Journal of Clinical Child & Adolescent Psychology, 44(5), 730–741, 2015

ISSN: 1537-4416 print/1537-4424 online
DOI: 10.1080/15374416.2014.910790

Routledge
Taylor & Francis Group

Comparing Client Outcomes for Two Evidence-Based Treatment Consultation Strategies

Beverly Funderburk, Mark Chaffin, Elizabeth Bard, Jenelle Shanley, and David Bard
University of Oklahoma Health Sciences Center

Lucy Berliner
School of Social Work, University of Washington

Posttraining expert case consultation is a key component of transporting and scaling up evidence-based treatments, and hopefully retaining their efficacy. Live practice observation and in vivo coaching is a strategy used in academic training environments, but is rarely feasible in field settings. Post hoc telephone consultation is a substitute strategy but does not approximate many aspects of live coaching. Live video technology offers a closer approximation but has not yet been sufficiently tested. Using a roll-out experimental design, this study compared client outcomes across doses of two posttraining expert consultation strategies—standard telephone consultation and live video coaching. The study was conducted during a two-state, 30-agency implementation involving 80 therapists and 330 cases receiving Parent–Child Interaction Therapy (PCIT). Child behavior problems fell from well above to below clinical cutoff values, with about a 1 standard deviation improvement in 14 sessions, which is within the range reported in laboratory efficacy trials. Symptom improvement was augmented by increased therapist dose of live video consultations. Phone consultation dose had no association with client level outcomes. PCIT benefits appear to be retained when the model is transported at scale into the field, and live video consultation appeared to offer small but significant advantages over telephone consultation as one element of an overall transport strategy.

This experimental trial compares two posttraining therapist consultation methods for an evidence-based treatment (EBT); the first is standard post hoc group phone consultation and the second is live video consultation and therapist skill coaching. The consultation conditions are compared for their dose-response impact on the EBT's main intended client outcome of reduced child behavior problems. The study was part of an implementation and scale-up initiative for a child EBT model at 30 participating agencies in two states.

Technology transfer between EBT development and field practice settings has been a priority for over a decade (Institute of Medicine, 2001). Clinical training and consultation are key activities in this process (Nadeem, Gleacher, & Beidas, 2013), although it should be borne in mind that that clinical training alone is rarely sufficient to achieve implementation. Transporting and implementing EBTs involves tasks unfolding across multiple phases and at multiple levels including individual treatment providers; organizational culture and climate; organizational leadership; and the outer community, system, and funding contexts (Aarons, Hurlburt, & Horwitz, 2011). There have been few true experimental tests of how variations in EBT training or consultation methods may impact downstream client outcomes during EBT scale-up (Herschell, Kolko, Baumann, & Davis, 2010).

Comparative outcome trials of posttraining consultation approaches have the potential to inform two questions—the field effectiveness of the EBT and how variations in implementation approach impact

Correspondence should be addressed to Mark Chaffin, University of Oklahoma Health Sciences Center, P.O. Box 26901, Nicholson Suite 4900, Oklahoma City, OK 73126. E-mail: mark-chaffin@ouhsc.edu

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hcap.

effectiveness. These types of studies may be described as "hybrid" implementation-effectiveness designs or pragmatic trials (Curran, Bauer, Mittman, Pyne, & Stetler, 2012; Glasgow, 2013).

One recent hybrid trial found that clients of paraprofessional home visitors assigned to receive live in vivo EBT coaching, as compared to post hoc verbal EBT consultation, achieved slightly better child welfare recidivism outcomes (Chaffin, Hecht, Bard, Silovsky, & Beasley, 2012). Given that child welfare recidivism is an outcome that is notoriously difficult to affect (Filene, 2012) even a small consultation strategy effect is noteworthy. However, the provider workforce, clientele, and service location in that trial differed from traditional mental health clinic settings. In this study, we examine the impact of a technology-based version of live consultation and in vivo coaching conducted in a more traditional clinic context and (a) employ an experimental design within a scaled-up multiagency implementation; (b) compare two competing, presumably active, and good quality posttraining consultation strategies; (c) use a proven efficacious EBT; and (d) track consultation strategy dose-response impact on the EBT's main intended client outcome.

Developing clinical competency with soft technologies such as a mental health EBT is a challenge. Cases present in complex and idiosyncratic ways, providers bring diverse preexisting practice loyalties and skills, and the nuances of model competency and expertise often are difficult or impossible to convey in even the best manual or workshop. Workshops or training courses alone may transfer basic model knowledge and some nascent skills but often are insufficient for achieving technology transfer or generating outcomes (Herschell, Kolko, Baumann, & Davis, 2010; Joyce & Showers, 2002). Posttraining consultation in the practice setting is believed to be critical for skill development (Beidas, Edmunds, Marcus, & Kendall, 2012; Nadeem, Gleacher, & Beidas, 2013). In their synthesis of the implementation research literature, Fixsen, Naoom, Blase, Friedman, and Wallace (2005) noted that

> newly learned behavior is crude compared to performance by a master practitioner. There are uncounted nuances of when and how to use the components in various combinations. . . . This functional and adaptable set of skills is developed in practice with the help of a consultant/coach who shares craft knowledge as he or she observes, describes and tutors the practitioner. (p. 44)

Drift occurs easily after training, and in the few instances where clinical outcomes have been studied, drift has been found to result in loss of effectiveness (Elliott & Mihalic, 2004; Schoenwald, Letourneau, & Halliday-Boykins, 2005). Avoiding drift, mastering the EBT, and gaining more nuanced aspects of EBT model competency in academic or laboratory research settings relies on the rich fidelity monitoring, expert consultation, mentoring, and in vivo coaching resources available in these contexts. These same resources are rarely present or feasible in implementation contexts, especially scale-up efforts, raising the question of how much and what type of posttraining consultation strategy should be used. There is evidence that some posttraining consultation is better than no consultation, at least for therapist knowledge and observed skills, if not always for downstream client outcomes (Edmunds, Beidas, & Kendall, 2013). Many consultation studies examine traditional post hoc types of consultation, not live skill coaching. There are some suggestions in observational studies of post hoc consultation that techniques similar to those used in live coaching (e.g., modeling, practice, rehearsal) may predict greater EBT uptake in the subsequent session, although the effects differ across therapist characteristics, and linkage with client outcomes has yet to be examined (Bearman et al., 2013).

