the content of the sessions. Thus, consent processes for telehealth sessions need to explain to families that the clinician cannot guarantee confidentiality on the client's end of the digital interaction, and clinicians also need to discuss with the family ways to maximize privacy (APA 2013). Clinicians also need to explain to families how they will be maintaining client privacy on their end so that families are more confident that their sessions will be confidential.

Prevention of Harm. Psychologists also have an ethical responsibility to exert every reasonable effort to prevent harm from coming to their clients during treatment (APA, 2010; Romani & Schieltz, 2017). In the clinic, the environment can be controlled, and safety precautions can be taken much more effectively and comprehensively than when the therapist is conducting services remotely via telehealth. For example, if the caregiver removes the child from the view of the camera, the therapist has no way to know what is happening except by asking the caregiver. If the child is hurt during a clinic-based session, the therapist can both see everything that occurs and potentially intervene or prevent further injury. If the child is hurt during a telehealth session, the therapist can contact local emergency services and guide the caregiver through managing

the situation but cannot physically help with an emergency (Crum & Comer, 2016). Additionally, if contact with the family is lost during an emergency (i.e., the videoconference ends abruptly), the clinician has no way to address the emergency due to the physical distance from the client. Therefore, in telehealth treatment, therapists have an ethical obligation to engage in more safety planning than they would for in-person treatment. This planning should take place from the first session and should include a discussion of the risks to safety in the environment and the availability of emergency supports. Therapist is also expected to help the caregiver decrease distractions and disruptions during session and ensure that the setting is the most conducive possible for therapeutic progress. All of these steps to set up the therapy setting are tasks that the clinician would usually perform independently in the clinic, but in the home setting, must learn to guide the caregiver to accomplish on their own (APA, 2013).

Competence. The American Psychological Association Ethical Guidelines dictate that psychologists must not operate outside their area of competence, as doing so can compromise the quality of the treatment provided to the patient, and thus produce poorer patient outcomes (APA 2010; Romani & Schieltz, 2017). When clinicians operate outside of their competence, it can cause patients to waste valuable time engaging in ineffective therapy. Additionally, such patients may adopt a belief that mental health services are ineffective in general, thus decreasing a likelihood of seeking those services in the future for themselves and their children. Therefore, following the expectation of practicing within one's area of competence ensures that families receive effective services currently, and ideally in the future as well. To achieve competence for telehealth services, a clinician must be competent at providing in-person services and also achieve the

additional competencies required to smoothly conduct their services virtually. This includes understanding how to translate their treatment into telehealth, how to use the required software and hardware required for session, and how to guide families through using this technology as well. Just as expected with in-person services, clinicians are expected to be familiar with the most current literature about their treatment in its telehealth translation (APA, 2013).

In the past, these perceived barriers to engaging in telehealth services have kept many clinicians from using telehealth (Glueckauf et al., 2018; McClellan, Florell, Palmer, & Kidder, 2020). However, avoidance of telehealth was based not on any proven lack of effectiveness but rather clinician concern about the ethical dilemmas presented above. More recent thinking in the field suggests that these ethical concerns can be addressed, and ethical violations prevented through thorough planning and telehealth training (APA, 2013; Webb & Orwig, 2015). This ability to train therapists to competence in telehealth, as well as a growing literature demonstrating the strengths of therapy conducted over telehealth increases the push in the field to expand further into this modality. However, as more and more therapists venture into virtual therapy, it will be important to continually evaluate the efficacy of existing treatments, ensuring that they retain their efficacy when delivered via a novel format.

Summary. Clinicians and researchers alike have described telehealth as the answer to many barriers to in-person treatment, including lack of access to transportation, feelings of stigma/judgement about being seen in a treatment clinic, and the time lost traveling back and forth to a clinic. Many subsequent studies have demonstrated that these theoretically postulated ways that telehealth can decrease barriers to treatment have

also been borne out in empirical analyses. Telehealth can decrease the cost of treatment for families and clinicians. It has been shown to improve treatment access for families who live far from clinics, lack reliable transportation, or for whom the costs of transportation preclude engagement in therapy. Despite these multitudinous benefits to telehealth, many therapists in particular remain skeptical about its widespread use. Barriers to widespread dissemination of telehealth include geographical areas with limited high-speed internet access and geographical licensure limitations for therapists requiring they only practice within the state where they are licensed. Additionally, lack of comfort, experience, and/or training with telehealth lead many therapists to continue in-person treatment to the exclusion of anything remote. Many therapists are concerned that a lack of specialized telehealth training (often difficult to access due to funding or availability) will decrease their efficacy and/or potentially cause harm to the patient. Meanwhile patients who may benefit from remote treatment are waiting for options available to them. Increasing access to families begins with improving/increasing therapist confidence in the effectiveness of treatments delivered via telehealth, which is the central goal of the current study.

Chapter II: Review of the Literature

Preliminary Results of Telehealth in Behavioral Health

Telehealth treatments were originally designed to address the barriers that rural families experienced related to treatment access. As evidence-based in-person treatments were adapted to a telehealth format, researchers needed to answer the question of whether the telehealth version could also be an efficacious treatment. In a review of videoconferencing telehealth interventions for children, Slone, Reese and McClellan (2012) found that treatments delivered virtually were equally effective when compared to their face-to-face versions. Recently, researchers examined the implementation of Trauma-Focused Cognitive-Behavioral therapy (TF-CBT) with children via videoconferencing. This pilot study found similarly large effect sizes (i.e., $d = 2.93$) to in-clinic TF-CBT, demonstrating promise for reaching rural youth who are historically underserved, particularly in the receipt of evidence-based treatments (Stewart, Orengo-Aguayo, Cohen, Mannarino, & de Arellano, 2017). Other researchers have treated anxiety remotely using a telehealth adaptation of Cool Kids, an evidence-based treatment for elementary-aged children. These researchers found that their telehealth version of this treatment resulted in significant decreases in severity of child anxiety, functional impact of that anxiety, and externalizing and depressive symptoms. Importantly, researchers found similar symptomatic improvement reported by both caregivers and children (McLellan, Andrijic, Davies, Lyneham, & Rapee, 2017).

Telehealth in Caregiver Coaching/Training

Over the past few years, psychologists have endeavored to disseminate telehealth services for behavioral caregiver training treatments specifically, assessing the efficacy of

these treatments conducted via videoconferencing, and their ability to decrease structural barriers to treatment attendance. Lindgren et al. (2016) evaluated a telehealth model for coaching caregivers to conduct Applied Behavior Analysis (ABA) with their children with Autism Spectrum Disorder (ASD) as an alternative to in-home ABA therapy conducted by a behavior analyst. This study found equivalent reduction in behavior whether the caregiver or clinician was leading the home sessions and found that having the clinician coach the caregiver at home greatly decreased the overall caregiver and clinician cost of conducting the treatment. Other studies of telehealth interventions for caregiver training for children with ASD have found improvements in imitation (Wainer & Ingersoll, 2015) and functional communication (Vismara, McCormick, Young, Nadhan, & Monlux, 2013). Additionally, researchers have found that conducting such treatments over the internet, which are often prohibitively expensive for families, decreases cost and increases access for families who live in rural, underserved areas (Ingersoll & Berger, 2015; Peterson, Piazza, Luczynski, & Fisher, 2017; Vismara et al., 2013).

Many of the evidence-based caregiver-training treatments for disruptive child behaviors share common components, which the field agrees serve as active ingredients for child behavior change (e.g., teaching caregivers how to set limits, how to reward positive behavior, and how to use time out appropriately). These programs go by several different names depending on the format of delivery (e.g., individual or group, online or in person) and the involvement of the child (e.g., as a topic of discussion or an active session participant). Although originally developed for in-person services, telehealth translations of these programs have demonstrated promise for producing similar

outcomes to in-person delivery formats, decreasing child disruptive behaviors from clinically elevated to below clinical levels. Reese and colleagues (2012) examined the outcomes of a group format of Triple P Positive Parenting Program delivered via videoconferencing technology. Similar to the course of Triple P delivered in person, caregivers in the telehealth group reported significant decreases in children's disruptive behaviors and decreases in caregivers' stress post-treatment. Another telehealth intervention for disruptive behaviors, Strongest Families, was compared to in-person therapy and both treatment modalities were found to produce equivalent decreases in child disruptive behavior and caregivers' stress (Olthuis et al., 2018). These treatments are both group parent training interventions, in which clinicians teach caregivers new strategies for behavior management during session, and caregivers practice those strategies with their children during the week. Group parenting interventions like these are often used for improving parenting skills preventatively, to decrease the likelihood that child behavior will increase to the level of clinical significance. When child behavior has escalated to clinically elevated levels, often a more intensive, individualized treatment is needed.

Parent-Child Interaction Therapy

One example of an intensive, individualized treatment for clinically elevated disruptive child behavior is Parent-Child Interaction Therapy (PCIT; Eyberg & Funderburk, 2011). PCIT is a well-established treatment for disruptive behaviors in young children with years of evidence backing its efficacy and effectiveness in many different populations. Originally developed to treat children with conduct disorder and oppositional defiant disorder (ODD), it was later found to be efficacious for increasing

compliance and decreasing disruptive behaviors for children with other diagnoses, including Attention-Deficit Hyperactivity Disorder (ADHD), Autism Spectrum Disorder (ASD), and a variety of anxiety disorders (Niec, 2018). PCIT is a parent-coaching treatment in which a caregiver and child dyad play together in one room, and a therapist coaches the caregiver from behind a one-way mirror. Caregivers are coached to acquire specific skills that align with an authoritative style of parenting. The treatment is divided into two parts. In the first phase of treatment, caregivers are taught to increase the warmth in their relationship by spending time playing with their children, providing specific positive verbal feedback and allowing the child to lead the playtime. In the second phase of treatment, caregivers are taught to give effective, direct commands to children and to enforce compliance with time out.

Conceptually, PCIT was developed from both behavioral and attachment research by Dr. Sheila Eyberg. Specifically, the simultaneous goals of increasing warmth in the caregiver-child relationship and increasing limits reflects Dr. Diana Baumrind's seminal parenting research suggesting that primary caregivers who displayed an authoritative (i.e., high warmth and high limits) parenting style precipitated child development that resulted in a higher likelihood of positive child outcomes than any other parenting style (Baumrind, 1966; Baumrind, 2013). Behaviorally, PCIT is grounded in an operant strategy for modifying child behavior that Eyberg learned from Dr. Constance Hanf. Within Dr. Hanf's two-part model, parents were first taught to use differential attention, decreasing disruptive behaviors by ignoring them, and increasing positive behaviors by providing intentional positive attention (McNeil & Hembree-Kigin, 2010). In the second

part of the model, parents were taught to use calm, consistent consequences for noncompliance, undoing coercive cycles of interaction with their child.

Eyberg learned this model of intervention from Hanf, and added strategies she had learned from play therapy, in which the parents followed the child's lead, provided undivided attention, described the child's actions (termed "behavior descriptions"), repeated child verbalizations (termed "reflections"), engaged (i.e., "imitated") in enthusiastic (i.e., "enjoy") play with the child, and used specific praise for positive behaviors (i.e., "labeled praises"). Although much of child *mis*behavior occurs outside of play, caregivers practice the PCIT strategies in a play-based setting, because young children do the majority of their learning and problem-solving during play (Eyberg, 1988).

Although clinicians in PCIT often model the skills they teach to caregivers when they are together in the clinic room, the majority of session duration takes place without the clinician in the room. The caregiver is instead coached through a one-way mirror. This caregiver-coaching model is vital for child behavior change during the preschool years, as the goal is to both optimize the amount of time that a child receives consistent behavioral consequences (i.e., the clinician is with them one hour per week, the primary caregiver much longer) and to undo destructive cycles of interaction between the caregiver and the child. Even with highly effective behavioral therapy, a child whose caregiver does not change their responses will likely have difficulty maintaining behavioral change after treatment ends (McNeil & Hembree-Kigin, 2010).

Telehealth with Parent-Child Interaction Therapy

PCIT is traditionally performed in a clinic setting but has recently begun to be piloted as a telehealth treatment in several clinics around the country (Comer et al., 2017) in an attempt to reach families who cannot come into the clinic weekly to receive treatment. PCIT is uniquely suited for a format like telehealth (Gurwitch et al., 2020), as usually the clinician is not in the room with the family during a PCIT session in the clinic, but rather behind the one-way mirror (Eyberg & Funderburk, 2011). During a telehealth session, instead of in the clinic room, the family is at home playing with their toys at home. This model has been implemented by clinicians over the last few years, and preliminary results suggest positive outcomes in child behavior and caregiver skill growth (Comer et al., 2017; Kohlhoff, Wallace, Morgan, Maiuolo, & Turnell, 2019). Kohlhoff and colleagues (2019) conducted several case studies and demonstrated that similar behavioral outcomes could be found between clinic-based and telehealth PCIT cases for families in very remote and rural settings. Comer and colleagues (2017) conducted a randomized controlled trial comparing clinic-based and telehealth delivered PCIT in a racially and ethnically diverse sample of children ages 2-5 and their caregivers. Comer et al. (2017) found that both treatments significantly decreased child disruptive behaviors and burden to caregivers, and that in some ways IPCIT produced more robust changes than clinic-based PCIT. IPCIT in Comer and colleagues' (2017) study was also associated with significantly fewer caregiver-perceived barriers to treatment.

Comer and colleagues (2017) utilized a randomized control trial to compare telehealth and traditional clinic-based PCIT. Although this study design allowed researchers to eliminate bias due to participants selecting their treatment modality, participant selection was also limited to those close enough to the clinic to attend

treatment in person (in case they were randomized to that condition). Although families are sometimes able/willing to accept either treatment format, families often select telehealth treatment because attending treatment in the clinic is not possible. Therefore, participants most in need of telehealth were arguably not included in the sample for Comer and colleagues' study. Further analyses are required to determine whether caregivers who self-selected into a telehealth treatment format will also demonstrate similar treatment gains to families who self-select into an in-person treatment format.

Despite Comer and colleagues' findings that caregivers reported fewer limitations to treatment in the IPCIT condition, they noted in their limitations that they were unable to reach families with geographical barriers to clinic-based treatment, as their use of a randomized controlled trial required families in the whole sample live close enough to the clinic to be able to receive treatment either virtually or in person. To address effectiveness of IPCIT in a more rural, remote population specifically, Fleming and colleagues (2020) conducted an open trial of the effectiveness of IPCIT in rural Australia. Results of this open trial produced significant improvement in both child disruptive behaviors and observed parenting behaviors, comparable to previous studies of clinic-based PCIT. To maintain the ecological validity of Fleming and colleagues' findings and the theoretical rigor of Comer and colleagues' study, the literature is missing an empirical comparison of the findings between IPCIT and clinic-based treatment in a sample of families who would naturally have selected into each location for services. This study is the first step toward closing that gap, and advancing IPCIT as a viable, comparable treatment for families who cannot attend in-person treatment and those who might choose telehealth options for a variety of other reasons.

Goals and Hypotheses

The goal of the current study is to examine the outcomes of PCIT delivered via telehealth as they compare to outcomes for clinic-based PCIT. Therefore, the goal is twofold. The first goal is to examine whether IPCIT produces significant improvements in child disruptive and adaptive behavior, child compliance, caregiver stress, and child-directed statements (i.e., positive statements and directive/corrective statements). The second goal is to compare outcomes of internet- and clinic-based treatment.

We hypothesize that families participating in IPCIT and clinic-based PCIT will both separately exhibit significant changes in frequency and quantity of child and caregiver outcomes from pre- to post-treatment. Regarding child outcomes, we expect that children in both IPCIT and clinic-based PCIT will exhibit reductions in the frequency and quantity of child disruptive behavior, child externalizing behavior, child adaptive skills and child compliance. Regarding parent outcomes, we expect that parents in both IPCIT and clinic-based PCIT will reduce the frequency and quantity of caregiver overall stress, caregiver stress related to child behavior, caregiver positive statements, and caregiver directive/corrective statements. Both IPCIT and PCIT are also hypothesized to produce increases in caregiver positive statements and decreases in caregiver directive/corrective statements. Second, we hypothesize that these treatment gains will not differ for families participating in IPCIT and clinic-based PCIT. Stated another way, we hypothesize that PCIT will be equally effective regardless of whether it was implemented using telehealth services (IPCIT) or using clinic-based services.

Chapter III: Method

Participants

Participants were children aged two to seven years old whose caregivers sought treatment through a Parent-Child Interaction Therapy program located in South Florida. Participants in this study represent a small segment of a larger study.

Children whose caregivers contacted the clinic for services were determined eligible if they fell in the appropriate age-range for services (i.e., 2-7), they lived in the county served through the grant-funder, and their caregivers reported clinically elevated behavioral concerns on either the Eyberg Child Behavior Inventory (ECBI; elevated $\geq$ 131; Eyberg & Pincus, 1999) or the externalizing scales for the Behavior Assessment System for Children, Third Edition (BASC-3; clinically elevated = *T* score $\geq$ 60; Reynolds & Kamphaus, 2015). If children were not eligible for services, they were referred to a different parenting program in the community more appropriate for less severe behavioral concerns. Treatment for all participants was provided free of charge to families, as it was grant-funded.

Participants in this study fall into two groups based upon their preferred method of service delivery. The first group of caregivers self-selected to receive treatment via telehealth, often because the in-person clinics available were too far from their homes to be reachable weekly during clinic hours, caregivers worked from home and traveling to the clinic took time out of their workday, or transportation was too unreliable to make it to weekly sessions. After March 2019, the in-person clinics closed temporarily, and all new families were enrolled in telehealth due to the COVID-19 outbreak. The second group of participants consisted of children from the same sample of referred children, but

whose caregivers elected to receive treatment at the clinic instead of via telehealth. All in-person families received treatment before the COVID-19 pandemic and so it was not necessary for any family to transition to telehealth after beginning in person treatment.

Children in both the IPCIT treatment ($N = 58$) and clinic-based control ($N = 140$) sample groups were ages 2-7 (IPCIT $M = 5.09$, $SD = 1.65$; clinic-based $M = 4.54$, $SD = 1.39$). In the IPCIT sample, 27 families (46.55%) included two caregivers in treatment, and for 31 families (53.45%) only one caregiver was actively enrolled. In the clinic-based treatment sample, 56 families (40.00%) included two caregivers in treatment, with 84 families (60.00%) only enrolling one. Out of the IPCIT sample, 14 families (24.14%) completed treatment before COVID-19, and 44 (75.86%) completed treatment after the beginning of the pandemic. See Table 1 for detailed information about the demographics of children and their caregivers in this sample including caregiver self-reported race/ethnicity, child sex, and caregiver highest education attained. Table 1 also presents treatment-level variables for participants, including the level of therapist's training and the language in which treatment was completed.

Procedures

Institutional Review Board Approval

This study was approved by the Institutional Review Board at the University of Miami, Miller School of Medicine, and approval from the principal investigators has been provided to use the data included in analyses for this dissertation. The dissertation study was determined to be exempt from review by the Institutional Review Board at Nova Southeastern University.

Table 1

Descriptive caregiver-, child-, and treatment-level variables, separated by treatment group

	IPCIT		Clinic-based	
	N	Percentage of IPCIT sample	*N*	Percentage of Clinic-based Sample
Identified Race/Ethnicity				
Hispanic[a]	33	56.90	99	70.71
White non-Hispanic	16	27.59	30	21.43
African American	4	06.90	5	03.57
Other	5	08.62	6	04.29
Caregiver Highest Level of Education Received				
High school degree[b]	4	06.90	14	10.00
Some college	4	06.90	20	14.29
Associates degree	3	05.17	20	14.29
Bachelor's degree	16	27.59	45	32.14
Graduate degree	31	53.45	41	29.29
Child Sex				
Male	35	60.34	102	72.86
Female	23	39.66	38	27.14
Level of clinician training				
Intern[c]	5	08.62	20	14.29
Postdoctoral fellow	41	70.69	47	33.57
Masters level clinician	8	13.79	55	39.29
Clinical faculty	4	06.90	18	12.86
Language therapy completed				
English	51	87.93	124	88.57
Spanish	7	12.07	16	11.43

Note: [a]Hispanic caregivers were used as the reference variable for dummy coding the categorical variable of caregiver ethnicity, and therefore "Hispanic caregiver ethnicity" does not appear as a unique predictor variable in the analyses below.

[b]Caregivers with a high school education were used as the reference variable for dummy coding the categorical variable of caregiver education, and therefore "High school caregiver education" does not appear as a unique predictor variable in the analyses below.

[c] "Intern therapist" was used as the reference variable for dummy coding the categorical variable of therapist level of training, and therefore "Intern therapist" does not appear as a unique predictor variable in the analyses below.

Recruitment

Families who were referred to the PCIT clinic in South Florida completed screening information about their concerns about their child's behaviors, their weekly availability, background information about their families, and their preferred location of services. Before the COVID-19 pandemic, families were offered their choice of six different clinics scattered around the county, or treatment via telehealth. After March 2020, the clinic moved to 100% remote services. In this case, when families called for screening, families were still offered either telehealth or in-person services, but it was communicated to families who preferred in-person services that it was unclear when physical clinics would re-open. Families were free to refuse telehealth and choose to receive clinic-based treatment instead at any time (pending the re-opening of the clinics), and there were no additional incentives for receiving one condition over another.

