

Florida State University Libraries

2014

Does Parent-Child Interaction Therapy Reduce Future Physical Abuse?: A Meta-Analysis

S. C. Kennedy, J. S. Kim, S. J. Tripodi, S. M. Brown and G. Gawdy



Running head: DOES PCIT REDUCE FUTURE ABUSE

Abstract

Objective: To use meta-analytic techniques to evaluating the effectiveness of Parent-Child Interaction Therapy (PCIT) at reducing future physical abuse among physically abusive families. **Methods:** A systematic search identified six eligible studies. Outcomes of interest were physical abuse recurrence, child abuse potential, and parenting stress. **Results:** Parents receiving PCIT had significantly fewer physical abuse recurrences and significantly greater reductions on the Parenting Stress Index than parents in comparison groups. Reductions in child abuse potential were non-significant, although 95% confidence intervals suggest clinically meaningful treatment effects. The studies examining physical abuse recurrence had a medium treatment effect ($g=0.52$), while results from pooled effect size estimates for child abuse potential ($g=0.31$) and parenting stress ($g=0.35$) were small. **Conclusions:** PCIT appears to be effective at reducing physical abuse recurrence and parenting stress for physically abusive families, with the largest treatment effects seen on long-term physical abuse recurrence. Applications to social work practice are discussed.

Keywords: Parent-child interaction therapy, child physical abuse, child welfare, meta-analysis

Introduction

Although the overall rate is declining, child maltreatment remains an enduring social problem. Federal data indicate that more than 650,000 children were the victims of child abuse and neglect in 2012 (U.S. Department of Health and Human Services [U.S. DHHS], 2013). According to the Fourth National Incidence Study of Child Abuse and Neglect, 60% of abuse cases involved physical abuse (Sedlak et al., 2010). Victims of childhood physical abuse (CPA) are at higher risk for experiencing a range of difficulties across the life span, including externalizing (e.g., acting-out) behavior problems (Herschell & McNeil, 2005), anxiety (Freeman & Fowler, 2009; Turner, Finkelhor, & Ormrod, 2006), depression (Springer, Sheridan, Kuo, & Carnes, 2007), and suicide attempts (Bruffaerts et al., 2010; Mandelli et al., 2011). Additionally, the U.S. DHHS (2013) estimates that over 80% of CPA is perpetrated by a biological parent. Therefore, the behavioral and mental health consequences of CPA generally occur within the context of the family system (Chaffin et al., 2004). These behaviors are theorized to contribute to parenting stress and thus, to subsequently increase the risk for the incidence of future physical abuse (Chaffin et al., 2004; Herschell & McNeil, 2005; Urquiza & McNeil, 1996). The coercive cycle model (Chaffin et al., 2004; Urquiza & McNeil, 1996) posits that harsh discipline is rewarded by the temporary compliance of the child and the child learns to ignore parental directives until the point of aggression. This process creates a hostile parent-child relationship in which force and coercion are both the cause and effect of the child's behavior (Chaffin et al., 2004).

A wealth of interventions are designed to halt this coercive cycle as a means to improving family dynamics. Parent-Child Interaction Therapy (PCIT) is a dyadic parenting intervention, in which the family system is altered through modifying the behavior of both the parent and child.

Although PCIT was designed to reduce severe behavior problems in very young children, the intervention has been increasingly applied to physically abusive or at-risk families in the years since Urquiza and McNeil (1996) first published a theoretical piece on the topic. Additionally, in 2004, PCIT was endorsed as a best-practice by the Kauffman Best Practices Project because it demonstrated a sound theoretical framework, a substantial research base, and could be easily translated to community practice. Since earning this endorsement, the use of PCIT with child-welfare involved families has increased dramatically. However, even when implemented with physically abusive or at-risk families, the majority of studies assess changes in child-level outcomes like externalizing behavior. Parent-level outcomes such as subsequent incidence of abuse, child abuse potential, and parenting stress are not always included in research design or synthesized in narrative reviews on the efficacy of the program with this population.

In a systematic review on the effectiveness of PCIT with physically abusive or at-risk families, Hershell and McNeil (2005) identified only one study (Chaffin et al., 2004) which reported physical abuse recurrence in the follow-up period for participants and controls. A second review comparing outcomes of PCIT and Triple P – Positive Parenting Program (an alternative social learning based parenting intervention; Sanders, Cann, & Markie-Dadds, 2003) for physically abusive families was published in 2007, but only reviews studies published through 2004 (Thomas & Zimmer-Gembeck, 2007). Further, the ability of either program to reduce abuse potential or future physical abuse were not examined as outcomes. Since the publication of these reviews, the methodological quality of studies evaluating PCIT with physically abusive or at-risk families has improved dramatically. Therefore, a comprehensive meta-analysis of available experimental and quasi-experimental studies is warranted.

The purpose of this meta-analysis is to examine the effectiveness of PCIT at reducing

future physical abuse for physically abusive or at-risk families. We operationally defined future physical abuse on three parent-level domains: physical abuse recurrence, child abuse potential, and parenting stress. Parenting stress was chosen as a salient proxy for the abuse construct as parenting stress is hypothesized to contribute to abuse potential for physically abusive or at-risk families (Urquiza & McNeil, 1996). Therefore, this meta-analysis reports only on outcome variables that represent changes in parenting behaviors, attitudes, or stress levels. An analysis of the Dyadic Parent–Child Interaction Coding System-II (DPICS–II; Eyberg, Bessmer, Newcomb, Edwards, & Robinson, 1994), a behavioral outcome measure which corresponds to parent behaviors coached in PCIT, was not possible due to heterogeneity of conceptualized coding procedures. Between-group differences on the DPICS-II were reported in a way that allowed combination for only two of the four studies which used the measure. Analysis of child-level outcomes (e.g., measures of externalizing behaviors) included in the original studies are not presented here.