The EBT examined in this study was PCIT. PCIT is a behavioral parent training model developed as a dyadic, parent-mediated treatment for early childhood disruptive behavior problems. Efficacy of the model is well established and dates back 20 years. Controlled laboratory trials have demonstrated the model's efficacy in decreasing child disruptive behaviors (e.g., Eisenstadt, Eyberg, McNeil, Newcomb, & Funderburk, 1993; McNeil, Capage, Bahl, & Blanc, 1999), increasing child compliance with parental requests (e.g., Eyberg & Robinson, 1982), improving the parent–child relationship (e.g., Eyberg, Boggs, & Algina, 1995), and reducing parental stress (e.g., Schuhmann, Foote, Eyberg, Boggs, & Algina, 1998). Child behavior improvements have been found to generalize from the controlled clinic setting to the home environment (e.g., Schuhmann et al., 1998), as well as from the home to school classrooms, and from treated children to untreated siblings (McNeil, Eyberg, Eisendstadt, Newcomb, & Funderburk, 1991). A meta-analysis of 13 controlled trials reported an effect size range of .61 to 1.45 compared with untreated children (Thomas & Zimmer-Gembeck, 2007). In addition to child behavior problem effects, the model has been used as the major component of interventions shown to lower the risk for recurrent abusive parental behavior (Chaffin, Funderburk, Bard, Valle, & Gurwitch, 2011; Chaffin et al., 2004). Scale-up of the model has been undertaken in several states and is being promoted by Substance Abuse and Mental Health Services Administration's National Child Traumatic Stress Initiative.

PCIT proceeds over a planned average course of about 14 sessions (but ranging up to about 24), during which therapists teach parents a set of parent–child interaction and child behavior management skills until skill criterions are reached and child behavior problems

have improved. A distinguishing feature of PCIT is that the parenting skills are coached in vivo during live parent–child interactions using a wireless earphone, over which the therapist coaches the parent from behind a one-way mirror. PCIT was developed and for years was trained almost exclusively within university-affiliated training programs, such as graduate degree programs, internships, and postdoctoral fellowships. In PCIT academic training settings, live practice coaching is the dominant consultation strategy. Indeed, the aphorism often heard among PCIT training faculty is, "The best model for teaching PCIT is PCIT." Similar to the way PCIT therapists directly coach parents in vivo, PCIT trainers directly coach academic trainees during live sessions. Traditional academic PCIT training progresses through two phases. First, new trainees complete basic didactic training, including the theory base for the model, the structure and set-up of treatment, session-by-session protocols, and basic skills. This first phase usually includes observing the work of master practitioners and role-plays of basic techniques, and usually takes about 1 full week. This first training phase is reasonably transferrable into implementation and scale-up settings. During the second academic training phase, trainees are mentored using a live co-therapy approach through several cases. The traditional second phase does not transfer well to field implementations and especially to scale-up initiatives because few agencies can import an expert trainer/mentor/co-therapist/coach into their agency for several months. This same dilemma may be shared by other EBTs that might favor live coaching as a mentoring or competency development strategy.

In practice, field implementation often skips this second phase of training, or offers an abbreviated version (e.g., a single site visit or two), and instead relies primarily on post hoc verbal phone consultation during which trainees will discuss impressions of their cases and their practice with a remote expert. Post hoc verbal consultation has the advantage of being readily feasible, and it is a venerable psychotherapy tradition, but it also has several potential disadvantages. There may be limited correspondence between how trainees conceptualize or describe their clients and their practice after the fact and how practice is behaviorally delivered in vivo (Herschell et al., 2010). Using video recordings as an adjunct to post hoc verbal consultation may help but still does not entirely replicate the live coaching experience. For example, a trainee may encounter an idiosyncratic practice issue, then present the recorded session for consultation. The consultant may make suggestions but cannot know whether the suggestions are implemented correctly until (a) the idiosyncratic implementation issue occurs again in a future session, (b) the trainee remembers what to do, and (c) the trainee's new response to that issue happens to be recorded and presented again for consultation. This process, which occurs immediately and in real-time during live coaching, may take weeks to occur in post hoc consultation using videos, if it occurs at all. We hypothesize that live coached consultation is superior, but there is limited evidence to confirm this, especially when it comes to ultimate client outcomes. The technology based live-coached consultation strategy tested in this study was designed to mimic key aspects of traditional academic live-coached PCIT training, but in a way that is reasonably feasible for multiagency scale-ups. The key research question is whether live video consultation will improve downstream client outcomes (which we view as the evidence-based practice bottom-line) relative to standard phone consultation.

METHOD

Participants

There were two classes of participants in the study—PCIT therapists ($n = 80$) and parent–child client dyads ($n = 330$). Participants were recruited from 30 participating agencies. Thirty-two agencies were solicited for participation, out of which 30 agreed. Recruitment among therapists within agencies was 100%. Two participating therapists withdrew over the course of the study. Therapists and agencies agreed to recruit all clients receiving PCIT, and this was monitored during regular clinical consultation. However, no data were made available to the study about clients who did not consent to participate. Among the therapists, 90% were female, 4% were ethnically Hispanic, 86% were Caucasian, 5% Asian, 5% Native, and 3% African American. Therapists were on average 37 years old, and most (92%) had a master's degree, predominantly in the disciplines of psychology (53%) or social work (26%). Of the therapists, 84% were full-time employees of their agency. Therapists had a mean of 6.3 years of clinical experience, with an average of 6.2 years delivering parenting services and 5.4 years delivering children's services. Most therapists endorsed a cognitive-behavioral theoretical orientation (65%) followed by family systems (18%), interpersonal (6%), behavioral (3%), psychodynamic (3%), or other (5%). On average, each therapist treated 4.1 study patients (range = 1–18) over the course of the study, enrolling their first study patient a median of 1 month after completing their PCIT training workshop. Child clients were 63% boys, with a mean age of 4.65 years ($SD = 1.6$). They were most often the biological children (72%) of the caregiver in their treatment dyad, followed by grandchildren (10%), adopted children (8%), foster children (5%), or other (6%). The mean baseline Eyberg

Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999) Intensity Score was 155, and the mean ECBI Problem Score was 19, both well above clinical range cutoff values of 132 and 15, respectively. Caregivers were 82% female; 9% were ethnically Hispanic, 70% were Caucasian, 12% African American, 1% Asian, 13% American Indian, and 4% unspecified. Of the caregivers, 38% were currently married, 13% cohabitating, 28% divorced or separated, 17% never married, and 4% other. Thirteen percent of caregivers had completed at least a bachelor's degree, 47% had some postsecondary education less than a bachelor's degree, 26% were high school graduates, and 12% had not completed high school. We excluded a small number (seven) of therapists and their eight cases because the therapist was trained in PCIT many years prior and had received extensive prestudy PCIT consultation. We also dropped 46 clients who had only a baseline/intake assessment and no treatment assessments because these cases had no exposure to their therapist's experimental conditions (i.e., an active treatment session that might be influenced by consultation) or if the density of contacts during the treatment episode was less than one per month (potentially confounding the session-by-session coding scheme for the growth models). We also excluded a very small number of sessions that occurred after a long hiatus from treatment. Excluded cases did not significantly differ from included cases on any demographic variable or baseline EBCI score. Nine clients withdrew from the study but still allowed use of data provided prior to their withdrawal. The final analysis sample included 276 cases and 73 therapists, with observed data on a total of 2,561 PCIT treatment sessions.