Treatment Conditions

Common to Both In-person and Telehealth PCIT. All treatments were provided by master's level clinicians, predoctoral intern trainees, postdoctoral fellows, and/or doctoral-level licensed clinicians. Any unlicensed clinicians providing services were provided one hour of supervision per week and live supervision during sessions as needed by a licensed doctoral-level psychologist. All therapists received training to become certified in PCIT, which includes completing 40 hours of didactic training, two PCIT cases to completion with consultation/supervision of a trainer, and completion of specific skills-based competencies (e.g., independently coaching several criterion sessions to fidelity, coding to fidelity, etc.). Caregivers attended two intake sessions to complete consent forms, a caregiver biopsychosocial interview and the DPICS-IV

assessment. This assessment entails observing the parent during 5-minute intervals in which the caregiver is instructed to follow the child's lead (i.e., Child-Led Play), the caregiver is instructed to lead the play (i.e., Parent-Led Play), and the caregiver is told to instruct the child to clean up the toys independently (i.e., Clean-Up). During the Child-Led Play DPICS-IV observation, the clinician also codes the caregiver's statements and behavior, tallying the number of statements that fall in the categories of neutral talk, behavior descriptions, reflections, labeled praises, unlabeled praises, questions, commands, and criticism/sarcasm. The clinician also codes the caregiver's ability to imitate the child's play, use enthusiasm, and ignore disruptive behaviors by indicating whether caregiver skills in each area are "satisfactory," or "need practice." For ignoring disruptive behavior, clinicians also have the option of marking "N/A" if there is no disruptive child behavior during the five minutes of coding. During the Parent-Led play and clean-up portions of coding, the clinician coded caregiver commands and follow through with commands as well as child compliance.

After the two intake sessions were completed, the clinician completed a teaching session for the first phase of treatment (i.e., Child Directed Interaction; CDI) with the caregivers alone in which caregivers were taught the skills to be practiced during CDI. After this session, caregivers were instructed to begin practicing the five minutes of daily play time in which they practice the CDI skills. Caregivers then completed up to five sessions of coaching for CDI. Some families moved on sooner than five sessions if they met criteria for CDI skills before they had completed five sessions, indicated by their using 10 behavior descriptions, 10 labeled praises, 10 reflections, and three or fewer questions, commands, and critical statements during the first five minutes of session.

During each CDI session, the clinician checked in with the family about their week, coded CDI skills silently for five minutes, and then coached the caregiver through a headset to engage with their child in the play using the specific positive attending and differential attention skills taught to them during the teaching session. After coaching, clinician completed check out with caregivers, including reviewing the change in child behavior and caregiver coded statements from the previous session, as well as setting goals and priorities for homework completion during the week.

When caregivers achieved criteria for CDI or completed five coaching sessions, they moved into the second phase of the treatment. To begin the second phase (i.e., Parent-Directed Interaction; PDI), caregivers again completed a teaching session in which they learned the strategies for using effective commands and the time out sequence. Caregivers then received PDI coaching sessions until they reached 18 weeks of treatment or met PCIT graduation criteria. Graduation criteria for graduating earlier consists of (a) meeting the CDI criteria described above, (b) meeting PDI criteria, (c) and caregivers' reporting an ECBI Intensity score of 114 or lower. PDI criteria is determined by the caregiver(s)' use of 75% effective, direct commands and 75% correct follow-through for child compliance/noncompliance during five minutes of Parent-Directed Play coded at the beginning of PDI sessions. During typical criterion-based PCIT, caregivers have two sessions to meet all three of these criteria, but in time-limited PCIT, caregivers were expected to meet all three criteria within the same session to qualify for early (i.e., before 18 weeks) graduation.

During PDI sessions, caregivers were sometimes coded in CDI or PDI (as dictated by the PCIT Protocol; Eyberg & Funderburk, 2011), and coached in CDI and PDI

including coaching caregivers through a time out sequence if the child did not listen to commands. Occasionally PDI sessions were longer than the typical hour-long therapy session due to time outs extending beyond the session time. One month after graduating, families were invited to complete a follow-up session in which they were coded in both CDI and PDI skills and coached in both as well.

Unique to Clinic-Based PCIT. In the clinic-based condition, families received PCIT in the clinic setting for each of their sessions. The clinician completed the check in and check out portion of session with the caregiver and child in the playroom. Coaching took place over headset with the caregiver and child in the playroom and the clinician watching and coaching through a microphone on the other side of a one-way mirror.

Internet-Based Parent-Child Interaction Therapy (IPCIT). In order to be sufficiently trained in telehealth, therapists completed four hours of training specific to setting up the technology with the family and continuing to adhere to PCIT fidelity while conducting treatment remotely. This training was provided by the tech lead doctoral-level faculty on the team, also certified as a trainer for PCIT. Training was assisted by other clinicians (i.e., post-doctoral fellows and other faculty) certified in PCIT and practiced in delivering PCIT via telehealth. Training consisted of didactics as well as role-plays to simulate telehealth sessions and practice troubleshooting technology with families. After this initial training, when therapists began their first cases, they had the opportunity to attend group communities of practice about telehealth implementation, and consultation on-request for more experienced clinicians to shadow their cases and/or assist them in troubleshooting specific to the telehealth setting. During the COVID-19 pandemic when

every clinician was remote, therapists also were given access to an extensive manual created by others on the team which detailed how to make PCIT successful via telehealth.

At the beginning of telehealth cases, caregivers were offered an additional telehealth set-up session. This session always started with a phone call to the caregiver and then transitioned to the Zoom environment once families understood how to open and use Zoom. During this session, a therapist or a member of the data team met with them virtually to help them navigate each step of setting up the sessions with the therapist (e.g., signing into the videoconference, pairing the Bluetooth headset, clamping the tablet to the stand, and setting it up so the play area was clearly visible, etc.). This telehealth session was included specifically to combat potential difficulties with tech literacy, as had been done in other evidence-based studies using technology (e.g., Comer et al., 2017). Unfortunately, we did not directly assess whether families perceived this training as beneficial. However, significantly lower treatment dropout rates in the context of IPCIT provide indirect evidence that this technology training was at least effective enough for the vast majority of our families to feel competent enough with technology to complete training.

Before COVID-19. For families enrolled in the IPCIT condition prior to the COVID-19 pandemic when all therapist-client contact transitioned to virtual, caregivers attended either one or two intake sessions in person in the clinic where their therapist was located. At that time, caregivers were provided with all of the materials needed to conduct IPCIT sessions at home, including a tablet with a protective shatter-proof case, several clamps and stands to set up the tablet, a Bluetooth headset for the caregiver to wear during sessions, a corded backup headset, charging cords for both the Bluetooth

headset and the tablet, and a manual with a thorough explanation of how to set up at home. At the end of the second intake session, therapists spent about 20 minutes explaining all of the materials to caregivers and practicing several of the steps that caregivers would be conducting at home during sessions (i.e., pairing Bluetooth headsets, opening Zoom sessions, turning Bluetooth on and off in the settings on the tablet).

Telehealth During COVID-19. After the onset of the COVID-19 pandemic in March 2020, both intake sessions were completed via telehealth, although the content of sessions was identical to those intakes completed in person. At that time, borrowing telehealth equipment (e.g., tablet, Bluetooth headset, etc.) became optional for families and the need for this equipment was assessed when families completed their initial screening and qualified for services. If families required equipment to complete sessions, they had the option to pick it up from staff at a centrally located clinic, or to coordinate to have it dropped off at their home. Return of equipment was coordinated in the same manner. Explanation of the use of the equipment took place during the optional tech session between the family and the clinician described above instead of in-person as it had before the pandemic. Approximately 15% of families during COVID-19 borrowed equipment to complete sessions.

Common to All Telehealth. After the two intake sessions, the caregiver completed all subsequent sessions online through a HIPAA-compliant Skype for Business or Zoom session emailed to the caregiver. All sessions in the clinic and online were recorded, and families were informed during the consent process that they would be recorded (and were given the option to opt out of recording). Families in the telehealth condition completed additional consent documentation that explained the risks of

telehealth delivery of services (e.g., the clinician was not physically available for emergency situations, confidentiality concerns related to telehealth format) in accordance with risks detailed by the APA guidelines for telepsychology (APA, 2013).

Before coaching sessions began, the clinician also reviewed a safety plan with each family, including collecting emergency contacts and discussing emergency services for their current location. Clinician also discussed with the family how to make the environment safe (e.g., hiding or removing unsafe items). A new safety plan was created with the family each time they signed in from a new location unfamiliar to the clinician. At the beginning of treatment, clinician also reviewed with the family how to set up the treatment environment to increase the effectiveness of the session, making sure distractions were limited as much as possible and child and caregiver engagement was optimized.

Measures

Caregiver-report measures were completed either online (via REDCap, a HIPAA-compliant online application for managing research surveys) or on paper. Measures were completed in the caregiver's preferred language (i.e., English or Spanish). Measures were completed by all participating caregivers in both conditions (i.e., clinic-based and telehealth delivered treatment) at the beginning and end of treatment. Caregivers completed all intake measures during the intake process which began at the first intake session and was expected to be completed by the CDI Teach, usually the third session. Graduation assessments were completed in the same week as the graduation session – the last of the 18 weeks of treatment. Measures are reported below and linked with the study analytic plan in Table 5. See Tables 2, 3, and 4 for means and standard deviations for

each measure at pre- and post-treatment, for the total sample as well as for the families in clinic-based and IPCIT treatment separately.

Demographic Questionnaire. As a part of the intake assessments, caregivers completed information about child and caregiver demographics, including caregiver and child age, sex, self-identified race and ethnicity, primary language, and caregiver highest level of education attained. Participants were able to write in their race or ethnicity as they identified themselves and their children. For the purposes of analyses due to small sample sizes for multiple specific ethnic identities, race/ethnicity was divided into: Hispanic, Non-Hispanic White, African American, and Other (i.e., English-Speaking Caribbean, Russian, Multiracial, and Brazilian). For primary language, participants had the options of English, Spanish or Creole, but also had the option to choose multiple languages or write in one that was not listed. For caregiver education, participants could indicate they had received some high school, a high school degree or GED, some college or technical school, a college graduate degree, some post-graduate work, or a graduate/professional degree.

Frequency and Quantity of Child Disruptive Behavior. The frequency and quantity of child disruptive behavior was assessed using the Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999), a 36-item caregiver-report measure of disruptive behavior in children aged 2-16. Caregivers reported about the frequency of each of 36 disruptive child behaviors (e.g., lies, hits others, whines) on a Likert scale from 1 (i.e., never) to 7 (i.e., always). Caregivers also reported the quantity of disruptive behaviors. More specifically, they reported whether they found each behavior to be a problem, answering just *yes* or *no*, and the yes answers were summed resulting in a total

score of the disruptive behaviors the caregiver found problematic (ranging from 0 to 36). Raw scores of 131 or higher (*T* score = 60) on the Intensity scale and 15 or higher on the Problem scale (*T* score = 60) are clinically significant. The ECBI has demonstrated internal consistency reliability for both the Intensity (α = .95) and Problem (α = .93) scales (Eisenstadt, McElreath, Eyberg, & McNeil, 1994). Validity and sensitivity have also been established (Boggs, Eyberg, & Reynolds, 1990). In the current study, both raw scores of the frequency and quantity of disruptive behaviors were utilized to screen children to indicate eligibility for treatment as well as to measure pre- to post-treatment change.

Child Externalizing and Adaptive Behavior. Child externalizing and adaptive behavior were assessed using the Behavior Assessment System for Children, Third Edition, Parent Rating Scale (BASC-3 PRS; Reynolds & Kamphaus, 2015), a caregiver-report measure of emotional, behavioral, and adaptive functioning of children aged 2-21 years. On this measure, caregivers respond about the frequency of specific child behaviors on a four-point scale (i.e., never, sometimes, often, almost always). The BASC-3 has well-established validity and reliability (α = 0.83-.96, test–retest = 0.85–0.92; Reynolds & Kamphaus, 2015). In this study, the Externalizing Problems subscales (i.e., Conduct Problems, Hyperactivity, Aggression) and composite scores were used to screen for eligibility for program enrollment. The Externalizing Problems composite *T* score (i.e., composite of Hyperactivity and Aggression subscales for children ages 2-5, and Hyperactivity, Aggression and Conduct Problems for children ages 6-11) and Adaptive Behaviors composite *T* score were used to assess child treatment progress from pre- to post-treatment.

Child Compliance. Child compliance was assessed using the Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV; Eyberg, Nelson, Ginn, Bhuiyan, & Boggs, 2013), which is a behavioral observation coding system unique to PCIT. Child compliance is coded during the parent-directed interaction portion of the observation, in which caregivers are instructed to lead the play and instruct the child to follow their rules in the play. Then compliance is quantified as the percentage of effective commands that the child follows within five seconds after the command is given. Child compliance is coded at many sessions during the Parent-Directed Interaction portion of PCIT, but for the purposes of this study, the intake and post-treatment percentages will be utilized to measure child improvement in compliance from the beginning to end of treatment.

Caregiver Stress Overall and About Child Behavior. Caregiver overall stress and stress about child behavior were assessed using the Parenting Stress Index, Fourth Edition: Short Form (PSI-4-SF; Abidin, 2012), a 36-item caregiver-report measure of caregiver stress evenly divided into three separate scales measuring Parental Distress (i.e., stress related to being a parent/caregiver), Parent-Child Relationship Dysfunction (i.e., stress related to conflict/tension in the relationship between the caregiver and the child), Difficult Child (i.e., stress specifically related to disruptive child behaviors) and Total Stress, which is a composite of all three scales. Caregivers answer items on this measure on a five-point scale from *strongly agree* to *strongly disagree*. Higher scores indicate that caregivers are experiencing higher amounts of stress in the respective category. Scores ranging from the 15th to 80th percentile fall within normal limits, and scores above the 85th percentile are considered to be clinically significant. The PSI-4 has

been found to have good overall internal consistency (α = 0.96; Abidin, 2012) on the Total Stress and Difficult Child scales (α = .81-.89; Barroso et al., 2016; Lamis et al., 2014; Reitman, Currier, & Stickle, 2002), which are the two subscales of the PSI-4 that are most relevant to PCIT and examined in PCIT research (e.g., Garcia et al., 2021). Only the Difficult Child and Total Stress scores were used in this study. Caregivers completed the PSI-4: SF at pre- and post-treatment, and the Difficult Child and Total Stress percentiles were used to calculate change in caregiver stress from pre- to post-treatment.

Caregiver Positive and Corrective/Directive Child-Directed Statements. Caregiver positive and corrective/directive child-directed statements were measured using the Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV; Eyberg, Nelson, Ginn, Bhuiyan, & Boggs, 2013), a behavioral observation coding system that measures specific caregiver statements and interaction strategies used during three five-minute parent-child interactions in which the caregiver is instructed to first allow the child to lead, then take the lead of the play, and last instruct the child to clean up the toys. The reliability and validity of the DPICS-IV coding categories has been established across many studies (Eyberg et al., 2013). For the five minutes when caregivers follow the child's lead, caregiver verbalizations are coded for the frequency of positive statements (coded separately as labeled praises, behavior descriptions, and reflections) and corrective/directive statements (coded separately as questions, commands, and critical or sarcastic statements). For purposes of analysis, the dependent variables utilized are the summed values of the positive statements and the corrective/directive statements (coded only during this first five minutes in which the caregivers follow the child's lead).

In this study, coding was used to assess change in caregiver statements from pre- to post-treatment. DPICS-IV coding for this study was conducted live by PCIT therapists who were trained by a certified PCIT trainer to 80% coding reliability (or higher) using the DPICS-IV manual (Eyberg et al., 2013). Observation assessments were completed for pre-treatment at the second intake session. For families in the clinic-based sample or those completing IPCIT before COVID-19, the intake DPICS-IV observation took place in person in the clinic. For all families who completed intake after March 2020, this observation took place virtually, with the family in their home.

Treatment Satisfaction. The Therapy Attitude Inventory (i.e., TAI; Brestan, Jacobs, Rayfield, & Eyberg, 1999) is a 10-item caregiver report measure of caregiver's satisfaction with behavioral caregiver training. Given at the end of treatment, it is designed to assess how treatments like PCIT have affected change in the ways intended, specifically improving caregiver-child relationships, decreasing disruptive behaviors, and improving parenting self-efficacy. Caregivers rate ten items on a five-point scale where a score of one indicates that they did not gain strategies in the program or that the behavior worsened; a score of five indicates high satisfaction with the results of the treatment. High internal consistency has been demonstrated for this measure ($\alpha = .91$). In the current study, the TAI was used to evaluate average caregiver satisfaction with treatment, and to assess whether caregiver treatment satisfaction differed between the telehealth and clinic-based samples at post-treatment.

Table 2

Descriptive statistics for outcome measures in the total sample

Measure assessed at pre- and post-treatment	Total Sample						
	Pre-treatment			Post-treatment			
	M (SD)	Min.	Max.	*M* (SD)	Min.	Max.	Potential Range of measure
Frequency of child disruptive behavior	*146.192* (29.237)	50	229	*89.537* (26.673)	36	183	36-252
Quantity of child disruptive behavior	*20.015* (6.728)	0	36	*8.595* (7.866)	0	36	0-36
Child Externalizing Behavior	*64.299* (10.482)	35	96	*53.958* (9.411)	35	93	Within normal limits < 60; At-risk = 60-69; Clinically elevated ≥ 70
Child Adaptive Behavior	*42.320* (8.544)	18	69	*48.556* (8.531)	29	68	Within normal limits ≥ 40; At-risk = 30-39; Clinically elevated < 30
Child compliance	*48.390* (40.797)	0	100	*79.776* (32.095)	0	100	0-100
Caregiver stress about child behavior	*78.035* (20.859)	2	99	*55.813* (27.109)	2	99	2-99
Caregiver overall stress	*64.788* (21.617)	1	99	*46.195* (24.710)	1	99	1-99
Positive caregiver statements	*6.091* (6.924)	0	46	*28.735* (12.392)	1	59	1-59
Corrective/directive caregiver statements	*29.584* (17.551)	0	106	*4.792* (6.117)	0	48	0-48

Table 3

Descriptive statistics for outcome measures in the clinic-based sample

Measure assessed at pre- and post-treatment	Clinic-Based PCIT						
	Pre-treatment			Post-treatment			
	M (SD)	Min.	Max.	*M* (SD)	Min.	Max.	Potential Range of Measure
Frequency of child disruptive behavior	*147.311* (30.224)	50	229	*88.129* (27.286)	36	183	36-252
Quantity of child disruptive behavior	*19.770* (7.065)	0	36	*8.246* (8.411)	0	36	0-36
Child Externalizing Behavior	*63.221* (10.710)	35	96	*52.891* (9.606)	35	93	Within normal limits < 60; At-risk = 60-69; Clinically elevated ≥ 70
Child Adaptive Behavior	*41.851* (8.930)	18	69	*48.409* (8.872)	29	68	Within normal limits ≥ 40; At-risk = 30-39; Clinically elevated < 30
Child compliance	*45.076* (40.134)	0	100	*76.976* (34.138)	0	100	0-100
Caregiver stress about child behavior	*76.600* (21.960)	2	99	*53.685* (27.826)	2	99	2-99
Caregiver overall stress	*63.600* (22.433)	1	99	*45.600* (26.445)	1	99	1-99
Positive caregiver statements	*6.541* (7.422)	0	46	*30.397* (12.296)	1	59	0-no upper limit
Corrective/directive caregiver statements	*31.904* (18.838)	0	106	*4.475* (6.484)	0	48	0-no upper limit

Table 4

Descriptive statistics for outcome measures in the IPCIT sample

Measure assessed at pre- and post-treatment	IPCIT						
	Pre-treatment			Post-treatment			
	M (SD)	Min.	Max.	*M* (SD)	Min.	Max.	Potential Range of Measure
Frequency of child disruptive behavior	*143.794* (27.078)	66	202	*92.633* (25.217)	51	152	36-252
Quantity of child disruptive behavior	*20.540* (5.962)	7	34	*9.350* (6.532)	0	26	0-36
Child Externalizing Behavior	*66.540* (9.693)	45	91	*56.250* (8.613)	40	75	Within normal limits < 60; At-risk = 60-69; Clinically elevated ≥ 70
Child Adaptive Behavior	*43.318* (7.630)	30	60	*48.867* (7.825)	33	65	Within normal limits ≥ 40; At-risk = 30-39; Clinically elevated < 30
Child compliance	*55.604* (41.621)	0	100	*85.877* (26.378)	0	100	0-100
Caregiver stress about child behavior	*81.11* (18.060)	18	99	*60.317* (25.156)	6	96	2-99
Caregiver overall stress	*67.333* (19.693)	22	99	*47.433* (20.789)	4	87	1-99
Positive caregiver statements	*5.113* (5.622)	0	28	*25.143* (11.929)	3	57	0-no upper limit
Corrective/directive caregiver statements	*24.532* (13.130)	0	62	*5.482* (5.219)	0	28	0-no upper limit

Dummy Coding Categorical Family- and Treatment-Level Variables

To ensure that the clinic-based and IPCIT modality effects were the result of the interventions themselves, and not systematic differences in therapist education level, skill, or experience, we controlled for therapist effects in all analyses. Therapists were grouped by level of training in psychology, with those completing their predoctoral internship in one group (three therapists), those completing postdoctoral fellowship in another group (four therapists), masters-level clinicians in a third group (four therapists), and doctoral-level faculty in the fourth and final group (five therapists).

For categorical variables selected to be included as covariates in the analysis (i.e., caregiver race/ethnicity, therapist level of training, caregiver level of education achieved), dummy coded variables were created. Caregiver level of education was dummy coded with "some high school or high school degree" as the reference variable. Caregiver race/ethnicity was dummy coded with "Hispanic" as the reference variable. Level of therapist training was dummy coded with "intern" as the reference variable, and the language in which PCIT was delivered was coded as 0 = English, 1 = Spanish. Additionally, child sex (i.e., 0 = male, 1 = female) and whether the second caregiver was involved with treatment (0 = second caregiver was not involved, 1 = second caregiver was involved) were also dichotomously coded and included as treatment covariates.