Specifically this study seeks to answer the following research questions: 1) Do abusive parents who receive PCIT significantly reduce physical abuse recurrence when compared to abusive parents in treatment-as-usual (TAU) comparison or no-treatment control groups?; 2) Do abusive parents who receive PCIT significantly reduce scores on measures of child abuse potential when compared to abusive parents in TAU or no-treatment control groups?; 3) Do abusive parents who receive PCIT significantly reduce scores on measures of parenting stress, when compared to abusive parents in TAU or no-treatment control groups?

Description of the Intervention

PCIT is a 14- to 20-week manualized dyadic parent-training intervention developed for children with externalizing behavior problems (Eyberg & Calzada, 1998; Eyberg & Robinson,

1983; Hembree-Kigin & McNeil, 1995). The intervention is grounded in social learning and attachment theories, employs a family systems approach, and targets children aged two to seven (Bandura, 1977; Bowlby, 1969). PCIT is unique in that the intervention goes beyond didactic psychoeducation, includes both parent and child in sessions (i.e., parent-child dyads are the typical unit of measurement), and uses in vivo coaching via a one-way mirror and a “bug in the ear” device to train parents to positively engage their child in play and respond appropriately to their child’s behavior (Eyberg & Calzada, 1998). The hypothesized mechanism of change is providing parents with behavior modification strategies so that they become agents of change in reducing their child’s externalizing behavior, which in turn promotes more positive parenting (Chaffin et al., 2004; Thomas & Zimmer-Gembeck, 2011, 2012).

PCIT consists of two phases. The goal of the first phase is to enhance the parent-child relationship (described as Child-Directed Interaction or CDI), and the second is to improve child compliance (described as Parent-Directed Interaction or PDI). Each phase consists of 7-10 sessions, the first of which includes at least one hour of didactic training. During the CDI phase, parents are trained to follow their child’s lead in play, ignore minor child misbehavior, to avoid criticism, sarcasm, or other negative verbalizations, and to increase PRIDE skills: Praise, Reflection, Imitation, Description, and Enthusiasm (Eyberg & Calzada, 1998). During the PDI phase, parents are coached in giving appropriate commands and using time-out strategies to obtain child compliance. Parent-child dyads do not move from the CDI to the PDI phase or from the PDI phase to completion, until they have mastered the skills of each module (Eyberg & Calzada, 1998). Therefore, treatment length is variable.

A hands-off version of PCIT is commonly used when implementing the program with a CPA population (Chaffin et al., 2004; Timmer, Urquiza, Zebell, & McGrath, 2005). The hands-

off model eliminates all physical holding and bottom-swats as a back-up to time-out, instead substituting a time-out room or some other non-physical second layer of discipline (Chaffin, Funderburk, Bard, Valle, & Gurwitsch, 2011; Chaffin et al., 2004; Foley, 2010; Galanter et al., 2012). Additionally, the hands-off model of PCIT works with children up to age 12, as the focus of change is the parent's behavior, rather than the child's behavior (Chaffin et al., 2004).

Conducting research on PCIT with child-welfare involved families raises ethical concerns about withholding treatment, therefore all of the identified studies compared PCIT to a TAU condition or a time-limited wait-list control condition. TAU represented a variety of didactic community-based parent-education delivered weekly to parents in a group format without any children present. TAU sessions were typically offered to mirror PCIT sessions in frequency and duration (Chaffin et al., 2011; Chaffin et al., 2004; Foley, 2010). In TAU sessions, parents learned about child development and developmentally appropriate expectations, principles of discipline, use of praise, communication strategies, stress management, and the ways in which parental personal problems affect children.

Likewise, implementation of PCIT, in most cases, was modified in some way to maximize participant exposure to the CDI and PDI phases even though participants were generally moved from one phase to another based on a timeline, rather than mastery of the skills. In one study (Foley, 2010), a group format was employed and all participants moved through the program together. Due to the necessary adjustments to PCIT identified above, the group format was not viewed as conceptually disparate from the other studies which met eligibility criteria. As stated previously, PCIT with child-welfare involved families identifies the parent as the primary target of change (Chaffin et al., 2011; Chaffin et al., 2004; Urquiza & McNeil, 1996). Therefore, this meta-analysis only synthesizes parent-level outcomes that represent changes in parenting

behaviors, attitudes, or stress levels. Specifically, the outcomes of interest are physical abuse recurrence, child abuse potential, and parenting stress.

Method

Study selection

To identify studies, two members of the research team performed independent searches of existing literature for all studies evaluating the effectiveness of PCIT with physically abusive or at-risk families. QUORUM reporting standards were used to guide the systematic identification of studies and subsequent analysis (Moher et al., 1999). The following electronic databases were searched in September 2013: Social Sciences Index and Abstracts (ASSIA), Campbell Collaboration Systematic reviews, Cochrane Central Register of Controlled Trials (CENTRAL), ISI Web of Knowledge, PILOTS, ProQuest Criminal Justice, PsycINFO, PubMed, Social Services Abstracts, and Sociological Abstracts. Potentially relevant studies were also searched for in the ProQuest Dissertations and Theses database, as the application of the PCIT intervention to this population is relatively recent. Studies were located using the following terms: Parent-Child Interaction Therapy or PCIT in the title and abus* or maltreat* or risk in the abstract. No date restrictions were placed on the search, although eligible studies were written in English language. Reference lists of all relevant articles, systematic reviews, meta-analyses, and dissertations/theses were checked to identify additional studies for inclusion. Studies from the grey literature (e.g., book chapters, conference papers, government reports, white papers) were included only if they were identified using the above search strategy. Two experts in the field were contacted as a means to identify potential in press or unpublished studies. The goal of this search strategy was to increase sample size and statistical power, and to minimize publication bias (Lipsey &

Wilson, 2001). Published studies are more likely to report statistically significant results than dissertations/theses, white papers, or unpublished studies (Begg, 1994).