Implementation context. Data were drawn from two publically funded implementation initiatives located in Oklahoma and Washington between August 2007 and July 2011. Participating agencies ($n = 30$) included a mix of state-run community mental health centers and health departments, tribal social services agencies, community-based nonprofits, church-affiliated nonprofits, and private for-profit providers. The study capitalized on the opportunity presented by these combined start-up efforts, which funded both the initial PCIT workshop trainings and some clinical service delivery. The research team's role in the overall implementation effort was limited to providing workshop training and all of the posttraining expert consultation received by the organizations. Initial PCIT workshops were conducted by two different PCIT training organizations that worked to align and harmonize their training approaches and used standard PCIT training materials. The study principal investigator (PI) reviewed and observed training delivered by both training teams and consulted with both training groups to promote consistency. Training predated development of PCIT International training and certification criteria (PCIT International, 2013). However, the study PI and other study investigators and trainers were closely involved in developing the official PCIT training criteria, and we believe that the training received by study therapists was congruent with subsequent official PCIT training standards.

Measures and Procedures

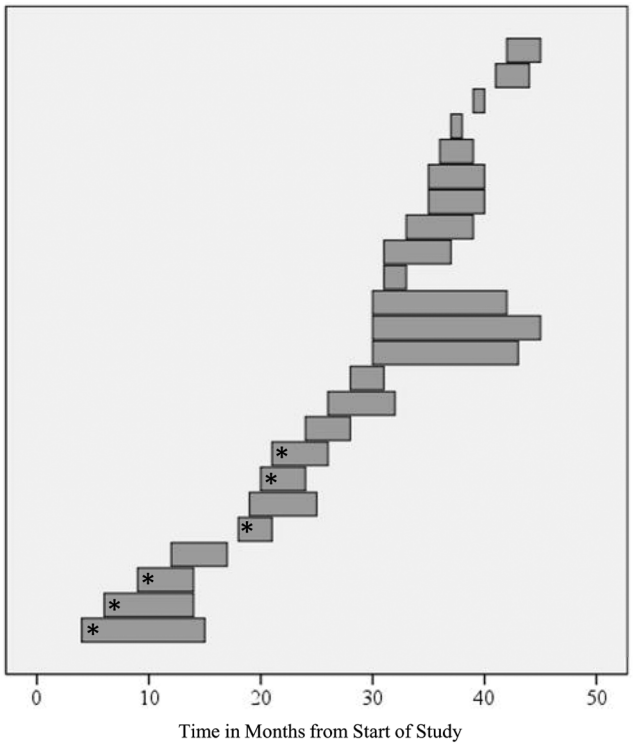
Consent procedures. All research procedures and the consent process were reviewed and approved by two university Institutional Review Boards, one in each state. Therapists were recruited by research assistants, with the understanding that agency administrators would not be informed about which of their therapists were study participants. Therapist study participation status did not affect the type or amount of training and consultation the therapist received. Clients were recruited by therapists and completed a written informed consent procedure authorizing their routine clinical data (e.g., basic demographic information, weekly ECBI, and Therapy Attitude Inventory [TAI] forms) to be shared with the study. Client study participation status did not affect the type of treatment received or data collected, only the sharing of that data with the study. Clients provided a separate consent to participate in live video consultation, which involved direct observation by a remote expert. No study related adverse events were reported.

Standard telephone consultation procedures. To preface our description of consultation conditions, it is important to distinguish consultation from supervision. By consultation we mean case review and guidance offered by an outside EBT model expert. By supervision, we mean the agency's internal, formal, and authoritative practice controls. The formal agency supervisor, not the external consultant, retained authority over practice. Consultation did not replace or supersede agency supervision; it augmented it. Telephone consultation used conference call technology and was scheduled weekly. Calls were led by a remote PCIT expert in Oklahoma and included an average of six therapist participants per call. Efforts were made to keep membership in a given phone consultation group intact. A process log and checklist was used to guide phone consultations, and both actual consultation sessions and process checklists were observed by the PI to monitor fidelity. Consultants also met regularly with the PI to troubleshoot consultation issues and promote consistency. Consultant fidelity was consistently high. The phone consultation protocol began with polling providers about PCIT sessions held since the previous meeting,

collecting parent skill coding and ECBI scores from recent sessions to track case outcomes, and then staff individual cases with an emphasis on reinforcing proper PCIT technique, offering examples for case-specific tailoring, tracking progress, and setting the agenda and goals for the subsequent session. Phone consultation was accomplished once every 12 days on average, and the total dose of phone consultation varied according to the experimental design (see Figure 1; range = 0–44 sessions).

Live video consultation procedures. Live video consultation used Internet-based encrypted audio-visual technology. The technology allowed a remote PCIT consultant in Oklahoma to (a) directly observe live parent–child interaction in the treatment room; (b) listen to the PCIT therapist coach the parent; (c) communicate directly and privately back and forth with the therapist and provide skill coaching to the therapist during the session; and (d) on what proved to be very rare occasions, directly take over live coaching of parenting skills with the parent in the remote treatment room in order to model skills for the therapist. A simplified diagram of the treatment room and equipment set up is provided in Figure 2. A process log and checklist was used to guide video consultations, and both video consultation sessions and checklists were observed by the PI for fidelity. Consultants met regularly to coordinate their consultation approaches. Fidelity was consistently very high. Live video consultation was initially scheduled to occur weekly for each therapist, and could occur with different cases a therapist might have. However, in practice, it occurred about half as often as planned. Live video consultation required a number of preconditions that weren't always met, including a kept and on-time PCIT appointment without any emergent crises that might divert from the EBT protocol, scheduling the single live video consultation equipped room at the agency, client consent to be observed via live video, working equipment and Internet connections, and correct operation of the equipment. The live video consultation phase at each agency lasted about on average about 4 months ($Mdn = 112$ days), with therapists receiving an average of eight live video consultations during this time (range = 0–19), or approximately one every 14 days on average, which was similar to the once every 12 days observed for phone consultation.