Data Analysis

First, propensity score analyses were conducted to address potential selection bias when participants self-select into treatment groups (Bai & Clark, 2019). Next, to minimize the chances that results were influenced by shifts in caregiver stress, child disruptive behaviors, and difficult child/caregiver interactions experienced by some

families due to the global pandemic, preliminary analyses were conducted to test whether results differed for participants who self-selected into IPCIT vs. those whose only modality option was IPCIT due to COVID-related clinic shutdowns. Specifically, repeated measures ANOVA analyses were completed to establish whether receiving IPCIT during COVID (i.e., yes or no) influenced change on each outcome measure. Next, we examined whether families in IPCIT and those in clinic-based PCIT improved on intervention targets and outcomes (Hypothesis 1) by assessing changes in pre- to post-treatment measures for both groups separately. Specifically, paired *t* tests were conducted for IPCIT and clinic-based PCIT separately, comparing pre-treatment to post-treatment measures. Then, we examined whether these hypothesized improvements differed for families in IPCIT vs. those in clinic-based PCIT (Hypothesis 2). Specifically, nine identical hierarchical linear regressions were completed in which the dependent variables were all measures of post-treatment caregiver and child outcomes detailed in Table 5. Predictors were entered in four steps, with the first step including family-level variables, the second step including treatment-level variables, the third step including the pretreatment scores of the outcome variables, and the fourth step including treatment modality (i.e., 0 = clinic-based treatment, 1 = IPCIT). Propensity scores were added to the variables at the beginning of the regression equation before the first step to control for possible bias due to self-selection into groups. Finally, treatment satisfaction was examined using descriptive statistics to evaluate satisfaction with treatment in both treatment conditions. A more detailed description of the data analyses used to address each of these issues along with the corresponding results is included in the Results section below.

Table 5

Hypotheses and Corresponding Analytic Plan

Hypotheses/Goals	Analyses	Outcomes	Measure used
IPCIT and clinic-based PCIT will both produce significant changes in caregiver and child outcomes	Paired *t* tests comparing pre- and post-test scores for outcome variables of interest	**Frequency of child disruptive behavior**	Eyberg Child Behavior Inventory Intensity raw score
		Quantity of child disruptive behavior	Eyberg Child Behavior Inventory Problem raw score
		Child externalizing behavior	The Behavior Assessment System for Children, Third Edition, Parent Rating Scale Externalizing Problems composite *T* score
		Child adaptive behavior	The Behavior Assessment System for Children, Third Edition, Parent Rating Scale Adaptive Behaviors Composite *T* score
		Child Compliance	The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV) Percentage of times child complied with an effective direct command during caregiver-led play
		Caregiver stress about child behavior	Parenting Stress Index – 4th edition, Short Form Difficult Child Percentile
		Caregiver overall stress	Parenting Stress Index – 4th edition, Short Form Total Stress Percentile
		Positive caregiver child-directed statements	The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV) Composite of Behavior Descriptions, Reflections, and Labeled praises coded during child-led play
		Directive/corrective caregiver child-directed statements	The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV)

			Composite of Questions, Commands, and Criticisms coded during child-led play
IPCIT and clinic-based PCIT will not differ in their effect on caregiver and child outcomes	**Hierarchical Linear Regression** 9 separate regressions completed with the DVs listed to the right *Propensity score is entered as a control variable Step 1: Family-level variables *(child age, child sex, caregiver highest level of education, caregiver race/ethnicity, involvement of second caregiver in treatment)* Step 2: Treatment-level variables *(language of treatment, level of therapist training)* Step 3: Intake assessment of all 9 outcome variables: *1) ECBI Intensity raw score* *2) ECBI Problem raw score* *3) BASC-3 Externalizing T score* *4) BASC-3 Adaptive behavior T score* *5) Child Compliance* *6) PSI-4:SF Total Stress percentile* *7) PSI-4: SF Difficult Child Percentile* *8) Caregiver composite positive statements* *9) Caregiver composite corrective/directive statements* Step 4: Treatment Sample Group (i.e., IPCIT or clinic-based treatment)	**Frequency of child disruptive behavior**	Eyberg Child Behavior Inventory Intensity raw score
		Quantity of child disruptive behavior	Eyberg Child Behavior Inventory Problem raw score
		Child externalizing behavior	The Behavior Assessment System for Children, Third Edition, Parent Rating Scale Externalizing Problems Composite *T* score
		Child adaptive behavior	The Behavior Assessment System for Children, Third Edition, Parent Rating Scale Adaptive skills composite *T* score
		Child compliance	The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV) Percentage of times child complied with an effective direct command during caregiver-led play
		Caregiver stress about disruptive child behavior	Parenting Stress Index – 4th edition, Short Form Difficult Child Percentile
		Caregiver overall stress	Parenting Stress Index – 4th edition, Short Form Total Stress Percentile
		Positive caregiver child-directed statements	The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV) Composite of Behavior Descriptions, Reflections, and Labeled praises coded during child-led play
		Directive/corrective caregiver child-directed statements	The Dyadic Parent-Child Interaction Coding System, Fourth Edition (DPICS-IV) Composite of Questions, Commands, and Criticisms coded during child-led play

Chapter IV: Results

Tables 2, 3, and 4 provide descriptive statistics for all study variables. Analyses are presented below in the order they are discussed in the analytic plan. First, a brief description of preliminary zero-order correlations among study variables is provided. Then, description of propensity score calculations is provided and sensitivity analyses examining whether differences between IPCIT and clinic-based PCIT are potentially due to the COVID-19 pandemic are presented. Next, substantive study hypotheses are explored. Aligning with the first hypothesis, changes in child and caregiver outcomes from pre- to post-treatment are explored separately in the IPCIT and clinic-based PCIT treatment modalities. Then, aligning with the second hypothesis, we examined whether changes in these outcomes differ between the IPCIT and clinic-based PCIT treatment modalities. Finally, we examined treatment satisfaction in both treatment modalities.

Both tests of statistical significance and effect size estimates are provided and were used together to evaluate the hypotheses. A central goal of the current study was to test a null hypothesis (i.e., that there is no statistically significant difference between two treatment groups), but statistical significance (or lack thereof) is highly dependent on sample size. Too large of a sample size may reveal a statistically significant but clinically meaningless result (Ferguson, 2009). Too small of a sample may hinder the detection of differences between treatment groups, such as when results demonstrate treatment group differences that are clinically meaningful, but where statistical significance is not achieved because the study is underpowered to detect small, but potentially meaningful effects, potentially leading to erroneous conclusions. Given the sample size differences between treatment modalities, as well as the low power for some analyses (particularly

those investigating only the IPCIT group), effect size estimates can demonstrate a more accurate measure of true effect in some cases than statistical significance alone (Ferguson, 2009; Wilkinson & Task Force on Statistical Inference, 1999).

In the current study, effect sizes were calculated as Cohen's *d* for paired *t* test calculations, with the cutoff values of 0.2 for small, 0.5 for medium and 0.8 for large effects (Cohen, 1992). Cohen's *d* was calculated using a variety of techniques, and the most conservative estimate is reported here. ANOVA effect sizes were calculated and presented as partial eta squared, with cutoff values of 0.01 for small, 0.09 for medium, and 0.25 for large. Effect sizes for the hierarchical linear regressions were calculated as R^2 reflected at each step in the R^2 change value, with the cutoff values of 0.02 for small, 0.13 for medium, and 0.26 for large. Effect sizes that failed to meet the threshold for a small effect will be described as minimal.

Correlations Among Study Variables

Intercorrelations among study variables are presented in Table 6. Correlations among many study variables provided preliminary support for several of the analytic steps outlined above. Significant positive associations of numerous covariates with treatment modality and/or specific outcome variables (i.e., post-treatment measures of each assessment) provides evidence for the importance of using propensity scores, as well as justification for the inclusion of each child-, caregiver-, and treatment-level variable included as covariates and additional predictors in subsequent analyses. For instance, therapist training at both the postdoctoral (significantly positively associated with use of IPCIT treatment) and masters' level (significantly negatively correlated with use of IPCIT treatment) were significantly correlated with treatment modality, as was

caregiver education (significantly positively correlated with IPCIT treatment). Additionally, use of the IPCIT treatment modality was associated with significantly higher pre-treatment child externalizing behavior, and significantly lower pre-treatment caregiver corrective/directive statements, suggesting that the difference in outcomes between treatment modalities needs to be further explored in more rigorous analyses.

Calculating Propensity Scores

Propensity scores are used for addressing selection bias in observational studies when the method of selection could create differences between groups but is not something that can necessarily be managed by controlling only one variable (Bai & Clark, 2019). The use of propensity scores is a recommended method for increasing internal validity when participants self-select into one condition or another, as there may be many pre-existing differences between groups, but they may have self-selected into specific groups for reasons that are not possible to manipulate experimentally, like where someone lives, or whether a participant has reliable access to transportation (both of which also may be influenced by income and ethnicity; Bai & Clark, 2019). Similarly, the use of propensity scores is common when circumstances preclude participants from being randomly assigned to one condition or the other. For example, propensity scores are often used to control differences between groups when one group has a rare diagnosis for which the incidence would not allow a large enough sample size to complete a randomized control trial. In the current study, this analysis allows families to be included in the IPCIT sample who would not be able to attend treatment in person, and to include families who completed treatment during COVID-19, when in-person treatment was not

possible, and families would therefore not have been able to be randomly assigned to treatment condition.

Propensity scores are calculated using a logistic regression model in which the grouping variable (i.e., describing the groups into which participants self-selected) is the dependent variable (in this case, treatment group: IPCIT vs. clinic-based PCIT). The covariates chosen for the analysis are somewhat at the discretion of the researcher, but it is recommended that they include at minimum all the available variables that are significantly related to either treatment group and/or the outcome variables (Bai & Clark, 2019). Intake measures of all outcome variables were included in the propensity score analysis in this study regardless of their impact on the treatment group, as covariates that significantly influence outcome variables are required to be included in the propensity score model (Bai & Clark, 2019; Brookhart et al., 2006; Rubin & Thomas 1996). Propensity scores are created for all participants with complete data (i.e., no missing data on predictors or outcome).

This logistic regression is used to calculate a propensity score for each participant, which is a new variable with a value between 0 and 1, indicating the probability of being placed in the group coded 1 on the outcome – conditional on the covariates included in the model. There are several ways to utilize the calculated propensity scores for controlling for selection bias. The strategy that conserves the most participants is to include the propensity score as a adjustment variable in subsequent analyses. Including propensity scores in this way conserves participants because it ensures that participants are not required to exactly "match" propensity scores across group (e.g., finding a participant with a .05 propensity score in the IPCIT group, and pairing them for

comparison with a participant with a .05 propensity score in the clinic-based PCIT group; Bai & Clark, 2019). When such "matching" is done, participants are not included in the analyses if they do not have an appropriate match within the other treatment group. Therefore, an exact matching process eliminates participants that would otherwise be included (i.e., because they are not missing data), thus needlessly discarding valuable data and drastically decreasing the sample size (Bai & Clark, 2019; Lane, To, Shelley, & Henson, 2012).

In the current study, in accordance with best practices for propensity score calculation (Bai & Clark, 2019), propensity scores were calculated (i.e., as a new variable for each participant) using a logistic regression model with treatment modality as the dependent variable (i.e., IPCIT vs. clinic-based PCIT). First, covariates were selected which were significantly correlated with treatment modality and/or one of the outcome variables. Predictors utilized to calculate the propensity score included family-level variables (e.g., caregiver level of education and race/ethnicity, child age and sex, involvement of a second caregiver in treatment), treatment-level variables (e.g., language of service delivery and level of therapist training), and pre-treatment assessments (i.e., frequency and quantity of disruptive child behavior [ECBI Intensity and Problem raw scores, respectively], child externalizing behavior [BASC-3 Externalizing composite *T* score], child adaptive behavior [BASC-3 Adaptive Behavior composite *T* score], child compliance [DPICS-IV child compliance], caregiver stress about child behavior and overall stress [PSI-4: SF Difficult Child percentile and Total Stress percentile, respectively], and caregiver positive and corrective/directive statements [DPICS-IV composite caregiver positive and corrective child-directed statements, respectively]). This

logistic regression produced a propensity score for each child (i.e., a score between 0 and 1) that indicated the child's probability of being placed in the telehealth treatment group (i.e., which was coded 1) based upon the variables included in the analysis (Guo & Fraser, 2010; Thoemmes & Kim, 2011).

Table 6 displays the final model of likelihood of treatment modality, or the significance of each covariate in predicting whether the families completed either IPCIT or clinic-based treatment. Within this model, the language of treatment delivery was a significant predictor of treatment group assignment, with Spanish-speaking families more likely to receive clinic-based services. Neither level of therapist training (postdoc trainees, masters trainees, faculty therapists), child age, caregiver level of education (some college, associate degree, bachelor's degree, graduate degree), child sex nor involvement of a second caregiver were significantly associated with treatment modality when included in the full model. Despite the lack of significance, several variables had an odds ratio of more than two (see Table 6), indicating that participants in those groups were more likely to receive IPCIT than clinic-based services (based upon OR, but not corroborated by statistical significance). Specifically, caregivers who identified their race/ethnicity as African American, those in the Other ethnicity category, caregivers with bachelor's degrees, and those with graduate degrees were more likely to receive treatment virtually.

Among intake measures included as factors in the logistic regression to produce the propensity scores, only frequency of child disruptive behavior emerged as significant. That is, the higher the caregiver reported frequency of child disruptive behavior at intake,

the more likely the family was to have received IPCIT treatment. Other intake measurements were not significantly related to treatment modality.

Table 6

Correlation table comparing demographic family- and treatment-level variables, intake measures for child and caregiver, and treatment modality

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Child age	-																						
2. Child sex	-.075	-																					
3. Treat. Lang.	-.071	.031	-																				
4. C2	.110	.010	-.084	-																			
5. PDTT	.116	.174*	-.229+	.064	-																		
6. TT-M	-.061	-.197+	.057	-.075	-.611+	-																	
7. TT - F	-.065	.008	.424+	.025	-.316+	-.242+	-																
8. CE – SC	-.020	-.080	-.038	-.096	-.083	.112	-.131	-															
9. CE – AD	.004	-.139	.164*	-.052	-.166*	.080	.123	-.135	-														
10. CE – BD	.010	.147*	-.037	.010	.064	-.033	-.027	-.248+	-.242+	-													
11. CE – GD	-.097	-.027	-.077	.103	.148*	-.133	-.033	-.241+	-.274+	-.504+	-												
12. CEth - Other	.107	-.019	-.088	.017	.138	-.024	-.086	-.023	.050	-.066	.000	-											
13. CEth–WNH	-.009	-.004	-.199+	-.007	.182*	-.196+	-.080	-.168*	-.050	-.004	.181*	-.133	-										
14. CEth – AA	.053	-.093	-.079	-.136	.049	.007	-.077	.216+	-.079	-.093	-.064	-.053	-.120	-									
15. DCB FQ	.018	.058	-.181*	.077	.041	.070	-.176*	.165*	-.035	-.172+	.058	-.002	-.049	.008	-								
16. DCB Q	.147*	.127	-.153*	.059	.053	-.005	-.082	.161*	.025	-.103	.058	-.017	-.042	-.048	.708+	-							
17. EXT	.016	.189+	-.220+	.050	.092	-.100	-.131	-.063	-.055	-.033	.042	-.096	.121	.027	.487+	.411+	-						
18. ADAPT	.014	-.187+	.231+	-.014	-.035	-.130	.209+	.041	.090	-.163*	.098	-.043	.010	.006	-.287+	-.158+	-.231+	-					
19. COMP	.047	.031	.199+	-.113	.036	-.010	.050	-.039	.004	.006	-.029	.019	-.095	.051	-.072	-.011	-.071	-.009	-				
20. DIFF	-.022	.070	-.225+	.020	.147*	-.018	-.238+	-.061	.015	-.006	.059	-.076	.176*	-.033	.454+	.413+	.500+	-.347+	-.102	-			
21. TS	-.002	.099	-.243+	.789	.113	.019	-.263+	-.101	.005	.036	.003	-.043	.134	-.041	.378+	.376+	.463+	-.383+	-.052	.851+	-		
22. POS	-.278+	.004	-.064	.005	-.037	.002	.021	.011	-.128	.009	.115	-.070	.003	.008	-.165*	-.229+	-.081	.127	-.002	-.162*	-.127	-	
23. COR/ DIR	-.337+	-.081	.135	-.018	-.140	.064	.100	.046	.069	.065	-.067	-.032	-.087	-.048	.034	-.086	-.120	-.105	-.054	-.121	-.189+	.036	-

24. TM	.159*	.108	-.011	.035	.306^{+}	-.257^{+}	-.104	-.088	-.112	-.057	.227^{+}	.071	.035	.111	-.056	.053	.149*	.080	.093	.101	.081	-.096	-.196^{+}

Note: CE: Caregiver education; C2: Involvement of second caregiver in treatment; SC: Some college; AD: Associate's degree; BD: Bachelor's degree; GD: Graduate degree CEth: Caregiver ethnicity; WNH: White Non-Hispanic; AA: African American; TT: Therapist training; PD: Postdoctoral; M: Master's level therapist; F: Faculty; DCB: Difficult Child Behaviors; FQ: Frequency; Q: Quantity; EXT: Child externalizing behaviors; ADAPT: Child adaptive behaviors; COMP: Child compliance; DIFF: Caregiver stress related to child behavior (i.e., measured by the PSI-4:SF Difficult child percentile); TS: Caregiver total stress (i.e., the PSI-4:SF Total Stress percentile); POS: Caregiver positive child-directed statements; COR/DIR: Caregiver corrective/directive child-directed statements; TM: Treatment modality (i.e., IPCIT or clinic)
All measures included (variables 15-23) are the intake measurement of each.
*$p < .05$, $^{+}p < .001$

Table 7

Model of likelihood predicting group assignment to IPCIT or clinic-based treatment

Predictors	B	*SE*	OR	*p*
Caregiver Race/Ethnicity[a]				
White non-Hispanic	-.219	.477	.617	.300
Other ethnicity	1.227	.908	2.315	.335
African American	1.530	1.064	3.267	.258
Highest level of caregiver education received[b]				
Some college	.456	1.008	1.437	.711
Associate degree	-.539	1.091	1.093	.928
Bachelor's degree	.746	.809	2.163	.323
Graduate degree	1.557	.813	5.348	.033
Second caregiver involved in treatment	.355	.410	1.238	.586
Child sex	.294	.440	1.065	.884
Child age	.190	.147	1.161	.294
Language PCIT delivered	1.549	.768	3.289	.102
Clinician level of training[c]				
Postdoctoral trainees	.889	.615	1.728	.335
Masters level clinicians	-.671	.729	.376	.142
Faculty therapists	-.903	.937	.281	.154
Intake assessments				
Child disruptive behavior frequency	-.025	.011	.976	.024
Child disruptive behavior quantity	.063	.045	1.047	.289
Child externalizing behaviors	.037	.025	1.048	.055
Child adaptive behavior	.032	.028	1.024	.376
Child compliance	.008	.005	1.006	.179
Caregiver stress about child behavior	.027	.020	1.016	.401
Caregiver total stress	-.008	.018	.995	.773
Caregiver positive statements	-.008	.030	.973	.352
Caregiver corrective/directive statements	-.015	.014	.984	.241

Note: Model Chi-square was significant; $X^2(26) = 47.804$, $p = .006$.
Model Cox & Snell $R^2 = .256$
[a]Dummy coded with reference variable caregiver race/ethnicity = Hispanic
[b]Dummy coded with reference variable caregiver highest level of education = some high school or high school degree
[c]Dummy coded with reference variable clinician level of training = predoctoral intern

Differences Related to COVID-19

Separate analyses were conducted to test whether any caregiver and child outcome differences were a product of treatment modality rather than an artifact of completing treatment before or during a global pandemic. Other studies have demonstrated that the stress of the COVID-19 pandemic (caused by a multitude of factors including job loss, racial trauma, lack of social contact, and remote schooling) adversely impacted child behavior and mental health and caregiver stress worldwide. Although individual caregivers in this sample were not assessed before and after the pandemic, based upon other studies it is reasonable to consider that families who received treatment during the pandemic may have outcomes discrepant from those who received treatment before the pandemic. Therefore, was necessary to determine that change over treatment did not differ between pre- and during-pandemic families. Within only the IPCIT sample, mixed model ANOVA analyses (i.e., with one between-subjects and one within-subjects factor) were completed comparing families who completed IPCIT before COVID-19 and those who completed it during the pandemic. Separate ANOVA analyses were completed to examine the change over time for frequency and quantity of child disruptive behaviors, child externalizing behavior, child adaptive behavior, child compliance, caregiver total stress and stress about child behavior, and caregiver positive and corrective/directive statements. Covariates included family-level variables (i.e., caregiver race/ethnicity, child sex, child age, caregiver education level completed) and treatment-level variables (i.e., language of treatment, level of therapist training, involvement of second caregiver). The *F* ratio reported below reflects the group by time interaction for each analysis.

None of these ANOVA models demonstrated a significant difference in outcomes between families who received IPCIT before COVID-19 and those who received it during the pandemic. No differences were found between families completing IPCIT during or before the COVID-19 pandemic in frequency (F[1, 52] = 1.285, p = .262) and quantity of child disruptive behavior (F[1, 52] = 2.063, p = .157), child externalizing behavior (F[1, 52] = .153, p = .698), child adaptive behavior (F[1, 52] = .144, p = .706), child compliance (F[1, 47] = 1.337, p = .253), caregiver stress about child behavior (F[1, 52] = .194, p = .662), caregiver total stress (F[1, 52] = 1.298, p = .260), caregiver positive statements (F[1, 47] = .752, p = .390) or caregiver directive/corrective statements (F[1, 47] = .346, p = .559). Power calculations revealed low power for these analyses (see Table 8), likely due to small sample size. As a result, effect size estimates were also calculated as partial eta squared. The majority of effect sizes did not meet the threshold for a small effect (i.e., they were below .01), with the exception of the quantity of child disruptive behavior, child compliance, and overall caregiver stress (which demonstrated small effects; i.e., above 0.01 but below 0.06). See Table 8 for means, standard deviations, and sample size for each group.