Review Criteria and Categorization

Several inclusion criteria were established for this meta-analysis on PCIT with physically abusive or at-risk families. First, eligible studies employed experimental or quasi-experimental research design. Case studies, one-group pretest-posttest designs, and one-group retrospective chart reviews were excluded. Second, only studies that reported physical abuse recurrence after treatment completion, or used a standardized proxy measurement of abuse potential or parenting stress were eligible for inclusion. Prior to conducting the search, particular measures of child abuse potential (e.g., Child Abuse Potential Inventory) and parenting stress (e.g., Parenting Stress Index) were not specified. Rather, any instrument that measured abuse or stress met inclusion criteria.

Applying these criteria, the search yielded an initial pool of 51 papers. See Figure 1 for a QUORUM flowchart. Titles were subsequently scanned and were excluded from review if the title indicated that the paper was a case study or focused exclusively on a sample of foster children and their foster or adoptive parents. Of these 51 titles, 31 titles were consistent with the search criteria. Abstracts determined relevance to the research question and 14 articles (eight peer-reviewed articles and six dissertations/theses) were read in full. Reasons for excluding studies at this stage were that the study created groups retrospectively from a pool of PCIT treatment completers (e.g., Lambdin-Shirley, 2008; Timmer, Ware, Urquiza, & Zebell, 2010), compared PCIT treatment completers to non-completers (e.g., Galanter et al., 2012), or PCIT was implemented among parents who did not have access to their children (e.g., women prisoners; Tempel, 2012). After a full-text review, six studies were retained:

four peer-reviewed articles (Chaffin et al., 2011; Chaffin et al., 2004; Thomas & Zimmer-Gembeck, 2011, 2012) and two dissertations (Foley, 2010; Terao, 1999). The research team developed a coding sheet before data extraction based on Lipsey and Wilson's (2001) recommendations. Two independent coders coded the included articles and then compared results. Any discrepancies or questions that emerged were reviewed and discussed by the entire research team and resolved.

Outcomes of interest

The current meta-analysis only considers parent-level outcomes that represent changes in parenting behaviors, attitudes, or stress levels. An analysis of child-level outcomes (e.g., measures of externalizing behaviors) reported in the original studies is not presented here.

Physical abuse recurrence.

Physical abuse recurrence was reported by two studies (Chaffin et al., 2011; Chaffin et al., 2004) and was obtained from the statewide child welfare administrative database. Surveillance effects were accounted for and any reports ruled out or screened out by child welfare workers were excluded. Both studies used intent-to-treat data and cases were followed for a mean of 904 and 850 days, respectively.

Child abuse potential.

The Child Abuse Potential Inventory (CAPI; Milner, 1986) is a 160-item dichotomous agree-disagree format instrument developed to estimate risk for committing child physical abuse. Subscales include parental distress, rigidity, unhappiness, problems with child and self, problems with family, and problems with others. The global CAPI has reported high internal consistency (0.92 to 0.96 for controls and 0.95 to 0.98 for abusers) and strong future predictive

validity (Chaffin & Valle, 2003; Milner, 1986, 1994). Milner (1986) suggests a cut score of 166 and a normative mean of 91.

Parenting stress.

Both the Parenting Stress Inventory (PSI; Abidin, 1990) and the Parenting Stress Index – Short Form (PSI-SF; Abidin, 1995) were used in the current analysis. The PSI consists of 101 items: 90 items forming composite scores for child and parent stress domains rated on a 6 point Likert scale, and 11 multiple choice items assessing parental life stressors. High scores on the child stress domain indicate that the specific qualities of their child make parenting more difficult. High scores on the parent stress domain indicate that parent functioning creates difficulties in the parenting relationship. The PSI-SF retains the qualities of the PSI using 36 items.

Statistical Procedures

Statistical analysis produced descriptive information on the characteristics of the included studies and effect sizes of each intervention for conceptually distinct outcomes of interest to this review. To avoid combining heterogeneous outcomes (i.e., physical abuse recurrence, child abuse potential, parenting stress), overall effect size estimates were calculated for each of the three outcome constructs.

For this meta-analysis, effect sizes and 95% confidence intervals were calculated using Comprehensive Meta-Analysis software 2.0 (Borenstein, Hedges, Higgins, & Rothstein, 2005). Hedges's *g* effect size statistic was calculated for all studies that reported mean scores and corresponding standard deviations. The effect size statistic standardizes reported study outcomes to allow interpretation, comparison, and analysis of results across studies (Lipsey & Wilson, 2001). Effect sizes go beyond the typical statistical significance normally reported with *p*-values

by estimating the magnitude and direction of the treatment effect (Littell, Corcoran, & Pillai, 2008).

Physical abuse recurrence was presented as a dichotomous outcome (presence/absence of re-report to a child welfare database), so an odds ratio effect size was calculated and then converted to Hedges's *g* effect size to help maintain consistency in interpreting the magnitude of treatment effect. Chaffin and colleagues' (2004) RCT contained two treatment groups – PCIT and Enhanced-PCIT (EPCIT). Results from PCIT and EPCIT groups were evaluated against a TAU comparison group. Because the other studies included in this meta-analysis used traditional PCIT, a decision was made prior to analysis that the current project would only present effect sizes for participants in the PCIT group, and would exclude results from EPCIT participants. Likewise, Chaffin et al. (2011) conducted a randomized disabling trial to parse out the effects of a six-session orientation used in both treatment and comparison conditions during the 2004 efficacy study. Chaffin et al. (2011) evaluated four discrete groups, only two of which corresponded to interventions delivered by Chaffin et al. (2004): PCIT plus a motivational orientation ($n=34$) and TAU plus a standard orientation ($n=42$). The corresponding author of both studies confirmed that the participants included in the current meta-analysis received equivalent interventions across studies and treatment conditions.