A third and final follow-up consultation condition also was used. After finishing video consultation, each therapist was assigned to receive weekly follow-up phone consultation for the duration of the study (range = 0–19 consultations). Follow-up consultation followed the same protocol as initial phone consultation.



* Indicates an agency that began with video consultation, without prior phone consultation. Others had phone consultation beginning at baseline and continuing until the start of video consultation. All agencies received follow-up phone consultation after video consultation ended

FIGURE 1 Observed roll-out time sequence for video consultation, across agencies.

ECBI. The ECBI is a 36-item parent report scale of externalizing behavior in children ages of 2 to 16 (Eyberg & Pincus, 1999) developed as the main outcome measure for PCIT. The main interpretive scale of the ECBI is the Intensity Scale, which reflects the presence of common child behavior problems. The Problem Scale codes whether behaviors on the Intensity Scale pose a problem for the parent. Internal consistency of the two scales is .95 and .93 (Colvin, Eyberg, & Adams, 1999), with interrater (mother–father) reliability coefficients of .69 and .61 (Eisenstadt, McElreath, Eyberg, & Bodiford McNeil, 1994), 12-week test–retest reliability coefficients of .80 and .85 (Funderburk, Eyberg, Rich, & Behar, 2003), correlations with the CBCL Externalizing Scale of .86 and .85 (Boggs, Eyberg, & Reynolds, 1990), and 10-month test–retest reliability coefficients of .75 (Funderburk et al., 2003). In our study sample, alpha was 0.94 to 0.96, and first session to second session correlations were .82 to .86. As a normal part of PCIT treatment, parents completed the ECBI at the beginning of each session, reflecting the child's behavior during the past week.

TAI. The TAI (Eyberg, 1974) is a parent-report questionnaire designed to assess parental satisfaction

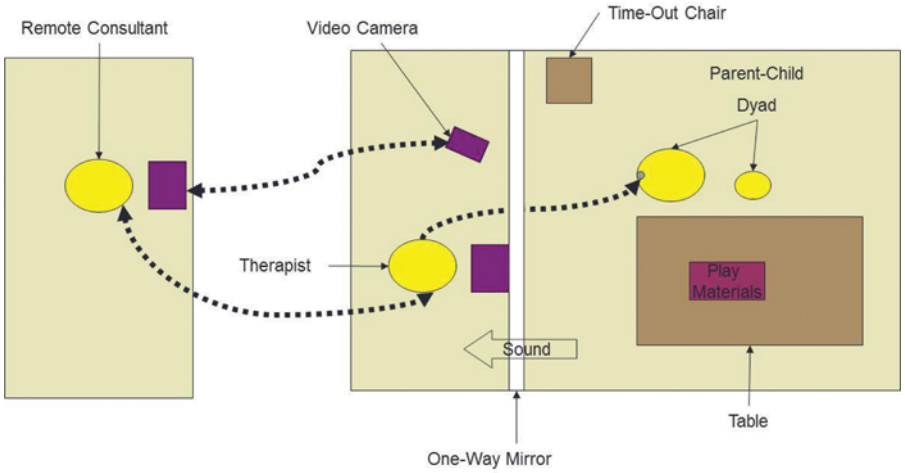


FIGURE 2 Schematic diagram of live video consultation set-up for Parent–Child Interaction Therapy.

with PCIT. Responses are given on a 1-to-5 scale. TAI scores usually show variability, so ceiling effects are less of a problem than sometimes seen with client satisfaction measures. Internal consistency is .88 (Eisenstadt et al., 1993) and 4-month test–retest stability is .85 (Brestan, Jacobs, Rayfield, & Eyberg, 1999). In the study sample, Session 1 to Session 2 correlations were .82, and alpha was 0.95.

Experimental design. The study used what is variously termed a stepped-wedge, rollout, or dynamic waitlisted experimental design (C. A. Brown & Lilford, 2006; C. H. Brown, Wyman, Gao, & Pena, 2006). Similar to an interrupted time series design, in this type of design service providers are randomly assigned to a particular spot in the implementation sequence. Implementation sequences varied by the time point in the overall implementation roll-out plan at which live video consultation was introduced (see Figure 1, which displays the actual observed video consultation start and end points for each agency). Six agencies were assigned to begin with video consultation, with no prior phone consultation. Video consultation was begun for other agencies after differing amounts of time spent in the standard phone consultation condition. Four agencies experienced only phone consultation. Data collection began for all agencies at about the same initial time point and continued across the entire study timespan as the implementation sequence unfolded. Data collection for a given client began at the client's intake and continued until service exit. Consultation type and dose was a therapist-level variable and was measured across all of a therapist's clients. Because we were manipulating consultation conditions at the therapist level, it is important to note that client outcomes could potentially be influenced by both their therapist's current consultation condition and any past consultation their therapist had received, given that learning from past consultations presumably can influence current practice. Also, because there was no manipulation of the treatment model, therapists could switch consultation conditions midtherapy for a given client. Thus, although consultation strategy was allocated as a therapist-level condition, the exact combinations of past and present consultation doses varied between therapists in the same agency, and the cumulative pattern was fairly unique for each client seen by a given therapist. Each therapist also had a fixed initial consultation sequence condition (i.e., starting with phone vs. video). The therapist's initial phase consultation condition (starting with phone vs. starting with live video) varied across therapists rather than within therapists and was fixed for all clients seen by that therapist.

Statistical analysis plan. Data were analyzed using three-level (sessions within clients within therapists) nonlinear growth models with three time-varying random dose effects, one for each type of consultation dose (phone, video, follow-up). Exploration of an initial four-level unconditional growth model (sessions within clients within therapists within agencies) suggested low ICC values for observations within agency clusters (ranging only up to .009 for any single outcome observation) relative to the therapist clusters (ranging up to .05). PCIT session number was selected as the growth rate denominator. Session-by-session outcome measures for up to 14 client sessions were modeled using a random intercept, random slope, and quadratic effect at the client level, with equality constraints imposed on score residual variances. The limit of 14 sessions was selected because it is a common PCIT treatment dose used in laboratory trials (e.g., McNeil et al., 1991) and because

only 20% of all cases in this study had additional session observations beyond this number. All latent variables in the growth model were allowed to covary at each level of the model. The three consultation doses (i.e., phone, video, or follow-up phone) were modeled as random time-varying covariates. Consultation dose effects were modeled beginning with the second client treatment session and lagged by one session so that the measured outcome at session N was predicted by the consultation dose the therapist had accrued across all cases through session $N-1$. This was done to ensure that outcomes were predicted only by therapist consultation doses to which the client was actually exposed. Cross-level interactions of dose by sequence (Consultation Dose × Whether Live Video Consultation Occurred First or Second in the Sequence) were tested, and dropped if they did not approach significance. Contrast terms were built in order to compare effect estimates for video versus phone, video versus follow-up and phone versus follow-up. Models were built and executed using MPlus 7.1 software using robust maximum likelihood estimation under the MAR assumption for missing outcome values in the growth models. No data were missing for consultation dose predictor variables. Programming syntax is is shown in the Appendix.