Change in Child and Caregiver Outcomes from Pre- to Post-Treatment

It was hypothesized that families in both IPCIT and clinic-based PCIT would exhibit improvements in child and caregiver behavioral outcomes and caregiver stress from pre- to post-treatment. To test this hypothesis, paired *t*-tests were computed separately for the IPCIT and clinic-based modalities. Table 9 presents the means, standard deviations, and effect sizes for paired *t*-test analyses of child and caregiver behavioral outcomes at the pre- and post-treatment assessments. Paired *t*-tests

demonstrated hypothesized improvements across outcome measures. More specifically, paired *t*-tests demonstrated improvements in frequency of child disruptive behavior (clinic-based: $t(136) = 21.35$, $p < .001$; IPCIT: $t(54) = 13.08$, $p < .001$), quantity of child disruptive behavior (clinic-based: $t(134) = 15.03$, $p < .001$; IPCIT: $t(54) = 10.74$, $p < .001$), child externalizing behavior (clinic-based: $t(129) = 11.765$, $p < .001$; IPCIT: $t(54) = 9.50$, $p < .001$), child adaptive behavior (clinic-based: $t(130) = -9.831$, $p < .001$; IPCIT: $t(54) = -5.99$, $p < .001$), child compliance (clinic-based: $t(125) = -7.355$, $p < .001$; IPCIT: $t(51) = -5.263$, $p < .001$), caregiver stress about child behavior (clinic-based: $t(131) = 9.37$, $p < .001$; IPCIT: $t(54) = 7.48$, $p < .001$), caregiver overall stress (clinic-based: $t(129) = 7.84$, $p < .001$; IPCIT: $t(54) = 7.73$, $p < .001$), caregiver positive statements (clinic-based: $t(124) = -21.405$, $p < .001$; IPCIT: $t(50) = -12.50$, $p < .001$), and caregiver corrective/directive statements (clinic-based: $t(125) = 16.48$, $p < .001$; IPCIT: $t(50) = 10.07$, $p < .001$). Effect sizes (i.e., Cohen's *d*) for all *t* tests were large, with the exception of child compliance, which was medium for both clinic-based and IPCIT treatment. See Table 9 for specific effect sizes.

Table 8

ANOVA analyses examining the difference between IPCIT patients before COVID-19 and those during COVID-19

Outcomes	Before or During COVID-19	*N*	Intake *M* (SD)	Post *M* (SD)	Sum of Squares	Partial Eta Squared*
Frequency of child disruptive behavior	Before	15	*139.800* (33.089)	*90.400* (23.173)	333.532	.009
	During	45	*147.200* (24.164)	*93.378* (26.068)		
Quantity of child disruptive behavior	Before	15	*18.667* (7.471)	*8.533* (6.128)	36.973	.019
	During	45	*21.311* (5.401)	*9.622* (6.706)		
Child externalizing behaviors	Before	15	*68.200* (9.151)	*57.200* (6.581)	16.704	.003
	During	45	*66.533* (9.694)	*55.933* (9.235)		
Child adaptive behaviors	Before	15	*42.933* (8.172)	*47.267* (5.284)	4.971	.001
	During	45	*43.267* (7.724)	*49.400* (8.489)		
Child Compliance	Before	15	*34.645* (38.969)	*88.698* (17.926)	3289.388	.059
	During	40	*60.010* (42.104)	*84.467* (29.252)		
Caregiver stress about child behavior	Before	15	*79.867* (21.735)	*61.267* (22.250)	73.759	.002
	During	45	*81.378* (17.329)	*60.000* (26.280)		
Overall caregiver stress	Before	15	*63.200* (22.517)	*43.933* (18.603)	549.989	.020
	During	45	*68.511* (18.350)	*48.600* (21.537)		
Caregiver positive statements	Before	14	*7.071* (7.539)	*26.643* (16.420)	40.978	.008
	During	41	*4.537* (4.985)	*24.610* (10.342)		
Caregiver corrective/ directive statements	Before	14	*23.071* (14.557)	*25.634* (13.386)	20.155	.006
	During	41	*4.643* (3.915)	*5.878* (5.618)		

Note: *Partial eta squared is included as an effect size estimate. Small effect size = 0.01. Medium effect size = 0.06. Large effect size = 0.14. Some overall sample sizes here exceed the IPCIT sample of 58 used for the wider analyses. Propensity scores in the wider analysis excluded some participants without complete data in all variables that contributed to the calculated propensity score variable.

Table 9

Paired t tests of the change in caregiver skill and child behavior from intake to post-treatment for IPCIT and clinic-based PCIT families

Measures assessed at pre- and post-treatment	Session Location	*N*	Pre-treatment scores *M (SD)*	Post-treatment scores *M* (*SD*)	df	*t*	*d**
Child disruptive behavior frequency	IPCIT	55	145.69 (27.17)	92.45 (25.07)	54	13.08*	1.764
	Clinic	137	147.12 *(30.20)*	88.37 *(27.29)*	136	21.35*	1.824
Child disruptive behavior quantity	IPCIT	55	20.89 (6.08)	9.51 (6.63)	54	10.74*	1.448
	Clinic	135	19.53 *(7.01)*	8.22 *(8.31)*	134	15.03*	1.293
Child externalizing behavior	IPCIT	55	66.80 (9.27)	56.42 (7.90)	54	9.50*	1.280
	Clinic	130	63.05 *(10.83)*	53.00 *(9.79)*	129	11.765*	1.032
Child adaptive behavior	IPCIT	55	43.13 (7.83)	48.65 (7.82)	54	-5.99*	-0.808
	Clinic	131	41.99 *(8.92)*	48.42 *(8.80)*	130	-9.831*	-0.859
Child compliance	IPCIT	55	54.64 (42.21)	87.92 (22.97)	54	-5.263*	-0.671
	Clinic	122	43.81 *(39.16)*	76.32 *(34.72)*	121	-7.501*	-0.679
Caregiver stress about child behavior	IPCIT	55	81.82 (18.26)	61.71 (24.35)	54	7.48*	1.009
	Clinic	132	75.51 *(22.02)*	53.36 *(27.90)*	131	9.37*	0.815
Total caregiver stress	IPCIT	55	67.84 (19.26)	48.78 (20.45)	54	7.73*	1.043
	Clinic	130	62.47 *(22.69)*	45.10 *(26.30)*	129	7.84*	0.687
Caregiver positive statements	IPCIT	51	5.43 (5.90)	25.20 (11.80)	50	-12.50*	-1.751
	Clinic	125	6.30 *(7.09)*	30.20 *(12.43)*	124	-21.405*	-1.915
Caregiver corrective statements	IPCIT	51	24.65 (12.76)	5.57 (5.27)	50	10.07*	1.409
	Clinic	126	32.06 *(19.47)*	4.51 *(6.44)*	125	16.48*	1.469

Note: *$p < .001$. Cohen's *d* is included as an effect size estimate. Small effect = 0.2; medium effect = 0.5; large effect = 0.8.

IPCIT and Clinic-Based PCIT Comparison of Treatment Outcomes

It was also hypothesized that the clinic-based and IPCIT modalities would produce similar improvements in treatment outcomes. To address this hypothesis, nine identical hierarchical linear regressions were completed in which the dependent variables were all measures of post-treatment caregiver and child outcomes detailed in Table 5. Predictors for the hierarchical regression were entered in steps, with the first step including family-level variables, the second step including treatment-level variables, and the third step including the pretreatment scores of the outcome variables (see Table 5 for specific variables entered at each step). Finally, the predictor of interest was added to the analysis – IPCIT or clinic-based treatment modality (i.e., 0 = clinic-based treatment, 1 = IPCIT) – at the fourth and final step of the hierarchical regression model. Propensity scores were added to the variables at the beginning of the regression equation before the first step to control for possible bias due to self-selection into groups. All of these individual variables were retained at each step due to their collective inclusion in the propensity score analysis. In accordance with best practices for increasing the robustness of a propensity score adjustment, it is recommended that any variables included in the original analysis to calculate the scores also be included in the analyses as individual predictors where the propensity scores are utilized for adjustment for selection bias (Bai & Clark, 2019; Schafer & Kang, 2008).

Table 10

Regression analyzing post-treatment frequency of disruptive child behaviors comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 177) = 1.483$, $p = .149$, $R^2 = .077$			
CE – Some college[a]	-8.919	8.808	-.108	.313
CE – Associate degree[a]	3.570	8.667	.042	.681
CE – Bachelors degree[a]	2.668	7.383	.047	.718
CE – Graduate degree[a]	9.190	7.383	.167	.215
Caregiver 2 involvement	-2.838	3.853	-.053	.462
CR – White non-Hispanic[b]	-1.112	4.734	-.018	.815
CR – Other ethnicity[b]	4.051	9.068	.033	.656
CR – African-American[b]	11.816	9.862	.090	.233
Child age	-1.015	1.425	-.057	.477
Child sex	5.965	4.459	.103	.183
Step 2: Treatment-level variables				
	$\Delta F(4, 173) = .945$, $p = .440$, $\Delta R^2 = .020$			
Language PCIT delivered	3.918	6.997	.046	.576
Postdoctoral therapist[c]	-9.635	6.116	-.180	.117
Masters level therapist[c]	-11.714	6.412	-.207	.070
Faculty therapist[c]	-10.332	8.428	-.120	.222
Step 3: Intake measures				
	$\Delta F(9, 164) = 3.560$, $p < .001$, $\Delta R^2 = .148$			
Child disruptive behavior frequency	.307	.105	.336	.004
Child disruptive behavior quantity	.033	.423	.008	.939
Child externalizing behavior	.397	.228	.157	.084
Child adaptive behavior	-.087	.259	-.028	.738
Child compliance intake	.029	.048	.044	.545
Caregiver stress about child behavior	-.221	.182	-.173	.226
Total caregiver stress	.108	.172	.087	.531
Caregiver intake positive statements	.018	.284	.005	.950
Caregiver intake directive/corrective statements	.031	.120	.020	.798
Step 4: Treatment modality				
	$\Delta F(1, 163) = .044$, $p < .835$, $\Delta R^2 = .044$			
IPCIT or clinic-based PCIT	.999	4.776	.017	.835

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 187) = 2.202$, $p = .002$, $R^2 = .245$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Frequency of Child Disruptive Behavior

Table 10 describes results for analyses examining whether post-treatment frequency of child disruptive behavior varied by treatment modality. Child-, caregiver-, and treatment-level variables (i.e., Steps 1 and 2) did not explain a significant amount of variance in post-treatment frequency of child disruptive behavior, and furthermore, effect sizes were small. Pre-treatment assessments (Step 3) explained a significant amount of variance, and exhibited a medium effect size, with pre-treatment assessment of frequency of child disruptive behavior achieving statistical significance. As hypothesized, post-treatment frequency of child disruptive behavior did not vary by treatment modality; results were nonsignificant, treatment modality did not explain a significant amount of variance in the outcome, and the effect size was small (see Table 10).

Quantity of Disruptive Child Behavior

Table 11 describes results for analyses examining whether post-treatment quantity of child disruptive behavior varied by treatment modality. Child- caregiver-, and treatment-level variables (i.e., Steps 1 and 2) did not explain a significant amount of variance in post-treatment quantity of child disruptive behavior. Effect size was small for child- and caregiver-level variables, and the effect size for treatment-level variables was minimal and below the threshold for a small effect size. Pre-treatment assessments (i.e., Step 3) explained a significant amount of variance and exhibited a medium effect size, with pre-treatment quantity of child disruptive behaviors achieving statistical significance. As hypothesized, post-treatment quantity of child disruptive behavior did not vary by treatment modality (i.e., Step 4); results were nonsignificant, treatment

modality did not explain a significant amount of variance, and the effect size was minimal (see Table 11).

Child Externalizing Behaviors

Table 12 describes results for analyses examining whether post-treatment child externalizing behaviors varied by treatment modality. Child- and caregiver-level variables (i.e., Step 1) explained a significant amount of variance in post-treatment child externalizing behaviors and exhibited a medium effect size, with child sex achieving statistical significance. Treatment-level variables (Step 2) did not explain a significant amount of variance and exhibited a small effect size. Pre-treatment assessments (i.e., Step 3) explained a significant amount of variance, and demonstrated a medium effect size, with pre-treatment child externalizing behavior achieving statistical significance. As hypothesized, post-treatment child externalizing behavior did not vary by treatment modality (i.e., Step 4); results were nonsignificant, treatment modality did not explain a significant amount of variance, and the effect size was minimal (see Table 12).

Table 11

Regression analyzing post-treatment quantity of disruptive child behaviors comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 175) = .904$, $p = .531$, $R^2 = .049$			
CE – Some college[a]	-3.413	2.711	-.140	.210
CE – Associate degree[a]	.881	2.703	.035	.745
CE – Bachelors degree[a]	-.313	2.290	-.019	.892
CE – Graduate degree[a]	.990	2.302	.061	.668
Caregiver 2 involvement	-1.056	1.174	-.066	.369
CR – White non-Hispanic[b]	-.027	1.451	-.001	.985
CR – Other ethnicity[b]	5.015	2.742	.137	.069
CR – African American[b]	4.611	2.985	.119	.124
Child age	-.271	.439	-.051	.538
Child sex	1.575	1.366	.092	.251
Step 2: Treatment-level variables				
	$\Delta F(4, 171) = .494$, $p = .740$, $\Delta R^2 = .011$			
Language PCIT delivered	1.480	2.148	.060	.492
Postdoctoral therapist[c]	-1.622	1.850	-.103	.382
Masters level therapist[c]	-2.319	1.939	-.139	.234
Faculty therapist[c]	-3.097	2.652	-.117	.245
Step 3: Intake measures				
	$\Delta F(9, 162) = 3.442$, $p = .001$, $\Delta R^2 = .151$			
Child disruptive behavior frequency	-.035	.032	-.129	.279
Child disruptive behavior quantity	.483	.129	.416	< .001
Child externalizing behavior	.136	.069	.182	.052
Child adaptive behavior	.077	.079	.983	.327
Child compliance intake	.001	.015	.006	.940
Caregiver stress about child behavior	.003	.055	.007	.963
Total caregiver stress	-.005	.052	-.013	.927
Caregiver intake positive statements	.025	.086	.022	.772
Caregiver intake directive/corrective statements	.024	.036	.054	.506
Step 4: Treatment modality				
	$\Delta F(1, 161) = .336$, $p = .563$, $\Delta R^2 = .002$			
IPCIT or clinic-based PCIT	-.839	1.448	-.049	.563

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 185) = 1.810$, $p = .017$, $R^2 = .213$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Table 12

Regression analyzing post-treatment child externalizing behavior comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 173) = 2.760$, $p = .003$, $R^2 = .138$			
CE – Some college[a]	-.147	2.769	-.005	.958
CE – Associate degree[a]	3.015	2.783	.098	.280
CE – Bachelors degree[a]	2.224	2.326	.110	.340
CE – Graduate degree[a]	3.606	2.315	.188	.121
Caregiver 2 involvement	-1.098	1.219	-.058	.369
CR – White non-Hispanic[b]	1.127	1.500	.051	.454
CR – Other ethnicity[b]	3.264	2.839	.075	.252
CR – African American[b]	5.824	3.092	.127	.061
Child age	.279	.455	.045	.541
Child sex	3.725	1.409	.183	.009
Step 2: Treatment-level variables				
	$\Delta F(4, 169) = 1.035$, $p = .391$, $\Delta R^2 = .021$			
Language PCIT delivered	.898	2.197	.031	.683
Postdoctoral therapist[c]	-.617	1.928	-.033	.749
Masters level therapist[c]	-1.331	2.010	-.067	.509
Faculty therapist[c]	-1.990	2.700	-.065	.462
Step 3: Intake measures				
	$\Delta F(9, 160) = 7.365$, $p < .001$, $\Delta R^2 = .247$			
Child disruptive behavior frequency	.023	.033	.070	.496
Child disruptive behavior quantity	-.034	.132	-.024	.800
Child externalizing behavior	.485	.072	.545	< .001
Child adaptive behavior	.034	.082	.032	.674
Child compliance intake	-.001	.015	-.004	
Caregiver stress about child behavior	-.052	.058	-.115	.372
Total caregiver stress	.017	.054	.039	.756
Caregiver intake positive statements	-.008	.089	-.006	.926
Caregiver intake directive/corrective statements	-.018	.038	-.034	.631
Step 4: Treatment modality				
	$\Delta F(1, 159) = .002$, $p = .960$, $\Delta R^2 < .001$			
IPCIT or clinic-based PCIT	.075	1.508	.004	.960

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 183) = 4.506$, $p < .001$, $R^2 = .405$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Child Adaptive Behavior

Table 13 describes results for analyses examining whether post-treatment child adaptive behavior varied by treatment modality. Child- and caregiver-level variables (i.e., Step 1) explained a significant amount of variance in post-treatment child adaptive behavior and exhibited a small effect size. Treatment-level variables (i.e., Step 2) did not explain a significant amount of variance and demonstrated a small effect size, with language of treatment delivery achieving statistical significance. Pre-treatment assessments (i.e., Step 3) explained a significant amount of variance, and demonstrated a large effect size, with pre-treatment child adaptive behavior achieving statistical significance. As hypothesized, post-treatment child adaptive behavior did not vary by treatment modality (i.e., Step 4); results were nonsignificant; treatment modality did not explain a significant amount of variance, and the effect size was minimal (see Table 13).

Child Compliance

Table 14 describes results for analyses examining whether post-treatment child compliance varied by treatment modality. Child-, caregiver-, and treatment-level variables (i.e., Steps 1 and 2) did not explain a significant amount of variance and exhibited small effect sizes. Pre-treatment assessments (i.e., Step 3) also did not explain a significant amount of variance and exhibited a small effect size, with pre-treatment positive caregiver statements achieving statistical significance. As hypothesized, post-treatment child compliance did not vary by treatment modality (i.e., Step 4), results were nonsignificant; treatment modality did not explain a significant amount of variance, and the effect size was minimal (see Table 14).

Table 13

Regression analyzing post-treatment child adaptive behavior comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 172) = 2.180$, $p = .021$, $R^2 = .112$			
CE – Some college[a]	3.233	2.353	.124	.171
CE – Associate degree[a]	2.575	2.401	.090	.285
CE – Bachelors degree[a]	-.611	1.975	-.033	.758
CE – Graduate degree[a]	.380	1.966	.022	.847
Caregiver 2 involvement	-1.131	1.041	-.066	.279
CR – White non-Hispanic[b]	-.735	1.274	-.037	.564
CR – Other ethnicity[b]	-3.747	2.417	-.095	.123
CR – African American[b]	-2.515	2.626	-.060	.340
Child age	-.214	.386	-.038	.581
Child sex	-.635	1.196	-.034	.596
Step 2: Treatment-level variables				
	$\Delta F(4, 168) = 1.999$, $p = .097$, $\Delta R^2 = .040$			
Language PCIT delivered	-3.725	1.870	-.139	.048
Postdoctoral therapist[c]	.367	1.636	.021	.823
Masters level therapist[c]	1.664	1.709	.092	.332
Faculty therapist[c]	3.991	2.292	.143	.084
Step 3: Intake measures				
	$\Delta F(9, 159) = 11.075$, $p < .001$, $\Delta R^2 = .326$			
Child disruptive behavior frequency	.008	.028	.026	.278
Child disruptive behavior quantity	-.123	.113	-.097	.788
Child externalizing behavior	.004	.061	.005	.945
Child adaptive behavior	.600	.069	.609	< .001
Child compliance intake	-.011	.013	-.055	.368
Caregiver stress about child behavior	.018	.049	.044	.724
Total caregiver stress	-.017	.047	-.042	.717
Caregiver intake positive statements	-.006	.076	-.005	.934
Caregiver intake corrective statements	-.006	.076	.088	.181
Step 4: Treatment modality				
	$\Delta F(1, 158) = 1.723$, $p = .191$, $\Delta R^2 = .006$			
IPCIT or clinic-based PCIT	1.680	1.279	.090	.191

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 182) = 6.197$, $p < .001$, $R^2 = .485$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Table 14

Regression analyzing post-treatment child compliance comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 163) = 1.343$, $p = .212$, $R^2 = .076$			
CE – Some college[a]	-15.478	12.324	-.153	.211
CE – Associate degree[a]	-12.041	12.003	-.116	.317
CE – Bachelors degree[a]	-6.483	10.332	-.094	.531
CE – Graduate degree[a]	-5.359	10.401	-.081	.607
Caregiver 2 involvement	-7.372	5.068	-.114	.148
CR – White non-Hispanic[b]	8.493	6.341	.112	.183
CR – Other ethnicity[b]	-4.910	11.546	-.034	.671
CR – African American[b]	-3.324	17.613	-.015	.851
Child age	1.941	1.919	.089	.313
Child sex	-4.949	5.849	-.071	.399
Step 2: Treatment-level variables				
	$\Delta F(4, 159) = 1.414$, $p = .232$, $\Delta R^2 = .032$			
Language PCIT delivered	-6.843	9.286	-.068	.462
Postdoctoral therapist[c]	6.492	8.079	.100	.423
Masters level therapist[c]	-3.976	8.467	-.057	.639
Faculty therapist[c]	17.629	11.463	.167	.126
Step 3: Intake measures				
	$\Delta F(9, 150) = 1.035$, $p = .415$, $\Delta R^2 = .052$			
Child disruptive behavior frequency	.108	.142	.096	.446
Child disruptive behavior quantity	-.575	.565	-.119	.310
Child externalizing behavior	.091	.298	.030	.760
Child adaptive behavior	.152	.342	.040	.658
Child compliance intake	.072	.063	.089	.259
Caregiver stress about child behavior	-.288	.246	-.186	.243
Total caregiver stress	.237	.235	.158	.314
Caregiver intake positive statements	.793	.385	.167	.041
Caregiver intake directive/corrective statements	.002	.156	.001	.989
Step 4: Treatment modality				
	$\Delta F(1, 149) = 1.935$, $p = .166$, $\Delta R^2 = .011$			
IPCIT or clinic-based PCIT	8.534	6.134	.120	.166

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was not statistically significant, $F(24, 173) = 1.279$, $p = .188$, $R^2 = .171$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Caregiver Stress

Caregiver Stress About Child Behavior. Table 15 describes results for analyses examining whether post-treatment caregiver stress about child behavior varied by treatment modality. Child- and caregiver-level variables explained a significant amount of variance in post-treatment caregiver stress about difficult child behaviors, and the effect size was medium, with child sex, African American and Other caregiver ethnicities achieving statistical significance. Treatment-level variables did not explain a significant amount of variance and exhibited a small effect. Pre-treatment assessments explained a significant amount of variance in post-treatment caregiver stress related to child behavior, and exhibited a medium effect size, with pre-treatment caregiver stress about difficulty child behaviors achieving statistical significance. As hypothesized, post-treatment caregiver stress about child difficult behaviors did not vary by treatment modality (i.e., Step 4); results were nonsignificant, treatment modality did not explain a significant amount of variance, and the effect size was minimal (see Table 15).