To ensure statistical independence of effect sizes, only one effect size was calculated per study per outcome. In cases where multiple measures were used to assess the same outcome construct in a study, a study-level average was calculated. Hedges's *g* effect sizes are reported such that positive estimates indicate results favoring the intervention group and negative effect sizes indicate results favoring the comparison or control group.

Both the fixed-effects and the random-effects model guides pooling of individual effect

size estimates in meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, 2010). Both models use weights based on sample size to account for variance of effect sizes across studies (Borenstein et al., 2010). The fixed-effects model assumes that variability in study effect sizes is no greater than would be expected due to sampling error (Lipsey & Wilson, 2001). Under the fixed-effects model, excess variability beyond sampling error is assumed to be non-random, and is explained by adding predictor variables to the model or calculating a mean. Additionally, the fixed-effects model posits the existence of a common effect size that is the same in all the studies analyzed, thus, the pooled effect size is an estimate of this true effect size (Borenstein et al., 2010). In a random-effects model, on the other hand, variability in effect sizes are assumed to be greater than would be expected due to sampling error (Lipsey & Wilson, 2001). The source of between-study variability is assumed to be randomly distributed and therefore, cannot be systematically explained using moderating (i.e., study descriptor) variables (Lipsey & Wilson, 2001). In the random-effects model, the effect sizes reported in each individual study are not assumed to estimate a common effect size, and the pooled effect size is interpreted as an estimate of the mean of treatment effects across studies (Borenstein et al., 2010). In this model, new weights and confidence levels are calculated to account for treatment effect variance across studies (Shadish & Haddock, 1994).

Although the Q statistic typically determines whether to use a fixed or random-effects model to interpret the data, this test lacks statistical power with small sample sizes (Lipsey & Wilson, 2001). Therefore, we used the I^2 index to quantify the effect of heterogeneity among the studies and provide a measure of the degree of inconsistency in each study's results (Higgins, Thompson, Deeks, & Altman, 2003). I^2 describes the percentage of variation across studies due to heterogeneity rather than chance. The I^2 value lies between 0% and 100%, where 0% indicates

no heterogeneity, and values increase as heterogeneity increases. High heterogeneity is an indication that variability is not due to sampling error alone (Higgins et al., 2003). For the current project, we used the classification where $I^2 = 25$ is small, $I^2 = 50$ is medium, and $I^2 = 75$ is large (Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006). Due to the small number of studies that measured the same outcomes within each time period, the authors were not able to conduct meaningful moderator analyses to assist with understanding what factors contributed to the heterogeneity.

Results

Search Results

PCIT treatment effect sizes were derived from all six studies and included a total of 571 parent-child dyads. This number reflects the exclusion of 109 dyads who received enhanced or otherwise non-comparable versions of PCIT or TAU in Chaffin et al.'s (2004) RCT (n=33) and Chaffin and colleague's (2011) RCT (n=76). See Table 1 for an overview of client, program, and study characteristics for included studies. The majority of parent participants were female (85.3%) with a mean age of 32.1 years. Children were majority male (66.8%) with a mean age of 5.3 years. Parents identified primarily as Caucasian (71.9%) and African-American (21.1%). Parents who identified as Bi-racial, Latino/a, Asian, Native-American, and Australian aboriginal comprised a minority of participants (< 7%). Several studies (Terao, 1999; Thomas & Zimmer-Gembeck, 2011, 2012) did not report complete demographic information on participants, and both Chaffin et al., (2011) and Chaffin et al. (2004) reported only parent demographics. When reported, the majority of participants were referred to PCIT through child protective, legal services, or social services (60%), with a substantial minority coming from other sources or self-referral (40%). Of the six studies, five used randomization to create groups. Chaffin et al. (2011),

Chaffin et al. (2004), and Terao (1999) randomized participants into an experimental group or a TAU comparison group; Thomas & Zimmer-Gembeck (2011, 2012) randomized participants into an experimental group or a no-treatment, wait-list control group. The sixth study (Foley, 2010) assigned participants to an experimental group or a TAU comparison group based on space availability at time of intake. All studies had quality control procedures in which PCIT was implemented using a treatment manual and fidelity checks. Across all six studies, PCIT was provided on average for 19.1 weeks and participants completed a mean of 16 total sessions. Five out of the six studies reported implementation of PCIT one time per week. All personnel delivering PCIT had specialized training: 40% were identified as PCIT experts, 40% clinical psychologists, and 20% social workers.

Effect Size Results

Physical abuse recurrence.

Two studies (Chaffin et al., 2011; Chaffin et al., 2004) evaluated long-term physical abuse recurrence within the parent-child dyad. Results from the Chaffin et al. (2004) study showed a large and statistically significant effect size estimate ($g = 0.77$, $p < 0.01$), with a 95% confidence interval that ranged from 0.22 to 1.32. Chaffin et al., 2011 reported a small and non-significant effect size estimate ($g = 0.29$, $p = 0.28$), with a 95% confidence interval that ranged from -0.24 to 0.82.

The I^2 index for physical abuse recurrence was 33.91 which indicated small heterogeneity and that a fixed-effects model was appropriate. The pooled effect size was interpreted as an estimate of the mean of treatment effects across studies. Results indicated a fixed-effects weighted mean effect size estimate of 0.52, suggesting that parents in the PCIT groups scored 0.52 standard deviations above those in the TAU comparison group. According to Cohen (1988),

this effect size is considered medium. The magnitude of the overall weighted mean effect size is positive and statistically significant ($p < 0.01$), with a 95% confidence interval of 0.14 to 0.90.