RESULTS

Covariate Balance Across Experimental Manipulations

We first tested for covariate or time-related imbalances in the design in order to identify potential covariates that might be needed in final models. Independence between the three experimental manipulations (therapist sequence assignment and therapist doses of phone and video consultation at client baselines) and baseline child, caregiver, and therapist characteristics were tested using a series of bivariate comparisons. Doses and sequence were not significantly associated with baseline ECBI scores, child age, child race, child ethnicity, household income, caregiver gender, caregiver race, caregiver ethnicity, caregiver's marital status, caregiver's education, caregiver's employment status, the child's relationship to the caregiver, therapist's years of clinical experience, therapist's full-time versus part-time working status, and the therapist's age. The median session spacing was one session every 10 days (mode = 7 days), and session spacing was uncorrelated with any independent variables or with the point in the overall study timeline that the client entered treatment. There was also no significant correlation between baseline ECBI scores and the point in the overall study timeline that the client entered treatment, suggesting that average baseline behavior problem severity of cases neither increased nor decreased over the course of the study. There were statistically significant associations between phone consultation dose and therapist discipline (psychology vs. social work), between consultation dose and the child's gender, and between sequence and initial PCIT training group. However, none of these three covariates approached significant prediction of ECBI latent growth variables and so were not included in final models.

Client Growth Model Outcome Analyses

For the ECBI Intensity score, which is the main child outcome of interest, no cross-level interactions with sequence approached significance and so were dropped from the final model. The intercept (or model estimated baseline) value for the ECBI Intensity scale was 152, which is well into the clinical range. These scores improved significantly in a nonlinear fashion (slope = −4.5 points/session, $SE = 0.84$, $p < .001$; quadratic = 0.18 points/session, $SE = .06$, $p = .004$). The quadratic component of the growth model had a nearly zero variance, and so was fixed to zero in the final model. The dose of phone consultation received by the therapist had virtually no association with client improvement rates (estimate = 0.03 points/phone consultation; $SE = .13$, $p = .85$). The dose of live video consultation was associated with significantly greater improvement rates (estimate = −0.46 ECBI Intensity points per live video consultation received; $SE = .23$, $p = .04$). In simple terms, this would mean that the rate of client improvement accelerated about half an ECBI point per session each time the therapist received one additional video consultation. Follow-up consultation doses were sparse, and their effect did not approach significance (estimate = .24 ECBI Intensity points/follow-up consultation, $SE = .72$, $p = .74$). The variance of the postvideo follow-up phone consultation effect was large across therapists (variance = 0.64), especially compared to the variances of phone and video effects (0.01 and 0.05, respectively). After accounting for consultation dose, the remaining therapist level variance in client improvement (i.e., growth slopes) was not significant (slope variance = 0.14, $p = .37$). The planned contrast between phone and live video consultation effect estimates was significant (effect = 0.48 ECBI Intensity points per consultation better improvement for live video consultation; $SE = .24$, $p = .04$). To depict these effects visually, four model-based estimates were constructed for two common hypothetical implementation consultation strategies, each at two implementation time points—receipt of weekly phone versus biweekly live video consultation for a case seen (a) by a new therapist treating his or her first PCIT case and accruing consultation dose over the course of the case, and (b) for a case seen by a therapist

after completing 4 months of biweekly live video consultation or 6 months of weekly phone consultation. These four trajectories are depicted in Figure 3.

The model was replicated for ECBI Problem Scores. Recall that each ECBI item reflects a child behavior problem and parents provide a response indicating how often that behavior occurs (the Intensity Score), and then are asked if they consider that level of occurrence to be a problem for them as a parent (the Problem Score). Although Intensity items tended to be answered completely, Problem items were often left blank, even if an Intensity response was provided, and some parents never provided a Problem response, even when their Intensity responses were complete. This created a substantially greater amount of missing data for the ECBI Problem score analysis, which combined with the complexity of the analytic model, resulted in model convergence problems. The ECBI Intensity and Problem Scores were significantly correlated when parents provided both ($r = .75$, $p < .001$, in our sample), and the overall pattern and direction of findings and significance in the ECBI Intensity model was mirrored in the ECBI Problem model, but due to model instability and convergence problems using a maximum likelihood estimator, we do not feel confident reporting these findings.

The growth model was repeated a third time using session-by-session Treatment Attitude Inventory scores as the outcome. Responses on the TAI items range from 1 (*very dissatisfied*) to 5 (*very satisfied*), with a maximum possible score of 75 (observed among 14% of participants) and a theoretical minimum of 15 (observed among no participants; observed minimum TAI = 34). This outcome evidenced an irregular distribution, due to common all-5 or all-4 responses. The overall mean score was 61, suggesting that parents were "somewhat satisfied" overall. To manage distributional irregularity, TAI scores were collapsed into 4 bins—50 and under, 51–61, 62–74, and 75, which yielded a unimodal and symmetrical score distribution, then modeled. Satisfaction increased significantly over sessions (slope = .08, $SE = .02$, $p \leq .001$), but the impact of consultation sequence or the dose of any consultation condition did not approach significance. There was a very small nonsignificant trend toward client satisfaction improving less with cumulative live video consultation dose (effect = −0.01 units/video consultation, $SE = .01$, $p = .14$).

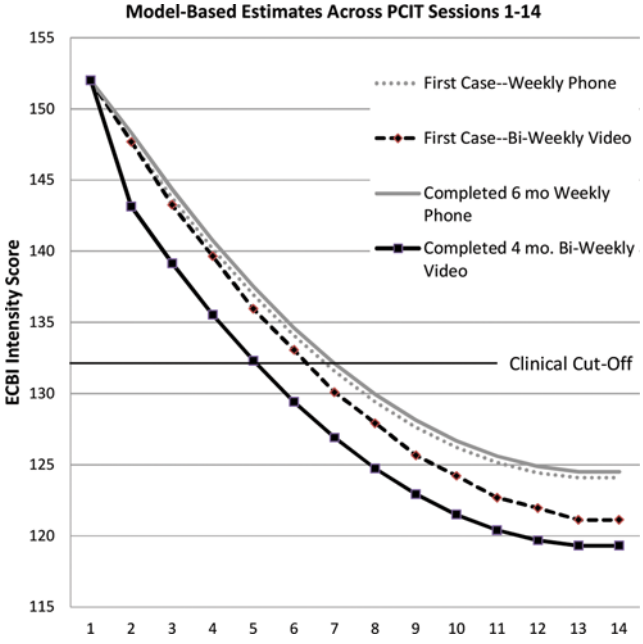


FIGURE 3 Model-based estimates across Parent–Child Interaction Therapy (PCIT) sessions for two common planned implementation strategies (weekly phone consultation vs. biweekly remote real-time consultation). *Note.* ECBI = Eyberg Child Behavior Inventory.