Table 15

Regression analyzing post-treatment caregiver stress about child behavior comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 172) = 3.239$, $p = .001$, $R^2 = .158$			
CE – Some college[a]	-6.450	8.206	-.078	.433
CE – Associate degree[a]	-.841	8.024	-.010	.917
CE – Bachelors degree[a]	5.848	6.878	.099	.397
CE – Graduate degree[a]	5.582	6.855	.100	.417
Caregiver 2 involvement	3.918	3.607	.072	.279
CR – White non-Hispanic[b]	6.716	4.489	.105	.137
CR – Other ethnicity[b]	19.140	8.399	.153	.024
CR – African American[b]	20.734	9.195	.157	.026
Child age	-.856	1.332	-.048	.521
Child sex	13.617	4.173	.230	.001
Step 2: Treatment-level variables				
	$\Delta F(4, 168) = 1.036$, $p = .390$, $\Delta R^2 = .020$			
Language PCIT delivered	-.657	6.456	-.008	.919
Postdoctoral therapist[c]	6.288	5.718	.115	.273
Masters level therapist[c]	3.634	5.942	.063	.542
Faculty therapist[c]	5.587	7.846	.065	.477
Step 3: Intake measures				
	$\Delta F(9, 159) = 5.595$, $p < .001$, $\Delta R^2 = .198$			
Child disruptive behavior frequency	-.109	.103	-.116	.290
Child disruptive behavior quantity	.588	.417	.146	.161
Child externalizing behavior	.382	.213	.148	.075
Child adaptive behavior	.107	.241	.034	.657
Child compliance intake	-.027	.044	-.041	.539
Caregiver stress about child behavior	.414	.169	.321	.015
Total caregiver stress	.070	.160	.055	.663
Caregiver intake positive statements	-.201	.264	-.053	.446
Caregiver intake corrective statements	.003	.112	.002	.975
Step 4: Treatment modality				
	$\Delta F(1, 158) = 461$, $p = .498$, $\Delta R^2 = .002$			
IPCIT or clinic-based PCIT	-2.998	4.415	-.051	.498

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 182) = 4.002$, $p < .001$, $R^2 = .378$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Total Caregiver Stress. Table 16 displays results for analyses examining whether post-treatment total caregiver stress varied by treatment modality. Child- and caregiver-level variables explained a significant amount of post-treatment total caregiver stress, and exhibited a small effect size, with child sex and Other caregiver ethnicity achieving statistical significance. Treatment-level variables did not explain a significant amount of variance and exhibited a minimal effect size. Pre-treatment assessments explained a significant amount of variance in post-treatment total caregiver stress, and exhibited a medium effect size, with pre-treatment total stress achieving statistical significance. As hypothesized, post-treatment total caregiver stress did not vary by treatment modality (i.e., Step 4); results were nonsignificant, treatment modality did not explain a significant amount of variance and exhibited a minimal effect size (see Table 16).

Table 16

Regression analyzing post-treatment caregiver overall stress comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 170) = 2.133$, $p = .024$, $R^2 = .111$			
CE – Some college[a]	1.755	7.616	.023	.818
CE – Associate degree[a]	-3.516	7.455	-.046	.638
CE – Bachelors degree[a]	4.001	6.384	.075	.532
CE – Graduate degree[a]	2.327	6.383	.046	.716
Caregiver 2 involvement	1.415	3.360	.028	.674
CR – White non-Hispanic[b]	5.772	4.196	.098	.171
CR – Other ethnicity[b]	16.353	7.793	.145	.037
CR – African American[b]	16.036	8.532	.134	.062
Child age	-.758	1.236	-.047	.541
Child sex	11.522	3.904	.213	.004
Step 2: Treatment-level variables				
	$\Delta F(4, 166) = .733$, $p = .571$, $\Delta R^2 = .015$			
Language PCIT delivered	-.572	6.096	-.007	.925
Postdoctoral therapist[c]	4.886	5.316	.098	.361
Masters level therapist[c]	2.937	5.515	.056	.595
Faculty therapist[c]	7.290	7.301	.093	.320
Step 3: Intake measures				
	$\Delta F(9, 157) = 5.975$, $p < .001$, $\Delta R^2 = .223$			
Child disruptive behavior frequency	-.124	.096	-.146	.196
Child disruptive behavior quantity	.483	.388	.132	.215
Child externalizing behavior	.261	.199	.111	.191
Child adaptive behavior	.041	.224	.014	.854
Child compliance intake	.009	.041	.015	.830
Caregiver stress about child behavior	-.134	.157	-.115	.395
Total caregiver stress	.626	.148	.549	< .001
Caregiver intake positive statements	-.199	.245	-.058	.417
Caregiver Intake directive/corrective statements	-.034	.105	-.024	.743
Step 4: Treatment modality				
	$\Delta F(1, 156) = 1.612$, $p = .206$, $\Delta R^2 = .007$			
IPCIT or clinic-based PCIT	-5.223	4.114	-.098	.206

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 180) = 3.598$, $p < .001$, $R^2 = .356$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Caregiver Positive Child-Directed Statements

Table 17 describes results for analyses examining whether post-treatment caregiver positive statements varied by treatment modality. Child- and caregiver-level variables explained a significant amount of variance in post-treatment caregiver positive statements and exhibited a small effect size. Treatment-level variables did not explain a significant amount of variance and exhibited a minimal effect size. Pre-treatment assessments explained a significant amount of variance and exhibited a small effect size, with pre-treatment caregiver positive child-directed statements achieving statistical significance. Contrary to hypothesized, post-treatment caregiver positive statements varied by treatment modality (i.e., Step 4); results were significant, treatment modality explained a significant amount of variance and exhibited a small effect size (see Table 17). Results of this final model indicated that caregivers in the IPCIT treatment modality had fewer post-treatment positive statements than those in the clinic-based treatment modality.

Caregiver Directive/Corrective Statements

Table 18 describes results for analyses examining whether post-treatment caregiver corrective/directive statements varied by treatment modality. Child-, caregiver-, and treatment-level variables (i.e., Steps 1 and 2) did not explain a significant amount of variance in post-treatment caregiver directive/corrective statements and exhibited small effect sizes. Pre-treatment assessments (i.e., Step 3) did not explain a significant amount of variance in post-treatment caregiver corrective/directive statements and exhibited a small effect size, with pre-treatment caregiver directive/corrective statements achieving statistical significance. Contrary to hypothesis, post-treatment caregiver

directive/corrective statements varied by treatment modality (i.e., Step 4); results were significant, treatment modality explained a significant amount of variance and exhibited a small effect size (see Table 18). Results of this final model indicated that caregivers in the IPCIT treatment modality had greater post-treatment corrective/directive statements than those in the clinic-based treatment modality.

Table 17

Regression analyzing post-treatment caregiver positive statements comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(1,) = 2.006$, $p = .036$, $R^2 = .110$			
CE – Some college[a]	-2.684	4.619	-.071	.562
CE – Associate degree[a]	-8.562	4.489	-.220	.058
CE – Bachelors degree[a]	-2.875	3.948	-.108	.468
CE – Graduate degree[a]	.163	3.983	.006	.967
Caregiver 2 involvement	-3.226	1.884	-.129	.089
CR – White non-Hispanic[b]	.143	2.388	.005	.952
CR – Other ethnicity[b]	2.761	4.277	.049	.520
CR – African American[b]	-2.967	6.539	-.036	.651
Child age	-.552	.707	-.066	.436
Child sex	-.694	2.160	-.026	.748
Step 2: Treatment-level variables				
	$\Delta F(1,) = .286$, $p = .887$, $\Delta R^2 = .006$			
Language PCIT delivered	-1.379	3.419	-.036	.687
Postdoctoral therapist[c]	1.109	2.959	.044	.708
Masters level therapist[c]	.124	3.099	.005	.968
Faculty therapist[c]	3.496	4.191	.088	.406
Step 3: Intake measures				
	$\Delta F(1,) = 1.957$, $p = .048$, $\Delta R^2 = .093$			
Child disruptive behavior frequency	.007	.053	.016	.314
Child disruptive behavior quantity	-.209	.207	-.111	.890
Child externalizing behavior	.136	.112	.113	.229
Child adaptive behavior	.063	.127	.042	.622
Child compliance intake	-.019	.024	-.060	.433
Caregiver stress about child behavior	.008	.091	.014	.432
Total caregiver stress	.070	.088	.120	.929
Caregiver intake positive statements	.463	.143	.253	.001
Caregiver intake directive/corrective statements	-.209	.207	.109	.192
Step 4: Treatment modality				
	$\Delta F(1,) = 5.139$, $p = .025$, $\Delta R^2 = .027$			
IPCIT or clinic-based PCIT	-5.118	2.258	-.188	.025

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was statistically significant, $F(24, 172) = 1.910$, $p = .010$, $R^2 = .236$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Table 18

Regression analyzing post-treatment caregiver directive/corrective statements comparing IPCIT and clinic-based PCIT families

Predictor	B	Standard Error	Beta	*p*
Step 1: Family-level variables				
	$F(10, 162) = .850$, $p = .582$, $R^2 = .050$			
CE – Some college[a]	.774	2.336	.041	.741
CE – Associate degree[a]	4.349	2.270	.226	.057
CE – Bachelors degree[a]	2.601	1.995	.198	.194
CE – Graduate degree[a]	1.565	2.013	.123	.438
Caregiver 2 involvement	.551	.946	.045	.561
CR – White non-Hispanic[b]	.480	1.189	.033	.687
CR – Other ethnicity[b]	-.188	2.163	-.007	.931
CR – African American[b]	-.061	3.307	-.001	.985
Child age	.407	.357	.099	.256
Child sex	1.519	1.092	.115	.166
Step 2: Treatment-level variables				
	$\Delta F(4, 158) = 1.541$, $p = .193$, $\Delta R^2 = .035$			
Language PCIT delivered	-.742	1.729	-.039	.668
Postdoctoral therapist[c]	-1.304	1.497	-.106	.385
Masters level therapist[c]	2.195	1.560	.166	.162
Faculty therapist[c]	-1.464	2.119	-.074	.491
Step 3: Intake measures				
	$\Delta F(9, 149) = 1.750$, $p = .083$, $\Delta R^2 = .087$			
Child disruptive behavior frequency	-.005	.027	-.021	.864
Child disruptive behavior quantity	.015	.104	.016	.884
Child externalizing behavior	-.087	.056	-.147	.123
Child adaptive behavior	.014	.064	.019	.829
Child compliance intake	.005	.012	.035	.654
Caregiver stress about child behavior	.059	.046	.199	.202
Total caregiver stress	-.042	.044	-.147	.343
Caregiver intake positive statements	-.092	.072	-.102	.201
Caregiver intake corrective statements	.093	.029	.270	.002
Step 4: Treatment modality				
	$\Delta F(1, 148) = 4.862$, $p = .029$, $\Delta R^2 = .026$			
IPCIT or clinic-based PCIT	2.518	1.142	.187	.029

Note. CE = Caregiver education; CR = Caregiver self-identified race/ethnicity.
All steps are presented with statistics they produced in the last step with all variables entered.
Full model was not statistically significant, $F(24, 173) = 1.534$, $p = .065$, $R^2 = .19$.
[a]Dummy coded categorical variable for "some high school or high school diploma"
[b]Dummy coded variable for caregiver ethnicity with "Hispanic" as the reference variable
[c]Dummy coded categorical variable for therapist level of training, with "intern" as the reference variable

Treatment Satisfaction

Descriptively, caregiver satisfaction with treatment was also analyzed to provide information about overall satisfaction with the treatment as a whole. Then levels of satisfaction between groups were compared. Specifically, on the Therapy Attitude Inventory, the average caregiver rating of treatment satisfaction was obtained. Then the means in the IPCIT and clinic-based modalities were compared via independent samples *t*-test to determine whether one treatment produced overall higher ratings of satisfaction. Treatment satisfaction is rated out of 50 on this assessment (minimum score possible is 10). Both the clinic-based ($M = 47.18$, $SD = 3.11$) and the IPCIT ($M = 46.58$, $SD = 4.07$) families reported overall high rates of satisfaction with treatment, and there was no significant difference between the treatment modalities [$t\ (180) = 1.08$, $p = .281$]. The IPCIT modality had a wider range of scores, with families rating down to a score of 32 ($N = 1$). The lowest rated score for families receiving clinic-based treatment was 38 ($N = 1$). Similar percentages of families reported the highest possible treatment satisfaction (i.e., 50; IPCIT = 30.2%, clinic-based = 28.1%).

Summary

To summarize, first paired *t*-tests analyzing the difference between pre- and post-test scores on the outcomes of interest demonstrated significant improvements for both the IPCIT and clinic modalities on child disruptive behavior frequency and quantity, child externalizing and adaptive behavior and compliance, caregiver overall stress and stress about child behavior, and caregiver positive and corrective/directive statements; effect sizes for these improvements ranged from 0.687 to 1.824, suggesting medium to large effects. Next, the families who completed telehealth before the COVID-19

pandemic were compared to those who completed treatment during the pandemic. No significant differences in treatment change were found for any of the outcome variables.

Next, a series of multiple regression analyses was conducted to determine whether improvements in treatment outcome variables varied by treatment modality. Within these regressions, multiple child-level and treatment-level variables were included as control variables, as well as the pre-treatment scores for each outcome variable of interest. The first step of the regression, composed of the family-level variables, was significant for the following outcome variables: post-treatment child externalizing behaviors, child adaptive behaviors, caregiver stress about child behavior, caregiver overall stress, and caregiver positive statements. The second step of the regression, composed of treatment-level variables, was not significant for any outcome variables, and neither were any of the individual variables included within that step. The third step of the regression, composed of pre-treatment assessments of outcome measures, was significant for the following outcome variables: post-treatment frequency of disruptive child behavior, quantity of child disruptive behavior, child externalizing behaviors, child adaptive behaviors, caregiver stress about child behavior, caregiver overall stress, and caregiver positive statements.

Although individual predictors from steps 1-3 were control variables, and specific relations with each outcome of interest were not hypothesized, several of the control variables emerged as significant. Although not consistent across the regressions, several are consistent with findings from previous literature. These results should be interpreted with caution, but nonetheless warrant reporting and discussion here to enable more rigorous replication to aid clinical work in this area in future research. Specifically, each

matching pre-treatment assessment of outcome was significantly associated with the post-treatment assessment of the same measure for each of the regression analyses with the exception of child compliance and caregiver corrective/directive statements. Other significant individual predictors of post-treatment measures included the following: child sex (i.e., for child externalizing behavior, caregiver overall stress and stress related to child behavior); language of treatment (i.e., for child adaptive behaviors); pre-treatment caregiver positive statements (i.e., for post-treatment child compliance); African American caregiver ethnicity (i.e., for caregiver stress related to child behavior); and Other caregiver ethnicity (i.e., for caregiver overall stress and stress related to child behavior).

Finally, did treatment improvements vary for families in the IPCIT vs. the clinic-based PCIT groups? Results suggested that treatment outcomes did not differ significantly by treatment modality for frequency and quantity of child disruptive behaviors, child externalizing and adaptive behaviors and compliance, and caregiver stress overall and related to child behavior, suggesting that treatment effects did not vary for families in IPCIT vs. those in clinic-based PCIT on child behavior or caregiver stress. These nonsignificant results were further confirmed by effect size analyses, which indicated differences in the minimal to small range.

In contrast, for caregiver positive and corrective/directive statements, treatment modality was a significant predictor of treatment outcomes, such that despite improvements being noted across modalities for both treatment groups, caregivers increased positive statements and decreased corrective/directive statements less in IPCIT than in clinic-based PCIT. Despite these significant results, effect sizes were small.

Finally, treatment satisfaction was high on average and did not significantly differ between IPCIT and clinic-based PCIT.

Chapter V: Discussion

Summary

The current study compared child and caregiver outcomes (i.e., child disruptive and adaptive behavior, child compliance, caregiver stress, and caregiver child-directed statements) for PCIT delivered through telehealth and clinic-based formats. This study represents an extension of previous literature of IPCIT outcomes and expands upon previous studies by a) including a comparison of caregiver statement outcomes (i.e., as well as child behavioral outcomes) between IPCIT and clinic-based samples, b) expanding the sample size of previous studies, representing the largest sample size of children and caregivers in a study of IPCIT, c) including a more diverse sample of families than previous studies of IPCIT, and d) examining IPCIT in a time-limited (i.e., 18 weeks) vs. criterion based format.

It was hypothesized first that families enrolled in both the IPCIT and clinic-based PCIT treatment modalities would demonstrate improvements in child behavior, caregiver stress and child-directed statements. Second, it was hypothesized that, controlling for potential differences between groups caused by selection bias, treatment outcomes would be not significantly differ between IPCIT and clinic-based PCIT treatment modalities.

With respect to the first hypothesis, results indicated that both IPCIT and clinic-based treatment modalities produced significant pre- to post-treatment improvements across targeted outcomes. More specifically, decreases were noted for child disruptive behaviors (both in the weekly frequency of those behaviors and the quantity caregivers deemed to be problematic), child externalizing behaviors, caregiver stress, and caregiver corrective statements, and increases were noted for child adaptive behavior and

compliance, and caregiver positive statements from pre- to post-treatment. These significant findings were supported by medium to large effect sizes.

With respect to the second hypothesis, results revealed that, for the most part, IPCIT and clinic-based treatment did not differ in post-treatment outcomes for child disruptive behaviors (i.e., both frequency and quantity), externalizing behaviors, adaptive behaviors and compliance, and caregiver stress (i.e., overall and about child behavior). These results emerged even after accounting for pre-treatment and demographic variables and controlling for selection bias using propensity scores. Effect size analyses support the outcomes across treatment modalities as well (i.e., effect sizes for group differences were minimal), thus lending more strength to the conclusion that the lack of differences across treatment modalities is real and not simply an artifact of lack of statistical power. These results did not extend to caregiver child-directed statements. Contrary to hypotheses, post-treatment caregiver child-directed statements differed between the treatment groups for both positive and corrective/directive statements. Effect sizes for these differences, however, were small.

Study findings have important clinical implications for service providers and represents an extension of Comer and colleagues' (2017) RCT, in which the efficacy of IPCIT was established in comparison to clinic-based PCIT. This study extends previous findings for IPCIT, representing the largest sample size to date examining outcomes of PCIT completed via telehealth. In the RCT comparing IPCIT and clinic-based PCIT (Comer et al., 2017), 18 families completed IPCIT, and 17 families completed clinic-based treatment. The more recent open-trial of IPCIT in rural Australia (Fleming et al., 2020) included 17 families who completed IPCIT treatment. This and the Comer (2017)

study represent the strongest studies to-date examining outcomes of IPCIT. Therefore, the current study represents a replication of previous findings in a much larger sample (i.e., $N = 198$), which is an important extension to add to the field of IPCIT.

Previous samples have also relied heavily on White, non-Hispanic samples and not included the diversity of race, ethnicity and education often represented in community agencies. Indeed, a notable strength of the current study is the enrollment and retention of families who are diverse with respect to caregiver race/ethnicity and education.

Previous studies of IPCIT have examined the outcomes of criteria-based PCIT, in which families must reach skill criteria before graduating, which varies in length depending on the family, but can sometimes last longer than 30 weeks (Comer et al., 2017). The current study operated on a time-limited PCIT protocol (i.e., 18 weeks from intake to graduation), in which no families took longer than 18 weeks in treatment, whether they had met traditional PCIT graduation criteria or not. The time-limited model produced similar improvements compared to traditional criteria-based PCIT in child behavior and caregiver stress (i.e., exemplified by similar decreases in ECBI intensity and problem scores, similar changes on child externalizing, and similar improvements to caregiver stress in studies where caregivers were required to meet criteria to graduate). Additionally, the shortened length of treatment decreases the time and financial burden to the family required to attend sessions (which often includes requested time off from school and/or work). As mentioned by Fleming and colleagues (2020), a shorter version of this highly effective treatment is more likely to be widely disseminated by community

health clinics, and more likely to be covered by a variety of insurance carriers, thus reaching a wider variety and larger number of families.