Child abuse potential.

Five studies (Chaffin et al., 2004; Foley, 2010; Terao, 1999; Thomas & Zimmer-Gembeck, 2011, 2012) examined child abuse potential using the CAPI. Standardized mean differences effect sizes (Hedges's g) were used for these continuous type outcome measures, with an effect size of zero indicating no effect. As Table 2 shows, effect size estimates across all studies ranged from 0.07 to 0.96. The I^2 index for child abuse potential was 57.30 which indicated medium heterogeneity and that a random-effects model would be more appropriate (Huedo-Medina et al., 2006). That is, effect size variability between studies was larger than would be expected given only sampling error. Therefore, individual effect size estimates from the primary studies represented a random sample of all possible effect sizes that could have been observed (Borenstein et al., 2010). The pooled effect size was interpreted as an estimate of the mean of treatment effects across studies. Results indicated a random-effects weighted mean effect size estimate of 0.31, which suggests that parents in the PCIT groups scored 0.31 standard deviations above those in the comparison or control groups. According to Cohen (1988), this effect size is considered small. The magnitude of the overall weighted mean effect size is positive with a 95% confidence interval of -0.002 to 0.620.

Although this effect size was not significant with an alpha of .05 ($p = 0.053$), the p -value suggests that the observed between group differences would only have occurred by chance in 5.3% of samples. Examination of the 95% confidence revealed encouraging results, as these data suggested with 95% confidence that the true CAPI score for parents in the PCIT group was between 62% lower and 0.2% higher than parents in comparison and control groups. While these

data reveal that pre-test to post-test scores on the CAPI could actually increase, the majority of the confidence interval falls on the side of PCIT reducing child abuse potential when compared to comparison and control conditions.

Parenting stress.

Four studies (Foley, 2010; Terao, 1999; Thomas & Zimmer-Gembeck, 2011, 2012) examined parenting stress outcomes using the PSI or PSI-SF. As Table 2 shows, Hedges's g effect size estimates across all studies ranged from 0.15 to 0.97. The I^2 index for parenting stress was 45.39 which indicated medium heterogeneity and that a random-effects model would be more appropriate (Huedo-Medina et al., 2006). That is, effect size variability between studies was larger than would be expected given only sampling error. Therefore, individual effect size estimates from the primary studies represented a random sample of all possible effect sizes that could have been observed (Borenstein et al., 2010). The pooled effect size was interpreted as an estimate of the mean of treatment effects across studies. Results indicated a random-effects weighted mean effect size estimate of 0.35, which suggests that parents in the PCIT groups scored 0.35 standard deviations above those in the comparison or control groups for parenting stress. According to Cohen (1988), this effect size is considered small. The magnitude of the overall weighted mean effect size is positive and statistically significant ($p=0.03$), with a 95% confidence interval of 0.04 to 0.65.

Discussion and application to social work

Since Urquiza and McNeil (1996) published a theoretical piece detailing the potential benefits of using PCIT with physically abusive or at-risk families, outcome research on the topic has become increasingly sophisticated and methodologically rigorous. The intervention's hypothesized mechanism of change follows social learning theory explanations of child

misbehavior and harsh or violent discipline strategies (Patterson, 1976). Changes occur by altering parent-child interaction patterns and de-escalating coercive interactions (Urquiza & McNeil, 1996). The parent then disrupts the coercive cycle – which posits that parents' harsh discipline is rewarded by temporary child compliance, but ultimately teaches the child to ignore parental directives until the point of aggression (Chaffin et al., 2004; Urquiza & McNeil, 1996). This process creates a hostile parent-child relationship in which force and coercion are both the cause and effect of the child's behavior (Chaffin et al., 2004). PCIT is designed to teach appropriate parent-child interaction and discipline, reduce the child's externalizing behavior, and promote more positive parenting (Eyberg & Calzada, 1998; Urquiza & McNeill, 1996). After receiving endorsement as best-practice by the Kauffman Best Practices Project (2004), PCIT has been widely implemented with children in child welfare contexts. As social workers are key agents in the child welfare system, the effectiveness of PCIT to reduce future physical abuse among physically abusive or at-risk families is especially salient to the field.

This meta-analysis evaluated the effectiveness of PCIT to reduce future physical abuse when implemented among physically abusive parent-child dyads. Future physical abuse was evaluated on three parent-level domains: physical abuse recurrence, child abuse potential, and parenting stress. Six studies (five RCTs and one quasi-experiment) met inclusion criteria (Chaffin et al., 2011; Chaffin et al., 2004; Foley, 2010; Terao, 1999; Thomas & Zimmer-Gembeck, 2011, 2012) and effect sizes were calculated using data from 571 participant dyads. The six studies reviewed here tentatively support the hypothesis that PCIT is an effective intervention for reducing future physical abuse in physically abusive parent-child dyads.

The current analysis suggests that physically abusive parents who participate in PCIT

interventions are significantly less likely to have physical abuse recurrence when compared to physically abusive parents who participate in didactic parenting classes (i.e., TAU). Effect sizes from the two studies which considered physical abuse recurrence as an outcome had moderate treatment effects ($g=0.52$), which suggests that participation in the PCIT intervention contributed to meaningful reductions in future physical abuse. This finding is strengthened by the RCT design and a median follow-up time in excess of 850 days for both studies.