DISCUSSION

There are two main findings from this hybrid implementation-effectiveness study. The first concerns the overall transportability aim of the study and the field effectiveness of PCIT. Children entered treatment with an observed mean ECBI Intensity Score of 155 ($SD = 36$), well above the clinical cutoff of 132, and this dropped in the growth model by an average of 4.5 ECBI Intensity score points per session, with the decreases attenuating in a nonlinear fashion over sessions. Model based estimates at the benchmark session 14 were in a range of 120 to 125, depending upon the therapist's consultation type and dose (Figure 3), which represents about a 1 standard deviation decrease in behavior problems. Many cases discontinued therapy prior to the 14th session (*Mdn* number of sessions = 8), and about 20% continued beyond 14 sessions, but the 14 session point was selected for generating model-based estimates because it is a common fixed-dose benchmark used in laboratory PCIT studies. Compared to our modeled 152–120 pre–post ECBI changes, laboratory trials have reported decreases of 181 to 106 (McNeil et al., 1991), 175 to 126 (Hood & Eyberg, 2003), 169 to 112 (Eisenstadt et al., 1993), 160 to 127 (Schuhmann et al., 1998), and 150 to 134 (Thomas & Zimmer-Gemback, 2012). The 1 standard deviation improvement estimated in this study is squarely within the 0.61–1.45 standard deviation improvement range reported in the Thomas and Zimmer-Gemback (2007) meta-analysis. The main conclusion we would draw from this is that the benefit seen in PCIT laboratory efficacy trials transferred reasonably well into a scaled-up multiagency field implementation. This is encouraging news for scaling up PCIT.

The second main finding concerns the link between our implementation independent variable and client outcomes. There were small but significant client outcome benefits realized by using a live video type of posttraining consultation strategy. The size of this finding is consistent with that found for in vivo provider coaching in our scale-up of a very different type of home-based EBT in a very different service setting (Chaffin et al., 2012). Combined, these two findings suggest that the scale-up consultation strategy selected can have a small but potentially meaningful impact on downstream client outcomes. Although we do not wish to overstate the clinical significance of an additional 5-point ECBI Intensity score endpoint improvement for any individual case, when it comes to scaled-up implementations potentially serving large numbers of children, the population-level impact across hundreds of cases may be worth considering. Endpoint differences are not the sole benefit to consider. The findings suggest that live video consultation may create better outcomes earlier in the consultation process, generate improved outcomes with fewer consultation sessions, and produce client benefits earlier in therapy. For example, given the quadratic change trajectory model for PCIT, we would estimate that behavior disordered children beginning with an ECBI score of around 152 could reach the clinical cutoff on the ECBI about two sessions earlier with a fully video consulted therapist compared to a fully phone consulted therapist. This may be meaningful given that about half of all cases are not retained past eight sessions. It also is important to note that small effects should be anticipated in this type of study, where all clients received a robust EBT (PCIT); no clients received a control, placebo, or weak treatment; and the manipulated experimental variables involved therapist consultation procedures, which touch on clients only via presumed mediational pathways.

Choosing an implementation consultation strategy involves both benefit and cost considerations. We believe that the cost landscape for video consultation is rapidly changing. Live video consultation incurs more consultant costs—it requires one-on-one consultant time with each therapist. There is also fixed equipment cost and equipment complexity to consider, although we believe these concerns are declining as the necessary hardware is becoming less expensive, more portable, more user friendly, and more widely used in day-to-day life. Technical challenges were considerable in executing this project because a number of therapists were unfamiliar with handling and troubleshooting the videoconferencing technology, but we believe that technologies and familiarity have improved now, only a few years later. Every year, more tech-savvy generations enter the mental health services workforce. We are aware of at least two current efforts that have successfully employed simple tablet-based applications that are far easier to use, more portable, and less costly in order to accomplish essentially the same sort of video consultation we tested in this study (S. Self-Brown, personal communication, October 18, 2013; A. Urquiza, personal communication, October 18, 2013). On the plus side, video consultation does not incur the same opportunity costs as phone consultation. Every unit of completed live video consultation, by definition, is billable as a direct client treatment visit, whereas phone consultation requires taking time away from direct services. As a hypothetical example, let's assume that a unit of consultant time is valued at $200, and the consulted organization has five PCIT therapists and receives $100 for each completed service visit with an 80% attendance rate. The opportunity + direct cost to the organization for each scheduled phone consultation would be $(\$100 \times 5 \times 0.8) + \$200 = \$600$. The comparable cost of five scheduled video consultations to the organization would be $(\$200 \times 5) = \$1{,}000$. The per-therapist-per-week difference would be $(\$1{,}000 - \$600)/5 = \$80$. Assuming that each therapist has an ongoing PCIT caseload of four across which consultation costs are amortized, with each case seen an average of 10 weeks (or about eight sessions), this yields a per-case cost differential of $(\$80 \times 10)/4 = \$200$. This $200 per-case cost differential could then be weighed against the predicted incremental per-case client outcome benefit at various points in treatment (see Figure 3) and/or a reduction in theory of about two sessions of dose required to reach the clinical cutoff value defining "recovery." This two-session difference would be comparably valued at about $200 plus whatever patient costs might be involved in attending sessions). We appreciate that actual cost balance may differ from simple cost estimates such as these. But to the extent that the benefits of video supervision might be maintained after video supervision and its associated costs end, the cost-effectiveness balance could move from somewhat favorable in principle to quite favorable in actuality.

The findings for live video consultation should be interpreted in the context of the findings for standard phone consultation. Phone consultation is a routine implementation practice, and like most post hoc verbal consultation strategies is a venerable psychotherapy tradition. We were unable to document that phone consultation results in better client outcomes. It is still possible that phone consultation may have prevented drift and erosion of benefit over time, so our findings are not inconsistent with those studies reporting that some form of post hoc verbal consultation is superior to no consultation at all (e.g., Bradshaw, Butterworth, & Mairs, 2007). Still, this is not encouraging given that improvement of competency over time and not merely prevention of deterioration is the intended goal of phone

consultation. The effect of follow-up phone consultation, after the live video consultation period, appeared highly variable, which could have been related to data sparseness during this study phase.