The current study also represents the first comparison of caregiver skills growth from pre- to post-treatment between IPCIT and clinic-based treatment. Fleming and colleagues (2020) established that the IPCIT families in their sample produced significant changes in caregiver statements from pre- to post-treatment, but they did not compare these families to families in clinic. In the current study, although both IPCIT and clinic-based PCIT both produced significant changes in both positive and corrective/directive caregiver statements from pre- to post-treatment, this improvement was stronger and more pronounced for clinic-based than for IPCIT. These findings suggest that clinic-based services are more effective at producing the PCIT protocol-prescribed changes in caregiver statements. The significant changes in caregiver statements for the IPCIT condition from pre- to post-treatment are consistent with limited previous IPCIT literature examining caregiver skill change (Fleming et al., 2020), and extensive literature establishing significant changes in clinic-based PCIT (Thomas et al., 2017). However, the only other existing empirical comparison of telehealth and clinic-based IPCIT (i.e., Comer et al., 2017) does not include an analysis of caregiver skill change, therefore making this the first study to examine whether caregiver child-directed statements improve similarly between telehealth and in-person PCIT.

Indeed, most telehealth parent coaching studies including caregiver skill change only examine whether caregiver skills changed significantly over the course of the treatment, and not whether that change was the same as the in-person version of that treatment (Benson et al, 2018; Corralejo & Rodriguez, 2018; Fleming et al., 2020;

Olthius et al., 2018; Unholz-Bowden et al., 2020; Vismara, McCormick, Young, Nadhan, & Monlux, 2013). There is some evidence to suggest that the effect sizes related to change in caregiver skills are smaller than that for child behavior change across studies of telehealth (Breitenstein, Gross, Christopherson, 2014). Replication of the findings from the current study in other samples will be necessary to determine if this finding of lower post-treatment caregiver skills in telehealth vs. in-person treatment is consistent across other populations and interventions.

Despite caregivers improving less in their child-directed statements in IPCIT, child behavior and caregiver stress improved similarly for IPCIT and clinic-based treatment. Although these findings require replication, possible implications include that for IPCIT, improvements in caregiver statements may be less important than child behavior or a different measurement of caregiver skill. Alternately, it may indicate that treatment in person is more effective than telehealth at improving caregiver statements for managing child behavior. Several factors could have caused caregivers to improve less in their child-directed statements in the telehealth sample than in the clinic.

Clinically, natural differences between the home and clinic environment are the most likely explanation for these observed differences between IPCIT and clinic-based PCIT in improvements in child-directed statements. For example, the more the child is engaged in the play, the more controlled the environment, and the more appropriate the toys are for child-directed play (e.g., creative, non-messy toys that don't have extensive rules), the easier it is for caregivers to speak at a high rate about positive child behaviors and avoid using commands (e.g., "come back into the room"), questions (e.g., "what do you want to play?") or criticism/sarcasm (e.g., "don't bother the dog"; Kohlhoff et al.,

2019). In other words, some contexts make it easier to follow the rules for caregiver statements in the clinic. Children are easier to keep engaged in the clinic, because they are in a small room with the caregiver, and the door to leave is closed (Kohlhoff et al., 2019). There are no other family members in another room, no television to turn on, and no pets to chase – all of which may be present in the home when completing a telehealth session (Lerman et al., 2020). In the clinic, the clinician can also control who else comes into the room during the play or coding, while in the home, the caregiver controls access to the room – sometimes unsuccessfully (e.g., siblings, family members, and pets sometimes come in unannounced; Wade et al., 2011). The more these intrusions and other distractions (e.g., doorbell, phone call) are controlled, the more the caregiver and the child can focus on the play and the more the caregiver is likely to use positive child-directive statements rather than corrective/directive statements to re-engage the child in the play.

The toys also play a substantial role both in the child's ability to engage in the play and in the caregiver's ability to use the desired PCIT statements (Kohlhoff et al., 2019). Toys that are the same from session to session or too simple or too complex for the child can make engagement more difficult. Anecdotally, there are some toys that make it easier for caregivers to positively describe play and to avoid directing or correcting the child. For example, some caregivers prefer for blocks to be built a specific way and may therefore be more able to avoid directing using pretend play toys like a doctor's kit. Other caregivers have a specific script in mind for pretend play and have difficulty allowing their child to deviate (e.g., building with pretend food, pretending doctor's tools are musical instruments). In those cases, it decreases caregiver directive

language to use more construction toys like blocks. In the clinic, the clinician sets up the toys for the caregivers. If there are specific toys that prompt more misbehavior, boredom or frustration from the child, more aggressive play, or otherwise more correction from the caregiver, the clinician typically avoids putting those in the room during subsequent sessions. Thus, the clinician structures the environment more and more over the course of treatment for both child and caregiver success with the desired skills (i.e., child positive behaviors and caregiver increase of positive statements and decrease of corrective/directive statements). The goal in telehealth is for the clinician to guide the caregiver through this structuring of the environment at home as well, but there is a limit to how much this can replicate a clinic setting (Lerman et al., 2020; Wade et al., 2011). Some families' homes do not have any rooms without other distractions, despite their best efforts. Additionally, choice of toys is constrained by the toys that are available to the family.

Logistically, there are several other factors that differ between telehealth delivery of PCIT and clinic-based delivery that could contribute to the observed difference in caregiver statements between the two. In the clinic, the PCIT therapist can model what the caregiver statements should look like directly with the child. This often occurs during check in with the family when the clinician is still in the room. During telehealth, the clinician often models the use of these skills virtually, sometimes directly to the child. However, the caregiver is not able to see how this would look in vivo when treatment takes place solely via telehealth. There is a chance that this lack of live modeling holds back caregiver statement change/improvement.

Sometimes coaching is more difficult/less accurate via telehealth as well, which may decrease the frequency of helpful coaching statements that the clinician can provide (Wade et al., 2011). For example, occasionally there are technological glitches via telehealth that make it more difficult to hear what the caregiver says, and thus those caregiver statements cannot receive feedback from the clinician (Lerman et al., 2020). Some caregivers use headsets that make it difficult to hear what the child says during the play (i.e., because high-quality Bluetooth technology is designed to filter out background noise). The difficulty hearing the child makes it more challenging to code reflections, or the caregiver's repetition of child statements, because it is not always possible to discern whether the caregiver's statement was a reflection of child speech or a novel statement. Visual limitations can also limit clinician coaching (Lerman et al., 2020; Wade et al., 2011), as in the clinic setting the clinician can see everything the child is doing and coach the parent accordingly, but in the home the child is much more likely to wander off-camera. When children could not be viewed, clinicians continued to coach, but were limited to vaguer coaching statements due to the inability to view the child's specific actions (i.e., "It sounds like he's continuing to drive his car." vs. "I see him driving toward you again!"; Wade et al., 2011). Giving these less-accurate coaching statements in IPCIT due to the differences from clinic-based PCIT may make it more difficult for caregivers to grow their skills as much virtually. This hurdle combined with a chaotic environment, the difficulty hearing one another, decreased clinician modeling, and access to appropriate toys likely all coalesce to make it more difficult both for caregivers to use the PCIT skills (i.e., increased positive statements and decreased directive/corrective statements) and for clinicians to coach them.

Despite the difference in improvement in caregiver statements between telehealth and clinic-based treatment, child behavior changed equally between groups. Given that child behavior by and large improved similarly despite the lack of corresponding improvement in caregiver statements, perhaps future studies would benefit from reconceptualizing this measure of caregiver behaviors. First, caregivers are observed on their use of child-directed statements in a structured task every time they are coded to increase standardization. However, over the course of treatment, the clinical focus is diffused from a play only situation to other situations throughout the day (e.g., mealtime, homework, getting dressed, transitions, etc.). Therefore, coding caregiver statements in a play situation becomes arguably less relevant to overall behavioral change over the course of treatment. More innovative research methods that track caregiver statements throughout the day or sample randomly from segments of the day would likely capture more accurately how caregivers generalize their skills and thus would be more likely to be related to positive child behavior change. In addition, such nuanced changes may not be captured by focusing on the raw number of positive statements or decrease for corrective statements. Instead, it may be more meaningful in future analyses to begin to capture positive statements as a proportion of overall statements (i.e., positive statements/all statements during coding), and do the same for corrective statements, as has been done in previous PCIT work (Bagner et al., 2016). Alternately, researchers may want to explore whether it is more meaningful (i.e., related to overall treatment progress variables) to analyze caregiver statements as the ratio between positive and corrective statements (i.e., positive statements composite/corrective statements; Phaneuf & McIntyre, 2011).

Within PCIT, child behavior is the primary outcome of interest and often the primary reason children are referred for treatment. Change in caregiver child-directed statements is hypothesized to be a mechanism through which child behavior improves. However, it is possible that the way PCIT therapists have traditionally measured caregiver statements (i.e., the coding that takes place during the structured observation of parent-child play) does not accurately assess the caregiver behaviors that directly influence child behavior change over the course of treatment. PCIT's caregiver skill focus during child-directed play involves decreasing the three corrective categories of statements and increasing the three positive caregiver statements measured, which are usually those emphasized in PCIT research, as they are easiest to measure according to strategies provided via the PCIT protocol (i.e., the DPICS-IV observation). These caregiver statements are hypothesized to be a proxy for the measurement of differential attention or decreasing caregiver attention to the child's negative behaviors and increasing positive attending to desirable behaviors (e.g., playing calmly, being gentle, sharing, using kind language).

However, the converse of positive attending—that is, ignoring—is assessed during the five-minute observation by a dichotomous rating of whether the caregiver is "satisfactory" or "needs practice" in their ability to use differential attention during play time (coders may alternatively indicate "N/A" if there were no behaviors that required ignoring). The technique of depending on caregiver statements to measure selective attention loses some of the qualitative nonverbal nuance in caregivers' behavior when ignoring disruptive behavior, including caregiver sighs of frustration, laughing at inappropriate comments, or alternately turning away until behavior improves and

providing more enthusiasm for more appropriate behavior. However, currently caregiver ignoring is not assessed directly in a way that lends itself well to tracking nuanced change over time. A dichotomous rating of ignoring does not capture the incremental improvements that caregivers make from session to session. A more detailed observational measure of selective attention specifically may align more closely with behavior change over the course of treatment than the measurement of caregiver positive or directive statements. Future research should explore behavioral observation measures which can be completed concurrently with the DPICS-IV observation to assess whether these two observational measures are related, and which is more related to behavior change over time, to defend the use of caregiver skill acquisition goalposts for treatment completion and progression.

Findings of the current study suggest that time-limited PCIT (e.g., 18 weeks) in person and via telehealth are both related to significant, similar changes in child disruptive behavior. As PCIT is currently practiced in many clinics, caregivers move from one phase of treatment (i.e., CDI) to the next (i.e., PDI), and on to graduate from treatment contingent upon their meeting the DPICS-IV criteria for caregiver statements (i.e., for CDI, 10 behavior descriptions, 10 reflections, 10 labeled praises and three or fewer questions, commands and critical/sarcastic statements; for PDI and graduation from treatment, re-demonstrating CDI criteria as well as demonstrating 75% effective commands and 75% correct follow-through for effective commands, as well as an ECBI Intensity score at or below 114). If the findings of this study hold and caregiver statements have limited relationships to child behavior change, then it may be clinically unethical to keep families in treatment indefinitely until they meet a specific threshold of

skills that has not been tied to the clinical outcomes that are pivotal for long-term child and caregiver functioning. Instead, a more time-limited model like that used in this study may be a better model to decrease the burden of time taken off from work/school for the family and increase their chances of successfully completing the treatment program and improving family functioning.

Caregivers in both treatment modalities reported high levels of satisfaction. Satisfaction did not differ significantly between the treatment modalities, consistent with previous telehealth literature for both telehealth in general and IPCIT specifically (Backhaus et al., 2012; Comer et al., 2017; Turvey & Myers, 2012). It is notable that even within the context of a pandemic and heightened worldwide stress, caregivers continue to value treatment and benefit from PCIT.

Importantly, caregivers in the IPCIT condition had varying levels of competence/comfort with technology, which many clinicians worry will impact family satisfaction with treatment when conducted via telehealth (Brooks, Turvey, & Augusterfer, 2013; Connolly et al., 2019; Kruse et al., 2018). Similar levels of caregiver-reported satisfaction with each condition suggests that previous provider concerns about families not wanting telehealth were likely unfounded. However, this study did not directly measure family technological comfort/familiarity or previous use of telehealth treatment, which could potentially influence the satisfaction with treatment. Moreover, clinicians in this study provided families with any requested/needed coaching about how to set up the equipment, as well as active troubleshooting strategies for Zoom, Bluetooth, and internet connectivity concerns. This model of technology support is likely more widely generalizable to community health/private providers than previous models which

include an on-call IT support specialist (Comer et al., 2017). However, this model (i.e., clinician tech support) still requires the treating clinician to have a basic competency of troubleshooting the requisite technology for sessions. This may require additional training for many clinicians, and such training may be inaccessible. Future research in this area should more thoroughly assess family satisfaction variables directly related to telehealth, like family comfort with the technology as well as willingness to use telehealth versus in-person services in the future, and how these variables are related to overall treatment satisfaction and outcomes.

Limitations and Future Directions

Lack of Treatment Non-Completers

Several methodological limitations warrant consideration within the current study. First, this study represents an examination of only those families who completed the full dose of PCIT treatment, to ascertain whether the full course of treatment presented via a different modality (i.e., telehealth) produces similar results to traditional clinic-based services. Those families who did not complete treatment were not included. Future studies should examine differences between groups for treatment non-completers as well, to determine whether telehealth is effectively reaching the families who need services and decreasing the barriers it is intended to decrease.

Reliance on Caregiver Report

A second methodological limitation is the study's reliance on caregiver-report measures for child disruptive behavioral outcomes, a strategy prone to bias for a variety of reasons (Carter, Briggs-Gowan, & Davis, 2004). For example, caregivers experiencing their own psychopathology may be prone to report more severe child misbehavior, as

might caregivers who are seeking to prove eligibility for services (e.g., Briggs-Gowan, Carter, & Schwab-Stone, 1996; Fergusson, Horwood, & Lynskey, 1995; Harvey, Fischer, Weieneth, Hurwitz, & Sayer, 2013). The particular behaviors often treated in PCIT (e.g., hyperactivity, defiance) are often the exact behaviors which produce the most ratings discrepancy between caregivers and other reporters, particularly teachers (Harvey et al., 2013). Due to these frequent, multifaceted discrepancies between caregiver and teacher report of child behaviors, it is recommended as a best practice in psychological assessment to obtain collateral reports of child behavior, often from teachers or direct observation of child behavior related to the same outcomes reported by the caregiver (Carter, Briggs-Gowan, & Davis, 2004). Although child compliance was included as a measure of direct observation of child behavior in this study, it only reflects change in one positive child behavior (i.e., compliance) over the course of treatment, when caregivers usually present to treatment with many behavioral concerns, including aggression, fighting/arguing, dawdling, hyperactivity, etc. Additionally, child compliance was not correlated with any of the caregiver-report measures of child behavior in the current study, indicating that it may not accurately capture the complexity of child behavior/misbehavior presented by the children in PCIT. Future studies' inclusion of teacher-report or direct observation assessments measuring the same constructs as the caregiver-report measures (e.g., frequency of aggression during session) could increase the strength of the findings that IPCIT/PCIT produces meaningful child outcomes over time and across settings.

Non-Randomized Design

A third limitation is the lack of a randomized control group. A quasi-experimental design was chosen for this study to increase the external validity of the findings. It was often not possible for families who self-selected into the telehealth condition to receive clinic-based treatment due to geographical constraints (e.g., living too far from the clinic to be randomized into the clinic condition, or lacking access to reliable transportation) or because the in-person clinics were closed temporarily during the COVID-19 pandemic. Therefore, the purpose of this study was more to analyze whether those who self-selected into IPCIT would exhibit similar outcomes to those who self-selected into clinic-based PCIT – a model of treatment selection/assignment more likely to occur in a variety of clinics completing this treatment, and not just those conducting clinical trials. However, the use of a quasi-experimental design includes inherent risks to internal validity. The lack of a randomized control group decreases the ability to attribute change in outcome variables to the intervention, as it introduces the possibility that differences between groups could more parsimoniously explain any changes or improvements produced (Miller, Smith, & Pugatch, 2020).

In this study, propensity scores were utilized to decrease differences between groups in the analyses and to increase internal validity of the findings. Although the use of propensity scores strengthens the quasi-experimental model, and is considered a strength of this study, it remains a less robust experimental design than a randomized controlled trial. Additionally, not all threats to internal validity are accounted for in the current study as the control group used is akin to a "treatment as usual" group in the form of the clinic-based PCIT. However, as child and caregiver outcomes are expected to change from pre- to post-treatment in both outcomes, effects of history, maturation, and

statistical regression (i.e., additional threats to internal validity) are not controlled. An increasingly robust model of analysis would include a waitlist control group, although often families of children with such severe behavioral concerns are loathe to wait for services for months when treatment may be available at other locations in the community. There are additionally ethical concerns related to not providing timely treatment to children with behavior problems.

Future research should combine the strengths of the RCT embodied by Comer and colleagues' (2017) examination of IPCIT, and the more representative sample utilized by this study and the open-trial conducted by Fleming and colleagues (2020), decreasing threats to internal validity by adding a waitlist control group, so that IPCIT is established as both more widely efficacious (i.e., given the more robust methodology), and still demonstrating external validity.

Inconsistent Availability of Treatment Modalities

A fourth limitation is the inconsistent availability of treatment modalities during the period of the study. Although one of the goals of this study was to expand current IPCIT literature to include an IPCIT sample who self-selected into telehealth, this model was somewhat complicated by the onset of the COVID-19 pandemic. That is, when the clinic transitioned to 100% remote services to decrease the spread of COVID-19, although families continued to have the option to select into a standing waitlist for in-person services (i.e., with the expectation that they would be contacted first when in-person clinics reopened), the number of families who elected to wait for in-person services decreased as the pandemic wore on. Selecting into in-person only services required families to wait for an indeterminate amount of time until one of the in-person

clinics opened. Therefore, unlike before the pandemic when family wait times were only determined by therapist schedule and availability, during the pandemic families who preferred in-person services were disproportionately disadvantaged regarding the availability of services. At the time of the transition to solely remote services, the clinic greatly increased the support to families uncomfortable with technology, but it is possible that some of the families who received telehealth treatment during the pandemic would have preferred in-person services had those services been readily available at that time. Therefore, although treatment selection bias was decreased using propensity score matching in these analyses, study of IPCIT should continue to examine differences between self-selected groups once it is safe again to open in-person clinics.

Lack of Treatment Follow-up

A fifth methodological limitation is the lack of an assessment of caregiver and child outcomes after the termination of treatment – a follow-up data point. Often after graduating from weekly treatment, child behavior returns closer to baseline than it was at the termination of treatment (Comer et al., 2017), potentially due to caregivers practicing less intensively than they did during active treatment. Future studies should examine caregiver and child outcomes for at least one follow-up point. Findings would be even more robust with more than one follow-up point establishing whether treatment gains hold similarly over time for both IPCIT and clinic-based treatment (i.e., as has been previously established with clinic-based PCIT; Boggs et al., 2005).

Generalization to Areas with Weaker Broadband Access

The sixth limitation of the current study regards its location within an area with relatively stable and reliable broadband internet access relative to other remote

populations. Within the South-Florida area where this study was conducted, population density is about 1,400 per square mile, meaning that it takes a longer time to drive shorter distances than in more rural areas of the country and driving across the county for mental health services takes prohibitively long (i.e., upwards of two to three hours). However, high population density also means that families have fairly consistent access to high-speed internet access necessary for telehealth services. Therefore, given the results from this study it remains unclear whether this type of treatment would generalize well to more rural locations. Previous studies have shown promise with IPCIT in rural settings in other countries (Fleming et al., 2020), and other caregiver-coaching treatments delivered via synchronous videoconferencing in rural settings in the United States (Riegler et al., 2020; Tse et al., 2015). However, millions of families in the United States continue to have difficulty accessing broadband internet in rural areas. During the COVID-19 pandemic, this lack of access to broadband internet prevented rural families from accessing remote schooling, and precluded caregivers working from home (Giorgi, 2020). Therefore, unless broadband is expanded further into such rural areas in the country, it is unlikely that families in these areas would be able to access virtual delivery of mental health services (Summers-Gabr, 2020).

Lack of Measurement for Caregiver Treatment Barriers

The last notable methodological limitation involves this study's lack of explicit measurement of barriers to treatment, which would have allowed for the direct assessment of between group (i.e., IPCIT and clinic-based) differences, to empirically demonstrate that telehealth decreases barriers and thus increases access to treatment. Telehealth has often been upheld as a strategy for overcoming barriers to traditional

clinic-based treatment, including transportation, caregiver hours of availability, distance from the clinic, available mental health providers in the area, etc. (CDC, 2020; Summers-Gabr, 2020). For many around the world beginning in early 2020, telehealth temporarily became the only option for receiving mental health services, as lockdown orders differed from country to country, and from state to state in the U.S. (Bennett, Ruggero, Sever, & Yanouri, 2020). Comer and colleagues' (2017) RCT of PCIT directly measured caregivers' perceptions of their experienced barriers to treatment and established that caregivers in the IPCIT condition indeed experienced fewer barriers as theoretically postulated. Although the current study extended Comer's study by further decreasing barriers to treatment (e.g., limiting the length of treatment to limit the time burden on families), treatment barriers were not directly measured. Future studies examining IPCIT should measure treatment barriers directly so they can be compared between groups, including barriers like access to dependable transportation, miles or amount of time traveled to reach the clinic, scheduling conflicts (e.g., with work, other child activities), finding care for other children, treatment relevance to the family, health problems of family members (i.e., child or caregiver), discomfort with the clinic setting, etc. This might be particularly important for treatment non-completers.