Although our meta-analysis found that physically abusive parents who participate in PCIT interventions have significant pre-test to post-test decreases on measures of parenting stress (i.e., PSI) when compared to parents in TAU or no-treatment, wait-list control conditions, the effect sizes estimate was small ($g=0.35$). Between group pre-test to post-test decreases on measures of child abuse potential (i.e., CAPI) were not significant with an $\alpha=0.05$ ($p=0.053$), and the effect size estimate was also small ($g=0.31$). While not statistically significant in this sample, examination of the effect size confidence interval suggests that the majority of potential true scores of child abuse potential would decrease for parents who receive the PCIT intervention when compared to parents in comparison and control conditions.

These results may indicate that measures of child abuse potential and parenting stress are not capturing the future physical abuse construct in a meaningful way. The small effect sizes may also be a function of research design, as the PCIT intervention itself was altered to improve delivery to this population. In traditional PCIT, parent-child dyads do not move from the CDI to the PDI phase or from the PDI phase to completion, until they have mastered the skills of each module (Eyberg & Calzada, 1998). Therefore, treatment length is variable, with participants receiving an average of 14 weekly sessions, with a range between 10 and 20 total sessions (Eyberg et al., 2001). When implementing PCIT with child-welfare involved families however,

ethical considerations supersede these guidelines, especially in instances where a no-treatment wait-list control group is used for comparison (Thomas & Zimmer-Gembeck, 2011, 2012).

Therefore, participants in the studies reviewed were generally moved from one phase to another based on a timeline, rather than mastery of the skills. The decision to restrict treatment exposure may underestimate effect size estimates in the reviewed studies.

The current analysis underscores the clinical benefits of using objective measures of future physical abuse as opposed to proxy risk measures of child abuse potential or parenting stress. In the current meta-analysis, rigorous objective outcomes produced the most sizable treatment effects, although this finding is tempered by a reduced sample size compared to standardized measures. Likewise, designs which allowed treatment to progress as parents mastered PCIT skills produced greater between-group variability than those studies which, for ethical reasons, restricted treatment to 12 weeks. The results of the current meta-analysis suggest that measures of physical abuse recurrence be incorporated into study designs whenever feasible. In instances where a no-treatment, wait-list control group is used, physical abuse recurrence analysis could be conducted for all participants even where subsequent between group analyses are not possible. To this end, Thomas & Zimmer-Gembeck (2011) found that PCIT treatment-completers were significantly less likely to be notified to child protection at one-month follow-up than those participants who dropped out of the study ($\chi^2=7.7$, $p<0.01$).

Our results also highlight the importance of identifying effective social work interventions for child-welfare involved families. Since the Family Preservation and Support Services Program Act was passed in 1993, the goal of child welfare intervention has been to safely maintain children in their biological parent's homes whenever possible (U. S. DHHS, n.d.). Therefore understanding and implementing interventions effective at reducing future

physical abuse are of critical importance to the children and families served by social workers.

PCIT goes beyond the didactic parenting classes typically offered to child-welfare involved families by involving both the parent and the child in therapy sessions, teaching parent and child to play and listen together. This creates an ideal vehicle for social work intervention.

Adjustments to the parent-child relationship can be made in real-time, allowing for coaching and support throughout the process. The analysis presented here suggests that PCIT may facilitate family preservation by halting the coercive cycle and improving family dynamics. Although analysis of physical abuse recurrence was limited, PCIT appears to contribute to the reduction of future physical abuse among this especially vulnerable population.

Limitations

Limitations exist in this review that warrant caution in the interpretation of the results. First, this meta-analysis synthesizes results from a small number of studies on three outcome constructs. Although the current analysis synthesized results from only six primary studies, the technique retains utility and importance (Littell et al., 2008). After receiving endorsement as a best-practice by the Kauffman Best Practices Project, PCIT has been increasingly implemented among child-welfare involved families, although typically outcomes are conceptualized in terms of child behavior. This review adds to the research base by providing an analysis of parent-level outcomes such as subsequent incidence of abuse, child abuse potential, and parenting stress. Analysis results are bolstered by the rigorous experimental or quasi-experimental design employed by the six included studies. Despite the use of experimental or quasi-experimental methods, however, the small number of studies in the current meta-analysis makes cross-study comparisons difficult. Second, only studies published in scholarly journals or available through a dissertations and theses database were included in this review. Informally published or

unpublished works were eligible for inclusion in the meta-analysis only if they were identifiable through reference harvesting or communication with two experts in the field. No such studies were located, thus publication bias is a potential threat to the internal validity of this study (Lipsey & Wilson, 2001). As studies reporting positive and significant results are more likely to be published, the synthesis of studies included in this review could be overestimating the effects of the intervention (Begg, 1994).

Third, participants in each of the studies were exposed to slightly different forms of PCIT (e.g., individual versus group format; time-limited versus skills-based; brief orientation preceding the intervention versus no orientation). And, in studies which compared PCIT to a TAU condition, TAU comprised a heterogeneous array of didactic parent education classes offered through a local agency. It is hypothesized, however, that comparing results between PCIT and TAU conditions underestimates, rather than overestimates treatment effects. Fourth, only two studies conducted between-group analyses of physical abuse recurrence. The raw recidivism rates analyzed here are known to be biased, as physical abuse recurrence did lead to the child being removed from the home in several cases (therefore eliminating future potential abuse reports for that child). For both studies, however, survival analyses and risk deprivation analyses suggested that only PCIT group membership predicted time to report, after controlling for other factors. And finally, participant samples were comprised of a mixture of physically abusive and at-risk families. It is unknown whether and how families with confirmed physical abuse might differ in their response to both PCIT and TAU treatment conditions when compared to at-risk families.

Declaration of interest: The authors declare no conflicts of interest.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abidin, R. R. (1990). *Parenting Stress Index Manual*. Charlottesville, VA: Pediatric Psychology Press.