Secondary findings concern client satisfaction with PCIT and observations about PCIT change trajectories. Although parents reported somewhat better behavior problem improvement with services delivered by a video consulted therapist, this did not translate into greater service satisfaction. This might raise a question of whether some parents experienced discomfort at being remotely observed (although no parent was observed without their voluntary authorization, and we were aware of no adverse events), but it does tend to mitigate any concern that greater improvement under video consultation is attributable to the placebo value of the technology or the cachet of having a remote expert direct treatment. Another secondary finding concerns PCIT change trajectories. This study captured session-by-session measures and because of the sample size was able to employ nonlinear growth models, in contrast to prior PCIT studies that have relied on simple pre–post measurement paradigms. This allowed us to investigate the overall shape of the change trajectory that might be expected in field implementation and document that it is curvilinear with most change occurring during early sessions.

Several study strengths and limitations need to be considered. Strengths include the number and diversity of study sites. Limitations include the fact that data were collected for clinical use by therapists, which may have impacted responses. The study did not follow therapists or cases for long after video consultation ended, and so we cannot judge whether its benefits were durable. The optimal dose and timing of video consultation also remains an open question. The study does not address the potentially important questions about the necessary length of external consultant involvement in an implementation, or when and how quality control functions might be handed off to local coaches or agency supervisors. We might speculate that live consultation is also advantageous when delivered by local agency supervisors, and that it might be easier to accomplish in this context. There is one feature we consider as both a strength and a limitation. The study was conducted in a range of real-world front line services settings. Consultation procedures were protocol driven, but there were no research staff at the clinical service sites, no limits placed on ordinary agency procedures, and no enforcement of client inclusion or exclusion criteria other than the clinician's judgment that PCIT was appropriate. Therapists entered and exited agency employ, took leave, had interruptions in job duties, and had varying PCIT caseloads in addition to their other work. We view this as a strength when it comes to the ecological validity of our findings. It may be a weakness when it comes to internal validity controls. This points to a need for replication studies examining consultation strategy findings, including attempting replication in smaller and more highly controlled contexts and employing more straightforward experimental designs. Roll-out type designs are well suited to the realities of implementation and scale-up (C. H. Brown, Wyman, Gao, & Pena, 2006), but they do introduce multiple design and analytic complexities. We also would point out that study findings were observed for a single EBT, and although we might expect that this would generalize to closely related EBTs, PCIT does have some unique characteristics.

The overall findings suggest some optimism about scaling up PCIT with reasonable transfer of efficacy trial benefits and document that more intensive consultation and competency development efforts can pay off in terms of small incremental client outcome benefits. Internet-based video technologies are one potentially promising resource that might be explored for accomplishing this. As the costs and technical challenges of live remote video strategies decrease, their appeal to EBT implementers may increase.

FUNDING

This research was supported by grant number U49CE001065 from the U.S. Centers on Disease Control and Prevention to Beverly Funderburk. We recognize the support of the CDC Foundation and the Doris Duke Charitable Foundation. We recognize the contributions of the following individuals to the project: Melanie Nelson, Jenelle Shanley, Sheila Eyberg, Briana Anderson, Naomi Perry, Jocelyn Savage, John McGuiness, Carol McCoy, Darden White, Amanda Pollock, Vicki Cook, Trish Gardner, Robin Gurwitch, Dee BigFoot, Alison Garrett, Eduardo Jones, Rebecca Glover, Linda Anne Valle, Beverly Fortson, Shannon Self-Brown, Jim Martindale, Christina Vasilev, Jacky Hoang, and the clinical staff and leadership of the 30 direct service agencies that graciously agreed to serve as study sites.

REFERENCES

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 4–23.

Bearman, S. K., Weisz, J. R., Chorpita, B. F., Hoagwood, K., Ward, A., Ugueto, A. M., & Bernstein, A. (2013). More practice, less preach? The role of supervision processes and therapist

characteristics in EBP implementation. *Administration and Policy in Mental Health and Mental Health Services Research*, *40*, 518–529.

Beidas, R. S., Edmunds, J. M., Marcus, S. C., & Kendall, P. C. (2012). Training and consultation to promote implementation of an empirically supported treatment: A randomized trial. *Psychiatric Services*, *63*, 660–665.

Boggs, S. R., Eyberg, S., & Reynolds, L. A. (1990). Concurrent validity of the Eyberg child behavior inventory. *Journal of Clinical Child Psychology*, *19*, 75–78.

Bradshaw, T., Butterworth, A., & Mairs, H. (2007). Does structured clinical supervision during psychosocial intervention education enhance outcome for mental health nurses and the service users they work with? *Journal of Psychiatric and Mental Health Nursing*, *14*, 4–12.

Brestan, E. V., Jacobs, J. R., Rayfield, A. D., & Eyberg, S. M. (1999). A consumer satisfaction measure for parent–child treatments and its relation to measures of child behavior change. *Behavior Therapy*, *30*, 17–30.

Brown, C. A., & Lilford, R. J. (2006). The stepped wedge trial design: A systematic review. *BMC Medical Research Methodology*, *6*, 54.

Brown, C. H., Wyman, P. A., Gao, J., & Pena, J. (2006). Dynamic wait-listed designs for randomized trials: new designs for prevention of youth suicide. *Clinical Trials*, *3*, 259–271.

Chaffin, M., Funderburk, B., Bard, D., Valle, L. A., & Gurwitch, R. (2011). A combined motivation and parent–child interaction therapy package reduces child welfare recidivism in a randomized dismantling field trial. *Journal of Consulting and Clinical Psychology*, *79*, 84–95.

Chaffin, M., Hecht, D., Bard, D., Silovsky, J. F., & Beasley, W. H. (2012). A statewide trial of the SafeCare home-based services model with parents in Child Protective Services. *Pediatrics*, *129*, 509–515.

Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., . . . Bonner, B. L. (2004). Parent–child interaction therapy with physically abusive parents: efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology*, *72*, 500–510.

Colvin, A., Eyberg, S., & Adams, C. (1999). *Restandardization of the Eyberg Child Behavior Inventory*. Available from http://www.pcit.org

Curran, G. M., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: Combining elements of clinical effectivenss and implementation research to enhance public health impact. *Medical Care*, *50*, 217–226.

Edmunds, J. M., Beidas, R. S., & Kendall, P. C. (2013). Dissemination and implementation of evidence-based practices: Training and consultation as implementation strategies. *Clinical Psychology: Science and Practice*, *20*, 152–165.