Future Directions

Despite limitations described above, the current study demonstrates a significant contribution to the literature demonstrating the potential for PCIT as an effective treatment delivered via telehealth. Future directions for both treatment, research, and training should include the further expansion of IPCIT, including understanding how to make this treatment equally effective and appealing for all groups of caregivers and

children who would benefit from PCIT and cannot otherwise access clinic-based treatment (e.g., due to lack of transportation, stigma about seeking services in a mental health setting, distance from the clinic, and other barriers to treatment described in detail earlier).

Within individual regression models, some child, family, and treatment-level variables emerged as significant; although not hypothesized, these results may serve as a springboard for hypotheses for possible replication in future work. The majority of these variables were not significant across outcomes; thus, caution is warranted in drawing conclusions about these findings without support from previous literature or corroboration from future independent analyses. However, they are described here to enable future studies to expand upon them as appropriate.

First, for the model predicting post-treatment child externalizing behaviors, girls were rated higher than boys, even when accounting for pre-treatment scores. This sex difference in externalizing by child sex is supported by previous studies of the development of aggression in young children, which suggest that boys (but not girls) are less likely to exhibit continued aggression when bedtime and mealtime routines are predictable (i.e., major treatment goals for PCIT; Rijlaarrsdam et al., 2016) and when parent discipline is consistent (Tung, Li, & Lee, 2012). Similarly, for caregiver stress related to child behavior and for total/overall caregiver stress, caregivers of girls reported significantly higher post-treatment stress than boys (i.e., intake stress was not significantly different), a finding potentially related to girls' higher post-treatment aggression contributing to continued heighted caregiver stress as well. This explanation is supported by previous literature, as child disruptive behavior and caregiver stress

demonstrate a robust positive feedback loop (Krahé et al., 2014), and externalizing behaviors are consistently viewed as less socially accepted for girls than for boys (Martin & Ross, 2005; Rubin, 2010).

Caregiver ethnicity (i.e., African American and Other ethnicity) was also significant in the model for post-treatment caregiver stress, with African American and caregivers in the Other ethnicity group reporting higher post-treatment stress. Although a unique finding between the individual analyses, it is potentially a reflection of the disproportionate stress experienced by BIPOC families during the COVID-19 epidemic (i.e., a large portion of the IPCIT sample), perhaps due to the greater likelihood of being exposed to uncontrollable stressors beyond parenting, including inflexible work arrangements, lack of sick leave, and exposure to COVID-19 (e.g., Iruka et al., 2021; Liu & Modir, 2020). Additionally, although found inconsistently in the current study, previous PCIT literature demonstrating similar disparities in stress reduction for African American families in PCIT (Fernandez, Butler, & Eyberg, 2011) also underscores the need for further clinical and research action in this area.

Caregiver ethnicity described as Other (i.e., English-Speaking Caribbean, Russian, Brazilian) was also significant in the model for post-treatment caregiver stress about child behavior and total stress. However, all caregivers reported total pre-treatment stress below clinically elevated levels, so there is a chance that differential improvements in stress were not clinically meaningful when families begin treatment with stress levels below clinical significance.

Caregiver pre-treatment positive statements also emerged as a significant predictor of child compliance. This finding is supported theoretically by PCIT's

connection drawn between increased positive parental attention and child compliance (Mandal, Olmi, Edwards, Tingstrom, & Benoit, 2000; Roberts, Tingstrom, Olmi, & Bellipanni, 2008; Schueler & Prinz, 2013), but has not been empirically supported by other researchers and therefore warrants replication before further conclusions can be drawn. In the model predicting child post-treatment adaptive behavior, the language of treatment delivery emerged as a significant predictor. However, this appears to the result of a ceiling effect, as pre-treatment scores for Spanish speaking families are already within normal limits and therefore had less room to change. Among the extraneous findings from the analyses here, some appear to be the result of measurement effects, but some appear to support findings from previous studies, and therefore warrant further consideration in future studies of PCIT, both clinic-based and IPCIT.

Findings demonstrating positive outcomes with IPCIT here show promise for reaching families who vary greatly in their income levels and therefore their access to technology of sufficient complexity for telehealth sessions. Previous studies into barriers to telehealth have mentioned the costs families may incur being prohibitive to their participation, particularly regarding purchasing the equipment needed for telehealth (e.g., Bluetooth, computer or tablet, webcam, etc.; Wen Yang et al., 2020). Indeed, the cost of in-person PCIT has been shown to be much higher than many other behavioral treatments due to the room-setup (i.e., rent for two rooms), and toys and technology needed for in-person coaching (Hare & Graziano, 2020; Washington State Institute of Public Policy, 2019).Within the current study this barrier was mitigated somewhat by families' ability to borrow technology from the clinic for the duration of treatment if requested. However, many families opted to use their own technology, and in lieu of a laptop or tablet, chose

to use their smartphone as video and audio devices. As adults in the United States are more likely to own a smartphone than a laptop or tablet (Pew, 2019), offering treatment effectively via smartphone in this manner can increase the reach and accessibility of this treatment. Indeed, previous studies have lauded smartphones as an effective method of reaching families with limited technological access (Lerman et al., 2020; Rios et al., 2018). Unfortunately, rural American adults are less likely to own a smartphone (71%) than those living in urban areas (83%; Pew, 2019), so while the use of smartphones ameliorates some of the cost barriers to telehealth, it upholds some of the barriers due to geography. Future studies of more widespread feasibility of rural/remote telehealth interventions should involve patient populations where broadband internet is less reliable and assess whether families can rely on their own equipment for services. Examining outcomes for families without any technological supplies provided by the treating clinicians would increase the generalizability of these findings and increase confidence that telehealth could be provided over longer distances between the treating clinician and the family, as without material support from the provider, the family bears the brunt of costs for the technology needed to participate in IPCIT services.

Families using smartphones for treatment were also sometimes able to bypass another barrier to telehealth treatment – broadband internet access. The use of smartphones enabled some caregivers with limited/inconsistent internet connection to use wireless data plans as a substitute or backup connection to enable treatment sessions to flow smoothly. However, the option to use a wireless data plan was only available for caregivers with unlimited data plans for their smartphone – access that is often limited by financial means. Internet connection was often influenced by more than just location (i.e.,

urban vs. rural) or financial capacity to buy high-speed internet. Often the video/audio quality of the session deteriorated the more that others in the home were using the internet, particularly during COVID-19 as more families needed to work from home or attend school remotely. Although this barrier to strong telehealth connection still disproportionately affected families of lower means, it was a more malleable barrier as therapists could schedule sessions when other family members did not need the internet or use video over the internet and audio over the phone to decrease the strain on the family's limited bandwidth. Although hotspots were available for families to borrow for this study, therapists were able to troubleshoot using several of the previously mentioned strategies, and caregivers often preferred these to using a new and unfamiliar technology.

A wide variety of barriers to mental health treatment can be addressed using telehealth, particularly on the level of individual families. However, IPCIT has yet to establish that it can increase access to telehealth on a more system-wide level. For example, although IPCIT can connect well-trained providers and patients who are very remote from one another, IPCIT alone will not be able to solve the wider population problem of the gap between the number of children who need behavioral intervention services and the number of providers available (Kazdin & Blase, 2011). More likely, IPCIT can serve as part of a wider constellation of differentiated services (e.g., smartphone apps) to begin to address the significant mental health gap.

Conclusion

Taken together, the current study yielded findings that have important implications for addressing barriers to mental health services. This study contributes to the existing IPCIT literature by providing support for the effectiveness of PCIT delivered

via telehealth in improving child behavior and parenting stress among a more diverse population with respect to SES, ethnicity and education than has been included in prior studies. This study also expands the previous research by utilizing a larger sample where caregivers (mostly) self-selected into IPCIT or clinic-based PCIT, which strengthens ecological validity. This study also examined the effectiveness of time-limited PCIT, which has support from previous studies in clinic-based PCIT, but has not previously been examined in IPCIT. This study also examined caregiver skill change over time, inviting the question of whether IPCIT or the caregiver skill measurement for PCIT require adjusting to ensure that meaningful caregiver skills improve over the course of PCIT, no matter where the treatment takes place. While the increased evidence for PCIT as a robust behavioral intervention across modalities adds further breadth to the range of effective telehealth treatments, further research is warranted to document intervention sustainability on a more system-wide level. It will be important for future work to continue to broaden the range of available, effective telehealth treatments, as it will provide mental health service providers and families with more options for virtual treatment selection to match the intervention more accurately to the presenting child behavior. Over time, the increase in effective individual services available over telehealth will serve a role in the effort to decrease barriers and increase child and family access to needed interventions, thus decreasing the challenges associated with child behavior issues, both for families and their children.

References

Abidin, R. R. (2012). *Parenting Stress Index (4th ed.).* Lutz, FL: PAR.

American Psychological Association (APA). (2010). *Ethical principles of psychologists and code of conduct.* Washington, DC: Author.

American Psychological Association (APA). (2013). *Guidelines for the practice of telepsychology.* Washington, DC: Author.

Backhaus, A., Agha, Z., Maglione, M. L., Repp, A., Ross, B., Zuest, D., Rice-Thorp, N. M., Lohr, J., & Thorp, S. R. (2012). Videoconferencing psychotherapy: A systematic review. *Psychological Services, 9*(2), 111-131. DOI: 10.1037/a0027924

Bagner, D. M., Coxe, S., Hungerford, G. M., Garcia, D., Barroso, N. E., Hernandez, J., & Rosa-Olivares, J. (2016). Behavioral parent training in infancy: A window of opportunity for high-risk families. *Journal of Abnormal Child Psychology*, *44*(5), 901-912.

Bai, H. & Clark, M. H. (2019). Basic concepts of propensity score methods. In *Propensity Score Methods and Applications,* pp. 1-17. SAGE Research Publications, Ltd.: Thousand Oaks. https://dx.doi.org/10.4135/9781071814253

Barroso, N. E., Hungerford, G. M., Garcia, D., Graziano, P. A., & Bagner, D. M. (2016). Psychometric properties of the parenting stress index-short form (PSI-SF) in a high-risk sample of mothers and their infants. *Psychological Assessment, 28*(10), 1331-1335. http://dx.doi.org/10.1037/pas0000257

Baumrind, D. (2013). Authoritative parenting revisited: History and current status. In R. E. Larzelere, A. S. Morris, & A. W. Harrist (Eds.) *Authoritative Parenting:*

Synthesizing Nurturance and Discipline for Optimal Child Development, pp. 11-34. American Psychological Association. DOI: 10.1037/13948-002

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development, 37*, 887–907. doi:10.2307/1126611

Becevic, M., Boren, S., Mutrux, R., Shah, Z., & Banerjee, S. (2015). User satisfaction with telehealth: Study of patients, providers, and coordinators. *The Health CCare Manager, 34*(4), 337-349. DOI: 10.1097/HCM.0000000000000081

Bennett, C. B., Ruggero, C. J., Sever, A. C., & Yanouri, L. (2020). eHealth to redress psychotherapy access barriers both new and old: A review of reviews and meta-analyses. *Journal of Psychotherapy Integration, 30*(2), 188-207. http://dx.doi.org/10.1037/int0000217

Benoit, D. A., Edwards, R. P., Olmi, D. J., Wilcynski, S. M., & Mandal, R. L. (2001). Generalization of a positive treatment package for child noncompliance. *Child & Family Behavior Therapy, 23*(2), 19-32. DOI: 10.1300/J019v23n02_02

Benson, S. S., Dimian, A. F., Elmquist, M., Simacek, J., McComas, J. J., & Symons, F. J. (2018). Coaching parents to assess and treat self-injurious behaviour via telehealth. *Journal of Intellectual Disability Research, 62*(12), 1114-1123. DOI: 10.1111/jir.12456

Boggs, S. R., Eyberg, S. M., Edwards, D. L., Rayfield, A., Jacobs, J., Bagner, D. & Hood, K. K. (2005). Outcomes of parent-child interaction therapy: A comparison of treatment completers and study dropouts one to three years later. *Child & Family Behavior Therapy*, *26*(4), 1-22. DOI: 10.1300/J019v26n04_01

Boggs, S. R., Eyberg, S., & Reynolds, L. A. (1990). Concurrent validity of the Eyberg Child Behavior Inventory. *Journal of Clinical Child Psychology, 19*(1), 75-78.

Bornheimer, L. A., Acri, M. C., Gopalan, G. (2018). Barriers to service utilization and child mental health treatment attendance among poverty-affected families. *Psychiatric Services, 69*(10), 1101-1104.

Breitenstein, S. M., Gross, D., & Christophersen, R. (2014). Digital delivery methods of parenting training interventions: A systematic review. *Worldviews on Evidence-Based Nursing, 11*(3), 168-176.

Breslow, N. E. (1996). Statistics in epidemiology: The case-control study. *Journal of the American Statistical Association*, *91*(433), 14–28.

Brestan, E., Jacobs, J., Rayfield, A., & Eyberg, S. M. (1999). A consumer satisfaction measure for parent– child treatments and its relationship to measures of child behavior change. *Behavior Therapy, 30*, 17–30. http://dx.doi.org/10.1016/S0005-7894(99)80043-4

Briggs-Gowan, M., Carter, A., & Schwab-Stone, M. (1996). Discrepancies among mother, child and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology, 24,* 749–765.

Bringewatt, E. H., Gershoff, E. T. (2010). Falling through the cracks: Gaps and barriers in the mental health system for America's disadvantaged children. *Children and Youth Services Review 32*:1291–1299.

Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology, 163*(12), 1149-1156. DOI: 10.1093/aje/kwj149

Brooks, E., Turvey, C., Augusterfer, E. F. (2013). Telemental health provider barriers to telemental health: Obstacles overcome, obstacles remaining. *Telemedicine and E-Health, 19*(6), 433-437. DOI: 10.1089/tmj.2013.0068

Cameron, E. E., Joyce, K. M., Delaquis, C. P., Reynolds, K., Protudjer, J. L. P., & Roos, L. E. (2020). Maternal psychological distress & mental health service use during the COVID-19 pandemic. *Journal of Affective Disorders, 276*, 765-774. https://doi.org/10.1016/j.jad.2020.07.081

Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry, 45*(1), 109-134.

Centers for Disease Control and Prevention. (2020). *Rates of mental and behavioral health service providers by county.* Atlanta, GA: Author. Retrieved from https://www.cdc.gov/childrensmentalhealth/stateprofiles-providers.html

Chou, T., Bry, L. J., & Comer, J. S. (2017). Overcoming traditional barriers only to encounter new ones: Doses of caution and direction as technology-enhanced treatments begin to "go live." *Clinical Psychology Science and Practice, 24*, 241-244. doi:10.1111/cpsp.12196

Cohen, J. E. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Comer, J. S., Furr, J. M., Cooper-Vince, C., Madigan, R. J., Chow, C., Chan, P., Idrobo, F., Chase, R., McNeil, C. B., & Eyberg, S. M. (2015). Rationale and

considerations for the internet-based delivery of parent– child interaction therapy. *Cognitive and Behavioral Practice, 22*, 302–316. http://dx.doi.org/10.1016/j.cbpra.2014.07.003

Comer, J. S., Furr, J. M., Miguel, E. M., Cooper-Vince, C. E., Carpenter, A. L., Elkins, R. M., Kerns, C. E., Cornacchio, D., Chou, T., Coxe, S., DeSerisyl, M., Sanchez, A. L., Golik, A., Martin, J., Myers, K. M., & Chase, R. (2017). Remotely delivering real-time parent training to the home: An initial randomized trial of internet-delivered parent-child interaction therapy (I-PCIT). *Journal of Consulting and Clinical Psychology, 85*(9), 909-917. http://dx.doi.org/10.1037/ccp0000230

Connolly, S. L., Miller, C. J., Lindsay, J. A., & Bauer, M. S. (2019). A systematic review of providers' attitudes toward telemental health via videoconferencing. *Clinical Psychology Science and Practice, 27*, 1-19. https://doi.org/10.1111/cpsp.12311

Conway, A., Rolley, J. X., Fullbrook, P., Page, K., & Thompson, D. R. (2013). Improving statistical analysis of matched case-control studies. *Research in Nursing & Health, 36,* 320-324. DOI: 10.1002/nur.21536

Corralejo, S. M. & Rodriguez, M. M. D. (2018). Technology in parenting programs: A systematic review of existing interventions. *Journal of Child and Family Studies, 27*, 2717-2731. https://doi.org/10.1007/s10826-018-1117-1

Crum, K. I. & Comer, J. S. (2016). Using synchronous videoconferencing to deliver family-based mental healthcare. *Journal of Child and Adolescent Psychopharmacology, 26*(3), 229-234. DOI: 10.1089/cap.2015.0012

Dadds, M. R., Thai, C., Diaz, A. M., Broderick, J., Moul, C., Tully, L. A., Hawes, D. J., Davies, S., Burchfield, K., & Cane, L. (2019). Therapist-assisted online treatment

for child conduct problems in rural and urban families: Two randomized controlled trials. *Journal of Consulting and Clinical Psychology, 87*(8), 706-719. http://dx.doi.org/10.1037/ccp0000419

de Figueiredo, C. S., Sandre, P. C., Catarina, L., Mázala-de-Oliveira, T, Chagas, L. S., Raony, I., Ferreira, E. S., Giestal-de-Araujo, E., dos Santos, A. A., & Bomfim, P. O. (in press). COVID-19 pandemic impact on children and adolescents' mental health: Biological, environmental, and social factors. *Progress in Neuropsychopharmacology & Biological Psychiatry.* https://doi.org/10.1016/j.pnpbp.2020.110171

Eisenstadt, T. H., McElreath, L. S., Eyberg, S. M., & McNeil, C. B. (1994). Interparent agreement on the Eyberg Child Behavior Inventory. *Child and Family Behavior Therapy, 16*(1), 21-27.

Eyberg, S. M. (1988). Parent-child interaction therapy: Integration of traditional and behavioral concerns. *Child & Family Behavior Therapy, 10,* 33-46.

Eyberg, S. M., & Funderburk, B. W. (2011). *Parent–child interaction therapy protocol.* Gainesville, FL: PCIT International.

Eyberg, S. M., Nelson, M. M., Ginn, N. C., Bhuiyan, N., & Boggs, S. R. (2013). *Dyadic parent child interaction coding system: Comprehensive manual for research and training (4th ed.).* Gainesville, FL: PCIT International.

Eyberg, S. M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory–Revised: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Ferguson, C. L. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5), 532-538. DOI: 10.1037/a0015808

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1995). The stability of disruptive childhood behaviors. *Journal of Abnormal Child Psychology, 23,* 379–96.

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, 46*(8), 837-849. DOI:10.1111/j.1469-7610.2004.00387.x

Fernandez, M. A., Butler, A. M., & Eyberg, S. M. (2011). Treatment outcome for low socioeconomic status African American families in Parent-Child Interaction Therapy: A pilot study. *Child & Family Behavior Therapy, 33*, 32-48. DOI: 10.1080/07317107.2011.545011

Fitzpatrick, O., Carson, A., & Weisz, J. R. (2020). Using mixed methods to identify the primary mental health problems and needs of children, *Child Psychiatry & Human Development,* https://doi.org/10.1007/s10578-020-01089-z

Fleming, G. E., Kohlhoff, J., Morgan, S., Turnell, A. Maiuolo, M., & Kimonis, E. R. (2020). An effectiveness open trial of internet-delivered parent training for young children with conduct problems living in regional rural Australia. *Behavior Therapy, in press.* https://doi.org/10.1016/j.beth.2020.03.001

Forehand, R., & Atkeson, B. M. (1977). Generality of treatment effects with parents as therapists: A review of assessment and implementation procedures. *Behavior Therapy, 8,* 575-593.

Fowles, T. R. Masse, J., McGoron, L., Beveridge, R. M., Williamson, A. A., Smith, M. A., & Parrish, B. P. (2017). Home-based vs. clinic-based parent-child interaction therapy: Comparative effectiveness in the context of dissemination and implementation. *Journal of Child and Family Studies, 27*, 1115-1129. DOI 10.1007/s10826-017-0958-3

Funderburk, B. W., Eyberg, S. M., Rich, B. A., & Behar, L. (2003). Further psychometric evaluation of the Eyberg and Behar rating scales for parents and teachers of preschoolers. *Early Education and Development, 14*(1), 67-81.