Abidin, R. R. (1995). *Parenting Stress Index (3rd Ed.)*. Odessa, FL: Psychological Assessment Resources, Inc.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399-409). New York: Russell Sage.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2005). *Comprehensive meta analysis (computer program)*. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1, 97-111. doi:<http://dx.doi.org/10.1002/jrsm.12>

Bowlby, J. (1969). *Attachment and loss: Vol. 1. attachment*. New York: Basic Books.

Bruffaerts, R., Demyttenaere, K., Borges, G., Haro, J. M., Chiu, W. T., Hwang, I. (2010). Childhood adversities as risk factors for onset and persistence of suicidal behaviour. *British Journal of Psychiatry*, 197, 20-27. doi:10.1192/bjp.bp.109.074716

*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., 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.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Eyberg, S. M., Bessmer, J., Newcomb, K., Edwards, D., & Robinson, E. (1994). *Dyadic Parent-Child Interaction Coding System—II manual*. Unpublished manuscript, University of Florida, Gainesville.
- Eyberg, S. M., & Calzada, E. J. (1998). *Parent-child interaction therapy: Procedures manual*. Unpublished manuscript, University of Florida. Retrieved from <http://pcit.phhp.ufl.edu/Presentations/PCIT%20Integrity%20Checklists%20and%20Materials%204-13-06.pdf>
- Eyberg, S. M., Funderburk, B. W., Hembree-Kigin, T., McNeil, C. B., Querido, J., & Hood, K. K. (2001). Parent-child interaction therapy with behavior problem children: One- and two-year maintenance of treatment effects in the family. *Child & Family Behavior Therapy, 23*, 1-20.
- Eyberg, S. M., & Robinson, E. A. (1982). Parent-Child Interaction Training: Effects on family functioning. *Journal of Clinical Child Psychology, 11*, 130–137.
- *Foley, K. P. (2010). *A comparison of parent-child interaction therapy and treatment as usual with families with a history of child abuse and neglect and intimate partner violence*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 3428772).
- Freeman, D., & Fowler, D. (2009). *Routes to psychotic symptoms: Trauma, anxiety and*

- psychosis-like experiences. *Psychiatry Research*, 169, 107-112.
doi:10.1016/j.psychres.2008.07.009
- Galanter, R., Self-Brown, S., Valente, J. R., Dorsey, S., Whitaker, D. J., Bertuglia-Haley, M., & Prieto, M. (2012). Effectiveness of parent-child interaction therapy delivered to at-risk families in the home setting. *Child & Family Behavior Therapy*, 34, 177-196.
- Hembree-Kigin, T., & McNeil, C. B. (1995). *Parent-child interaction therapy*. New York: Plenum.
- Herschell, A.D., & McNeil, C.B. (2005). Theoretical and empirical underpinnings of parent-child interaction therapy with child physical abuse populations. *Education and Treatment of Children*, 28, 142-162.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, 327(7414), 557-560.
doi:http://dx.doi.org/10.1136/bmj.327.7414.557
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I² index? *Psychological Methods*, 11, 193-206.
- Kauffman Best Practices Project (2004). *Closing the quality chasm in child abuse treatment: Identifying and disseminating best practices: Findings of the Kauffman best practices project to help children heal from child abuse*. Charleston, SC: National Crime Victims Research and Treatment Center.
- Lambdin-Shirley, N. (2010). Effectiveness of parent-child interaction therapy (PCIT) with substance abusing mother-child dyads. (Master's thesis). Retrieved from ProQuest Dissertations and Theses. (Order No. 871477296).
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA:

Sage.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic Reviews and Meta-analysis*. New York: Oxford University Press.

Mandelli, L., Carli, V., Roy, A., Serretti, A., & Sarchiapone, M. (2011). The influence of childhood trauma on the onset and repetition of suicidal behavior: An investigation in a high risk sample of male prisoners. *Journal of Psychiatric Research*, 45, 742-747. doi:10.1016/j.jpsychires.2010.11.005

Milner, J. S. (1986). *The Child Abuse Potential Inventory: Manual* (2nd ed.). DeKalb, IL: Psytec.

Milner, J. S. (1994). Assessing physical child abuse risk: The child abuse potential inventory. *Clinical Psychology Review*, 14, 547–583. doi: 10.1016/0272-7358(94)90017-5

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement. *Lancet*, 354, 1896-1900.

Patterson, G. R. (1976). The aggressive child: Victim and architect of a coercive system. In E. J. Mash, L. A. Hamerlynck, & L. C. Handy (Eds.), *Behavior modification and families* (pp. 267–316). New York: Brunner/Mazel.

Sanders, M. R., Cann, W., & Markie-Dadds, C. (2003). The triple P-positive parenting programme: A universal population-level approach to the prevention of child abuse. *Child Abuse Review*, 12, 155-171. doi:http://dx.doi.org/10.1002/car.798

Sedlak, A.J., Mettenburg, J., Basena, M., Petta, I., McPherson, K., Greene, A., & Li, S. (2010). *Fourth National Incidence Study of Child Abuse and Neglect (NIS–4): Report to Congress*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

- Shadish, W. R., and Haddock, C. K. (1994). Combining Estimates of Effect Size. In H. Cooper and L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 261–284). New York: Sage.
- Springer, K.W., Sheridan, J., Kuo, D., & Carnes, M. (2007). Long-term physical and mental health consequences of childhood physical abuse: Results from a large population-based sample of men and women. *Child Abuse & Neglect*, 31, 517-530.
- Tempel, A. B. (2012). Evaluation of parent-child interaction therapy and a standard parenting program within a women's state prison facility. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 1146297772).
- *Terao, S. Y. (1999). Treatment effectiveness of parent-child interaction therapy with physically abusive parent-child dyads. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 304537489).
- Thomas, R., & Zimmer-Gembeck, M. (2007). Behavioral outcomes of parent-child interaction therapy and triple P-positive parenting program: A review and meta-analysis. *Journal of Abnormal Child Psychology: An Official Publication of the International Society for Research in Child and Adolescent Psychopathology*, 35, 475-495. doi:10.1007/s10802-007-9104-9
- *Thomas, R., & Zimmer-Gembeck, M. (2011). Accumulating evidence for parent-child interaction therapy in the prevention of child maltreatment. *Child Development*, 82, 177-192. doi:10.1111/j.1467-8624.2010.01548.x
- *Thomas, R., & Zimmer-Gembeck, M. (2012). Parent-child interaction therapy: An evidence-based treatment for child maltreatment. *Child Maltreatment*, 17, 253-266. doi: <http://dx.doi.org/10.1177/1077559512459555>