Eisenstadt, T. H., Eyberg, S., McNeil, C. B., Newcomb, K., & Funderburk, B. (1993). Parent–child interaction therapy with behavior problem: Relative effectiveness of two stages and overall treatment outcome. *Journal of Clinical Child Psychology*, *22*, 42–51.

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

Elliott, D., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science*, *5*, 47–53.

Eyberg, S. M. (1974). *Therapy attitude inventory*. Available from Sheila Eyberg, Department of Clinical Psychology, Box J-165, JHMHC, University of Florida, Gainesville, FL 32610.

Eyberg, S. M., Boggs, S. R., & Algina, J. (1995). Parent–child interaction therapy: A psychosocial model for the treatment of young children with conduct problem behavior and their families. *Psychopharmacology Bulletin*, *31*, 83–91.

Eyberg, S. M., & Matarazzo, R. G. (1980). Training parents as therapists: A comparison between individual parent–child interaction training and parent group didactic training. *Journal of Clinical Psychology*, *36*, 492–499.

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

Eyberg, S. M., & Robinson, E. A. (1982). Parent–child interaction training: Effects on family functioning. *Journal of Clinical Child & Adolescent Psychology*, *11*, 130–137.

Filene, J. (2012, May). *Meta-analytic review of components associated with home visiting programs* [Final report]. Washington, DC: Pew Charitable Trusts.

Fixsen, D., Naoom, S. F., Blase, D. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication# 231). Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute.

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*, 67–82.

Glasgow, R. E. (2013). What does it mean to be pragmatic? Pragmatic methods, measures, and models to facilitate research translation. *Health, Education, and Behavior*, *40*, 257–265.

Herschell, A. D., Kolko, D. J., Baumann, B. L., & Davis, A. C. (2010). The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. *Clinical Psychology Review*, *30*, 448–466.

Hood, K. K., & Eyberg, S. M. (2003). Outcomes of parent–child interaction therapy: Mothers' reports of maintenance three to six years after treatment. *Journal of Clinical Child and Adolescent Psychology*, *32*, 419–429.

Institute of Medicine. (2001, March). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: The National Academies.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: ASCD.

Nadeem, E., Gleacher, A., & Beidas, R. S. (2013). Consultation as an implementation strategy for evidence-based practices across multiple contexts: Unpacking the black box. *Administration and Policy in Mental Health and Mental Health Services Research*, *40*, 439–450.

McNeil, C. B., Capage, L. C., Bahl, A., & Blanc, H. (1999). Importance of early intervention for disruptive behavior problems: Comparison of treatment and waitlist-control groups. *Early Education and Development*, *10*, 445–454.

McNeil, C. B., Eyberg, S. M., Eisenstadt, T. H., Newcomb, K., & Funderburk, B. (1991). Parent–child interaction therapy with behavior problem children: Generalization of treatment effects to the school setting. *Journal of Clinical Child and Adolescent Psychology*, *20*, 140–151.

PCIT International. (2013, March 2). Training requirements for certification as a PCIT therapist. Retrieved from http://www.pcit.org/training-guidelines/certified-pcit-therapist-training-requirements-5/

Schoenwald, S. K., Letourneau, E. J., & Halliday-Boykins, C. A. (2005). Predicting therapist adherence to a transported family- based treatment for youth. *Journal of Clinical Child & Adolescent Psychology*, *34*, 658–670.

Schuhmann, E. M., Foote, R. C., Eyberg, S. M., Boggs, S. R., & Algina, J. (1998). Efficacy of parent–child interaction therapy: Interim report of a randomized trial with short-term maintenance. *Journal of Clinical Child Psychology*, *27*, 34–45.

Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of parent–child interaction therapy and Triple P—Positive Parenting Program: A review and meta-analysis. *Journal of Abnormal Child Psychology*, *35*, 475–495.

Thomas, R., & Zimmer-Gembeck, M. J. (2012). Parent–child interaction therapy: An evidence-based treatment for child maltreatment. *Child Maltreatment*, *17*, 253–266.

APPENDIX

Sample MPlus 7.1 Syntax

```
VARIABLE: NAMES ARE
Client_id therapist_id ! client and therapist level ID variables
Y1-Y14; ! ECBI Intensity Scores for sessions 1-14
NP1-NP13 ! cumulative dose of phone consultation at sessions 1-13
NR1-NR13 ! cumulative dose of remote video consultation at sessions 1-13
NF1-NF13 ! cumulative dose of follow-up phone consultation at sessions 1-13

CLUSTER IS THERID;
WITHIN ARE NP1-NP13 NR1-NR13 NF1-NF13;
! setting multi-level structure
ANALYSIS:
TYPE IS TWOLEVEL RANDOM;
ESTIMATOR IS MLR;
! three-level growth model (growth + two-level) with
! additional random effects, using robust maximum likelihood
MODEL:
%WITHIN%
I S Q | Y1@0 Y2@1 Y3@2 Y4@3 Y5@4
Y6@5 Y7@6 Y8@7 Y9@8 Y10@9
Y11@10 Y12@11 Y13@12 Y14@13;
! client level growth model with random intercept (I) linear (S) and quadratic (Q) slopes

Y1-Y14 (resid);
! equality constraint on residuals

I WITH S; I WITH Q; Q WITH S;
! growth parameter covariances
! plus three time-varying random dose effects:
SP | Y2 ON NP1;
!thru
SP | Y14 ON NP13;
! random slope for lagged phone dose
SR | Y2 ON NR1;
thru
SR | Y14 ON NR13;
! random slope for lagged remote video dose
SF | Y2 ON NF1;
thru
SF | Y14 ON NF13;
! random slope for lagged follow-up dose
%BETWEEN%
IB SB QB | Y1@0 Y2@1 Y3@2 Y4@3 Y5@4
Y6@5 Y7@6 Y8@7 Y9@8 Y10@9
Y11@10 Y12@11 Y13@12 Y14@13;
! therapist level growth model with intercept, linear and quadratic slopes
Y1-Y14@0;
! between level variability in client residuals
! customarily constrained to zero in growth models

IB WITH SB; IB WITH QB; SB WITH QB;
!therapist level growth parameter covariances

[SR] (MEANR);
[SP](MEANP)
[SF] (MEANF);
! naming therapist level point estimates for dose-effect coefficients
MODEL CONSTRAINT:
NEW(DIFFPR DIFFPF DIFFRF);
DIFFPR = MEANP - MEANR;
DIFFPF = MEANP - MEANF;
DIFFRF = MEANR - MEANF;
! building therapist level dose-effect point estimate contrasts
```