Garcia, D., Blizzard, A. M., Peskin, A., Rothenberg, W. A., Schmidt, E., Piscitello, J., Espinosa, N., Salem, H., Rodriguez, G. M., Sherman, J. A., Landa, A. L., Weinstein, A., Davis, E. M., Parlade, M. V., Garcia, A., Perez, C., Rivera, J. M., Martinez, C., & Jent, J. F. (2021). Rapid scale up of telehealth for PCIT during the COVID-19 pandemic: Implementation and clinical implications. *Prevention Science.* Advanced online publication. https://doi.org/10.1007/s11121-021-01211-0

Ginn, N. C., Clionsky, L. N., Eyberg, S. M., Warner-Metzger, C., & Abner, J. P. (2015). Child-directed interaction training for young children with Autism Spectrum Disorders: Parent and child outcomes. *Journal of Clinical Child & Adolescent Psychology,* 1-9. DOI:10.1080/15374416.2015.1015135

Giorgi, A. (2020). COVID-19 and rural broadband: Progress, problems, and a long way to go. *The Daily Yonder.* Retrieved from https://www.dailyyonder.com/covid-19-and-rural-broadband-progress-problemsand-a-long-way-to-go/2020/04/28/

Glueckauf, R. L., Maheu, M. M., Drude, K. P., Wells, B. A., Wang, Y., Gustafson, D. J., & Nelson, E. (2018). Survey of psychologists' telebehavioral health practices: Technology use, ethical issues, and training needs. *Professional Psychology: Research and Practice, 49*(3), 205-219. http://dx.doi.org/10.1037/pro0000188

Gruber, J., Prinstein, M. J., Clark, L. A., Rottenberg, J., Abramowitz, J. S., Albano, A. M., Aldao, A., Borelli, J. L., Chung, T., Davila, J., Forbes, E. E., Gee, D. G., Hall, G. C. N., Hallion, L. S., Hinshaw, S. P., Hofmann, S. G., Hollon, S. D., Joormann, J., Kazdin, A. E., Klein, D. N., La Greca, A. M., Levenson, R. W., MacDonald, A. W., III, McKay, D., McLaughlin, K. A., Mendle, J., Miller, A. B., Neblett, E. W., Nock, M., Olatunji, B. O., Persons, J. B., Rozek, D. C., Schleider, J. L., Slavich, G. M., Teachman, B. A., Vine, V., & Weinstock, L. M. (2020). Mental health and clinical psychological science in the time of COVID-19: Challenges, opportunities, and a call to action. *American Psychologist*, Advance online publication. http://dx.doi.org/10.1037/amp0000707

Guo, S., & Frasher, M. W. (2010). *Propensity score analysis: Statistical methods and applications.* Thousand Oaks, CA: Sage Publications

Hare, M. M. & Graziano, P. A. (2020). The cost-effectiveness of parent-child interaction therapy: Examining standard, intensive, and group adaptations. *Administration and Policy in Mental Health Services Research*. https://doi.org/10.1007/s10488-020-01083-6

Harvey, A. G., Gumport, N. B. (2015). Evidence-based psychological treatments for mental disorders: Modifiable barriers to access and possible solutions. *Behaviour Research and Therapy, 68,* 1-12. http://dx.doi.org/10.1016/j.brat.2015.02.004

Harvey, E. A., Fischer, C., Weieneth, J. L., Hurwitz, S. D., & Sayer, A. G. (2013). Predictors of discrepancies between informants' ratings of preschool-aged children's behavior: An examination of child characteristics, and family functioning. *Early Childhood Research Quarterly, 28*, 668-682. http://dx.doi.org/10.1016/j.ecresq.2013.05.002

Hennekens, C. H. & Buring, J. E. (1987). Case-control studies. Pp. 132-152. In: Mayrent, S. L., (Ed.). *Epidemiology in Medicine.* Philadelphia: Lippincott Williams & Wilkins.

Hook, J. N., Watkins Jr., C. E., Davis, D. E., Owen, J., Van Tongeren, D. R., & Ramos, M. J. (2016). Cultural humility in psychotherapy supervision. *American Journal of Psychotherapy, 70*(2), 149-166.

Ingersoll, B. & Berger, N. I. (2015). Parent engagement with a telehealth-based parent-mediated intervention program for children with autism spectrum disorders: Predictors of program use and parent outcomes. *Journal of Medical Internet Research, 17*(10), DOI: 10.2196/jmir.4913: 10.2196/jmir.4913

Iruka, I. U., Curenton, S. M., Sims, J., Escayg, K-A., Ibekwe-Okafor, N., & RAPID-EC. (2021). *Black parent voices: Resilience in the face of the two pandemics—COVID-19 and racism.* Researchers Investigating Sociocultural Equity and Race (RISER) Network.

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67,* 300–307.

Jacobson, N. S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.

Kazdin, A. E. (1995). Child, parent, and family dysfunction as predictors of outcome in cognitive-behavioral treatment of antisocial children. *Behaviour Research and Therapy, 33,* 271-281.

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science, 6,* 21–37. http://dx.doi.org/10.1177/1745691610393527

Kaminski, J. W. Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36,* 567-589. DOI: 10.1007/s10802-007-9201-9

Kazdin, A. E. (1990). Premature termination from treatment among children referred for antisocial behavior. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 31*, 415-425.

Kohlhoff, J., Wallace, N., Morgan, S., Maiuolo, M., & Turnell, A. (2019). Internet-delivered parent-child interaction therapy: Two clinical case reports. *Clinical Psychologist, 23*, 271-282. DOI: 10.1111/cp.12184

Kraetschmer, N. M., Deber, R. B., Dick, P., & Jennett, P. (2009). Telehealth as gatekeeper: Policy implications for geography and scope of services. *Telemedicine and E-Health, 15*(7), 655-663. DOI: 10.1089/tmj.2009.0004

Krahé, B., Bondü, R., Höse, A., & Esser, G. (2014). Child aggression as a source and consequence of parenting stress: A three-wave longitudinal study. *Journal of Research on Adolescence, 25*(2), 328-339. DOI: 10.1111/jora.12115

Kruse, C. S., Karem, P., Shifflett, K., Vegi, L., Ravi, K., & Brooks, M. (2018). Evaluating barriers to adopting telemedicine worldwide: A systematic review. *Journal of Telemedicine and Telecare, 24*(1), 4-12. DOI: 10.1177/1357633X16674087

Lamis, D. A., West, L. M., Mehta, N., Lisco, C., Tarantino, N., Wilson, C., & Kaslow, N. J. (2014). Existential well-being, drug abuse, and parenting stress in African American women. *International Journal of Mental Health & Addiction, 12*, 686-699. DOI: 10.1007/s11469-014-9503-4

Lane, F. C., To, Y. M., Shelley, K., & Henson, R. K. (2012). An illustrative example of propensity score matching with education research. *Career and Technical Education Research, 37*(3), 187-212. DOI: 10.5328/cter37.3.187

Lerman, D. C., O'Brien, M. J., Neely, L., Call, N. A., Tsami, L., Schieltz, K. M., Berg, W. K., Graber, J., Huang, P., Kopelman, T., & Cooper-Brown, L. J. (2020). Remote coaching of caregivers via telehealth: Challenges and potential solutions. *Journal of Behavioral Education, 29*, 195-221. https://doi.org/10.1007/s10864-020-09378-2

Lieneman, C. C., Brabson, L. A., Highlander, A., Wallace, N. M., & McNeil, C. B. (2017). Parent-child interaction therapy: Current perspectives. *Psychology Research and Behavior Management*, *10*, 239-256.

Lindgren, S., Wacker, D., Suess, A. Schieltz, K., Peizel, K., Kopelman, T., Lee, J., Romani, P., & Waldron, D. (2016). Telehealth and autism: Treating challenging behavior at lower cost. *Pediatrics*, *137*(S2). e20152851O

Little, L. M., Wallisch, A., Pope, E., & Dunn, W. (2018). Acceptability and cost comparison off a telehealth intervention for families of children with autism. *Infants & Young Children, 31*(4), 275-286). DOI: 10.1097/IYC.0000000000000126

Liu, S. R. & Modir, S. (2020). The outbreak that was always here: Racial trauma in the context of COVID-19 and implications for mental health providers. *Psychological Trauma: Theory, Research, Practice, and Policy, 12*(5), 439-442. http://dx.doi.org/10.1037/tra0000784

Luxton, D. D., Pruitt, L. D., & Osenbach, J. E. (2014). Best practices for remote psychological assessment via telehealth technologies. *Professional Psychology: Research and Practice, 45*(1), 27-35. DOI: 10.1037/a0034547

Mandal, R. L., Olmi, D. J., Edwards, R. P., Tingstrom, D. H., & Benoit, D. A. (2000). Effective instruction delivery and time-in: Positive procedures for achieving child compliance. *Child & Family Behavior Therapy, 22*(4), 1-12. DOI: 10.1300/J019v22n04_01

Martin, J. L., & Ross, H. (2005). Sibling aggression: Sex differences and parents' reaction. *International Journal of Behavioral Development, 29,* 129–138. DOI:10.1080/01650250444000469.

McCabe, K. M., Sakamoto, M., Rosas, Y. G., La, R., Zerr, A., & Yeh, M. (2021). Keeping an "I" on PRIDE: Measuring imitation in parent-child interaction therapy. *Behavior Therapy, 52*, 28-38.

McLellan, L. F., Andrijic, V., Davies, S., Lyneham, H. J., & Rapee, R. M. (2017). Delivery of a therapist-facilitated telecare anxiety program to children in rural communities: A pilot study. *Behaviour Change, 34*(3), 156-167. DOI: 10.1017/bec.2017.11

McClellan, M. J., Florell, D., Palmer, J., & Kidder, C. (2020). Clinician telehealth attitudes in a rrurral community mental health center setting. *Journal of Rural Mental Health, 44*(1), 62-73. http://dx.doi.org/10.1037/rmh0000127

McNeil, C. B. & Hembree-Kigin, T. L. (2011). *Parent-child interaction therapy, 2nd edition.* New York: Springer Science.

Mental Illness, Bethesda, MD, National Institutes of Health, 2017. https://www.nimh.nih.gov/health/statistics/mental-illness.shtml

McGoron, L. & Ondersma, S. J. (2015). Reviewing the need for technological and other expansions of evidence-based parent training for young children. *Children and Youth Services Review, 59*, 71-83. http://dx.doi.org/10.1016/j.childyouth.2015.10.012

Miller, C. J., Smitch, S. N., & Pugatch, M. (2020). Experimental and quasi-experimental designs in implementation research. *Psychiatry Research, 283*, 1-7. https://doi.org/10.1016/j.psychres.2019.06.027

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674–701.

Mosher, D. K., Hook, J. N., Captari, L. E., Davis, D. E., DeBlaere, C., & Owen, J. (2017). Cultural humility: A therapeutic framework for engaging diverse clients. *Practice Innovations, 2*(4), 221-233. http://dx.doi.org/10.1037/pri0000055

Niec, L. N. (2018). Parent-child interaction therapy: A transdiagnostic intervention to enhance family functioning. *Handbook of Parent-Child Interaction Therapy, pp. 3-15.* https://doi.org/10.1007/978-3-319-97698-3_1

Niven, D. J., Berthiaume, L. R., Fick, G. H., & Laupland, K. B. (2012). Matched case-control studies: A review of reported statistical methodology. *Clinical Epidemiology, 4*, 99-110.

Ofonedu, M. E., Belcher, H. M. E., Budhathoki, C., & Gross, D. A. (2016). Understanding barriers to initial treatment engagement among underserved families seeking mental health services. *Journal of Child and Family Studies, 26*, 863-876. DOI 10.1007/s10826-016-0603-6

Olthuis, J. V., McGrath, P. J., Cunningham, C. E., Boyle, M. H., Lingley-Pottie, P., Reid, G. J., Bagnell, A., Lipman, E. L., Turner, K., Corkum, P., Stewart, S. H., Berrigan, P., & Sdao-Jarvie, K. (2018). Distance-delivered parent training for childhood disruptive behavior (Strongest Families): A randomized controlled trial and economic analysis. *Journal of Abnormal Child Psychology, 46*, 1613-1629. https://doi.org/10.1007/s10802-018-0413-y

Patallo, B. J. (2019). The multicultural guidelines in practice: Cultural humility in clinical training and supervision. *Training and Education in Professional Psychology, 13*(3), 227-232. http://dx.doi.org/10.1037/tep0000253

Patterson, G. R., Capaldi, D., & Bank, L. (1991). An early starter model for predicting delinquency. In D. Pepler & R. K. Rubin (Eds.), *The development and treatment of childhood aggression.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Pew Research Center. (2019, June 12). *Mobile Fact Sheet.* https://www.pewresearch.org/internet/fact-sheet/mobile/

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroer, J. C., & Huang, L.N. (2013). Mental health surveillance among children – United States, 2005-2011. *Morbidity and Mortality Weekly Report, 62(Suppl. 2)*, 1–35.

Perry, K., Gold, S., & Shearer, E. M. (2019). Identifying and addressing mental health providers' perceived barriers to telehealth utilization. *Journal of Clinical Psychology,* 1-10. DOI: 10.1002/jclp.22770

Peterson, K. M., Piazza, C. C., Luczynski, K. C., & Fisher, W. W. (2017). Virtual-care delivery of applied-behavior-analysis services to children with autism spectrum disorder and related conditions. *Behavior Analysis: Research and Practice, 17*(4), 286-297. http://dx.doi.org/10.1037/bar0000030

Phaneuf, L. & McIntyre, L. L. (2011). The application of a three-tier model of intervention to parent training. *Journal of Positive Behavior Intervention*, *13*(4), 198-207. DOI:10.1177/1098300711405337

Pierce, B. S., Perrin, P. B., Tyler, C. M., McKee, G. B., & Watson, J. D. (2020). The COVID-19 telepsychology revolution: A national study of pandemic-based changes in U.S. mental health care delivery. *American Psychologist, online first publication.* http://dx.doi.org/10.1037/amp0000722

Reardon, T., Harvey, K., Baranowska, M., O'Brien, D., Smith, L., & Creswell, C. (2017). What do parents perceive are the barriers and facilitators to accessing psychological treatment for mental health problems in children and adolescents? A systematic review of qualitative and quantitative studies. *European Child and Adolescent Psychiatry, 26*, 623-647. DOI 10.1007/s00787-016-0930-6

Reese, R. J., Slone, N. C., Soares, N., & Sprang, R. (2012). Telehealth for underserved families: An evidence-based parenting program. *Psychological Services, 9*(3), 320-322. DOI: 10.1037/a0026193

Reitman, D., Currier, R. O., & Stickle, T. R. (2002). A critical evaluation of the parenting stress index-short form (PSI-SF) in a head start population. *Journal of Clinical Child and Adolescent Psychology, 31*(3), 384-392. DOI: 10.1207/S15374424JCCP3103_10

Reynolds, C. R., & Kamphaus, R. W. (2015). *Behavior Assessment System for children manual* (3rd ed.). Bloomington, MN: Pearson.

Riegler, L. J., Raj, S. P., Moscato, E. L., Narad, M. E., Kincaid, A., & Wade, S. L. (2020). Pilot trial of a telepsychotherapy parenting skills intervention for veteran families: Implications for managing parenting stress during COVID-19. *Journal of Psychotherapy Integration, 30*(2), 290-303. http://dx.doi.org/10.1037/int0000220

Rijlaarsdam, J., Tiemeier, H., Ringoot, A. P., Ivanova, M. Y., Jaddoe, V. W. V., Verhulst, F. C., & Roza, S. J. (2016). Early family regularity protects against later disruptive behavior. *European Child and Adolescent Psychiatry, 25*, 781-789. DOI: 10.1007/s00787-015-0797-y

Rios, D., Kazemi, E., & Peterson, S. M. (2018). Best practices and considerations for effective service provision via remote technology. *Behavior Analysis: Research and Practice, 18*, 277–287. https://doi.org/10.1037/bar00 00072

Roberts, D. S., Tingstrom, D. H., Olmi, D. J., & Bellipanni, K. D. (2008). Positive antecedent and consequent components in child compliance training. *Behavior Modification, 32*(1), 21-38. DOI: 10.1177/0145445507303838

Romani, P. W. & Schieltz, K. M. (2017). Ethical considerations when delivering behavior analytic services for problem behavior via telehealth. *Behavior Analysis: Research and Practice, 17*(4), 312-324. http://dx.doi.org/10.1037/bar0000074

Rubin, K. H., Coplan, R. J., Chen, X., Bowker, J., & Macdonald, K. (2010). Peer relationships in childhood. In M. Bornstein & M. Lamb (Eds.), *Developmental science: An advanced textbook* (6th ed., pp. 519–570). New York: Psychology Press.

Rubin, D. B. & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics, 52*(1), 249-264. https://doi.org/10.2307/2533160

Schafer, J. L. & Kang, J. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods, 13*(4), 279-313. http://dx.doi.org.access.library.miami.edu/10.1037/a0014268

Schueler, C. M. & Prinz, R. J. (2013). The role of caregiver contingent responsiveness in promoting compliance in young children. *Child Psychiatry and Human Development, 44*, 370-381. DOI: 10.1007/s10578-012-0331-0

Scott, S., Augimeri, L. K., & Fifield, J. (2017). In S. Bailey, P. Chitsabesan, P. Tarbuck (Eds.), *Forensic child and adolescent mental health: Meeting the needs of young offenders.* (pp. 239-253). New York, NY: Cambridge University Press.

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology, 3*(111), 1-6. DOI: 10.3389/fpsyg.2012.00111

Shachar, C., Engel, J., & Elwyn, G. (2020). Implications for telehealth in a postpandemic future: Regulatory and privacy issues. *JAMA, 323*(23), 2375-2376. DOI:10.1001/jama.2020.7943

Sloan, N. C., Reese, R. J., & McClellan, M. J. (2012). Telepsychology outcome research with children and adolescents: A review of the literature. *Psychological Services, 9*(3), 272-292. DOI:10.1037/a0027607

Smit, F., Cuijpers, P., Oostenbrink, J., Batelaan, N., de Graaf, R., & Beekman, A. (2006). Costs of nine common mental disorders: Implications for curative and preventive psychiatry. *Journal of Mental Health Policy and Economics, 9*(4), 193–200.

Stewart, R. W., Orengo-Aguayo, R. E., Cohen, J. A., Mannarino, A. P., & de Arellano, M. A. (2017). A pilot study of trauma-focused cognitive-behavioral therapy delivered via telehealth technology. *Child Maltreatment, 22*(4), 324-333. DOI: 10.1177/1077559517725403

Summers-Gabr, N. M. (2020). Rural-urban mental health disparities in the United States during COVID-19. *Psychological Trauma: Theory, Research, Practice, and Policy, 12*(S1), S222-S224. http://dx.doi.org/10.1037/tra0000871

Swan, A. J., Carper, M. M., & Kendall, P. C. (2016). In pursuit of generalization: An updated review. *Behavior Therapy, 47*, 733-746. http://dx.doi.org/10.1901/jaba.2000.33-167

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research, 46*, 90-118. doi:10.1080/00273171.2011.540475

Thomas, R., Abell, B., Webb, H. J., Avdagic, E. & Zimmer-Gembeck, M. J. (2017). Parent-child interaction therapy: A meta-analysis. *Pediatrics, 140*(3), 1-15. DOI: 10.1542/peds.2017-0352

Tse, Y. J., McCarty, C. A., Stoep, A. V., & Myers, K. M. (2015). Teletherapy delivery of caregiver behavior training for children with attention-deficit hyperactivity disorder. *Telemedicine and e-Health, 21*(6), 451-458. DOI: 10.1089/tmj.2014.0132

Tung, I., Li, J. J., & Lee, S. S. (2012). Child sex moderates the association between negative parenting and childhood conduct problems. *Aggressive Behavior, 38,* 239-251. DOI: 10.1002/ab.21423

Turvey, C. L. & Myers, K. (2013). Research in telemental health: Review and synthesis. In K. Myers & C. L. Turvey (Eds.) *Telemental Health: Clinical, Technical, and Administrative Foundations for Evidence-Based Practice.* Amsterdam: Elsevier.

Unholz-Bowden, E., McComas, J. J., McMaster, K. L., Girtler, S. N., Kolb, R. L., & Shipchandler, A. (2020). Caregiver training via telehealth on behavioral procedures: A systematic review. *Journal of Behavioral Education, 29*, 246-281. https://doi.org/10.1007/s10864-020-09381-7

Vismara, L. A., McCormick, C., Young, G. S., Nadhan, A., & Monlux, K. (2013). Preliminary findings of a telehealth approach to parent training in autism. *Journal of Autism and Developmental Disorders, 43*, 2953-2969. DOI: 10.1007/s10803-013-1841-8

Wade, S. L., Oberjohn, K., Conaway, K., Osinska, P., & Bangert, L. (2011). Live coaching of parenting skills using the internet: Implications for clinical practice. *Professional Psychology: Research and Practice, 42*(6), 487-493. DOI: 10.1037/a0025222

Wainer, A. L. & Ingersoll, B. R. (2015). Increasing access to an ASD imitation intervention via a telehealth parent training program. *Journal of Autism and Developmental Disorders, 45*, 3877-3890. DOI: 10.1007/s10803-014-2186-7

Washington State Institute for Public Policy. (2019, December). *Benefit-Cost Results.* https://www.wsipp.wa.gov/BenefitCost?topicId=5

Webb, C. & Orwig, J. (2015). Expanding our reach: Telehealth and licensure implications for psychologists. *Journal of Clinical Psychology in Medical Settings, 22,* 243-250. DOI: 10.1007/s10880-015-9440-9

Weisz, J. R., Sandler, I. N., Durlak, J. A., & Anton, B. S. (2005). Promoting and protecting youth mental health through evidence based prevention and treatment. *American Psychologist, 60,* 628–648. http://dx.doi.org/10.1037/0003-066X.60.6.628

Wen Yang, H., Burke, M., Isaacs, S., Rios, K., Schraml-Block, K., Aleman-Tovar, J., Tompkins, J., & Swartz, R. (2020). Family perspectives toward using telehealth in

early intervention. *Journal of Developmental and Physical Disabilities.* https://doi.org/10.1007/s10882-020-09744-y

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychological journals: Guidelines and explanations. *American Psychologist, 54*, 594–604. http://dx.doi.org/10.1037/0003-066X.54.8.594

World Health Organization. (2020). *The impact of COVID-19 on mental, neurological and substance use services: Results of a rapid assessment.*

Zickuhr, K., & Smith, A. (2012). Digital differences. Pew Research Center Internet and American Life Project. Retrieved from http://pewinternet.org/Reports/2012/Digital-differences.aspx