Timmer, S. G., Urquiza, A. J., Zebell, N. M., & McGrath, J. M. (2005). Parent-child interaction therapy: Application to maltreating parent-child dyads. *Child Abuse & Neglect*, 29, 825-842. doi:10.1016/j.chiabu.2005.01.003

Timmer, S. G., Ware, L. M., Urquiza, A. J., & Zebell, N. M. (2010). The effectiveness of parent-child interaction therapy for victims of interparental violence. *Violence and Victims*, 25, 486-503.

Turner, H. A., Finkelhor, D., & Ormrod, R. (2006). The effect of lifetime victimization on the mental health of children and adolescents. *Social Science & Medicine*, 62, 13-27. doi:10.1016/j.socscimed.2005.05.030

Urquiza, A. J., & McNeil, C. B. (1996). Parent-child interaction therapy: An intensive dyadic intervention for physically abusive families. *Child Maltreatment*, 1, 134-144.

U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2013). Child maltreatment 2012. Retrieved from <http://www.acf.hhs.gov/programs/cb/research-data-technology/statistics-research/child-maltreatment>

U.S. Department of Health and Human Services, Administration for Children and Families, Child Welfare Information Gateway. (n. d.). Family Preservation and Support Services Program Act of 1993, P.L. 103-66. Retrieved from https://www.childwelfare.gov/systemwide/laws_policies/federal/index.cfm?event=federalLegislation.viewLegis&id=23

Table 1

Overview of client, program, and study characteristics for included studies

	Chaffin et al., 2011	Chaffin et al., 2004	Foley, 2010	Terao, 1999	Thomas & Zimmer- Gembeck, 2011	Thomas & Zimmer- Gembeck, 2012
Sample Size	76 parent-child dyads	110 parent-child dyads*	49 parent-child dyads	34 parent-child dyads	150 parent-child dyads	152 parent-child dyads
Created groups using randomization	Yes	Yes	No	Yes	Yes	Yes
Referral Source	Child protective services/legal	Child protective services/legal	Social service agency	Child protective services/legal	Various sources	Various sources
Professional Implementing PCIT	Trainees and PCIT certified experts	Trainees and PCIT certified experts	--	Social Worker	Clinical Psychologist	Clinical Psychologist
Duration of intervention (weeks)	22-24	22-24	12	14	24.3	--
Frequency of Intervention	Weekly	Weekly	Weekly	Weekly	Weekly	--
Quality Control Procedures	Yes	Yes	Yes	Yes	Yes	Yes
Child Gender (%)						
Male	--	--	61.0	64.7	71.0	70.4
Female	--	--	39.0	35.3	29.0	29.6
Parent Gender (%)						
Male	25.0	35.0	16.0	--	0.0	0.0
Female	75.0	65.0	84.0	--	100.0	100.0
Child Age (mean)	--	--	6.8	4.9	5.0	4.6
Parent Age (mean)	29.0	32.0	32.4	31.6	33.5	33.9
Parent Race/Ethnicity (%)						
Caucasian	55.3	52.0	91.8	--	--	--
African American	18.4	40.0	6.1	--	--	--
Latino/a	7.9	4.0	--	--	--	--
Bi-racial	--	2.0	2.0	--	--	--
Asian	--	1.0	--	--	--	--
Native American	11.8	1.0	--	--	--	--
Australian	--	--	--	--	--	74.0
Aboriginal	--	--	--	--	--	1.4
Other	7.0	--	--	--	--	--

*Note. 33 dyads in Enhanced PCIT (EPCIT) group were eliminated from subsequent statistical analysis.

Table 2

Parent-Child Interaction Therapy Effect Sizes

Outcome	Study Authors	Study Design	Sample Size* (dyads)	Effect Size (95% CI)
Physical abuse recurrence	Chaffin et al., 2011	RCT	76	0.29 (-0.24, 0.82)
	Chaffin et al., 2004	RCT	77	0.77 (0.22, 1.32)
Child abuse potential	Chaffin et al., 2004	RCT	77	0.96 (0.46, 1.45)
	Foley, 2010	Quasi-experiment	43	0.32 (-0.28, 0.91)
	Terao, 1999	RCT	34	0.05 (-.06, 0.71)
	Thomas & Zimmer-Gembeck, 2011	RCT	148	0.07 (-0.27, 0.41)
	Thomas & Zimmer-Gembeck, 2012	RCT	151	0.20 (-0.12, 0.53)
	Foley, 2010	Quasi-experiment	43	0.97 (0.34, 1.59)
Parenting stress	Terao, 1999	RCT	34	0.38 (-0.28, 1.05)
	Thomas & Zimmer-Gembeck, 2011	RCT	148	0.23 (-0.11, 0.57)
	Thomas & Zimmer-Gembeck, 2012	RCT	151	0.15 (-0.18, 0.47)
	Foley, 2010	Quasi-experiment	43	0.97 (0.34, 1.59)

*Note. Sample size indicates the post-test sample used in effect size calculation.

Figure 1

QUORUM Flowchart

