

Behavioral Outcomes of Parent-Child Interaction Therapy and Triple P—Positive Parenting Program: A Review and Meta-Analysis

Rae Thomas · Melanie J. Zimmer-Gembeck

© Springer Science+Business Media, LLC 2007

Abstract We conducted a review and meta-analyses of 24 studies to evaluate and compare the outcomes of two widely disseminated parenting interventions—Parent-Child Interaction Therapy and Triple P-Positive Parenting Program. Participants in all studies were caregivers and 3- to 12-year-old children. In general, our analyses revealed positive effects of both interventions, but effects varied depending on intervention length, components, and source of outcome data. Both interventions reduced parent-reported child behavior and parenting problems. The effect sizes for PCIT were large when outcomes of child and parent behaviors were assessed with parent-report, with the exclusion of Abbreviated PCIT, which had moderate effect sizes. All forms of Triple P had moderate to large effects when outcomes were parent-reported child behaviors and parenting, with the exception of Media Triple P, which had small effects. PCIT and an enhanced version of Triple P were associated with improvements in observed child behaviors. These findings provide information about the relative efficacy of two programs that have received substantial funding in the USA and Australia, and findings should assist in making decisions about allocations of funding and dissemination of these parenting interventions in the future.

Keywords Meta-analysis · Behavioral parenting interventions · Effectiveness of parenting interventions · Child behavior problems · Systematic review

R. Thomas (✉) · M. J. Zimmer-Gembeck
Gold Coast Campus and Griffith Psychological Health Research
Centre, School of Psychology, Griffith University, GPY,
PMB 50 GCMC, QLD 9726, Australia
e-mail: r.thomas@griffith.edu.au

M. J. Zimmer-Gembeck
e-mail: m.zimmer-gembeck@griffith.edu.au

Behavioral parent training interventions are widely used because the evidence shows they are efficacious (Serketich & Dumas, 1996). Although parenting interventions focusing on child behavior management are widespread there are a variety of theoretical foundations and delivery formats for these interventions (Kumpfer & Alvarado, 2003). Nevertheless, behavioral parent training, where parents participate in skills training, is included in many interventions with demonstrated efficacy (Serketich & Dumas, 1996). Often, behavioral parent training involves discussing differential reinforcement and timeout procedures, and an emphasis is placed on the role the parent plays in the development and maintenance of child problem behaviors (Hollenstein, Granic, Stoolmiller, & Snyder, 2004). Yet, interventions might only include parents, might involve family skills training, where the parents and children are taught skills and have time to practice these together with a therapist, or interventions may involve family therapy in which the family unit receives intervention with less of the content directed at skill development.

Clinical benefits of behavioral parent training interventions were described in one meta-analysis of 26 published studies (Serketich & Dumas, 1996). In another systematic review (Barlow & Stewart-Brown, 2000), 16 randomized controlled trials of group parenting interventions were determined to be efficacious for improving child behavior when compared to waitlists or other comparison groups. In this review, however, the heterogeneity of the interventions and study designs, and the difficulties in generalizing the findings were emphasized. There have been few reviews that have systematically compared the efficacy or effectiveness of parenting interventions that have different primary methodologies. Hence, the evidence of efficacy of parent interventions pertains to the broad category of behavioral parent training rather than particular intervention types. In past reviews, summaries included one or two published controlled trials

to represent each of a range of studies that used differing training methodologies.

Recently, recommendations were made regarding the principle characteristics of effective child and family interventions (Kumpfer & Alvarado, 2003; Naton et al., 2003). Kumpfer and Alvarado differentiated intervention methodologies with respects to service delivery, child only interventions, parent only interventions, and family-focused interventions. Family-focused prevention interventions were cited as, in general, producing moderate to very large effect sizes, whereas child only interventions often produced very small effects. Although this paper was not a systematic review of interventions, the distinction between intervention methodologies and their differing effect on family functioning highlights the importance of examining the efficacy of specific types of parenting interventions, and the potential importance of such findings for governmental policy and decision-making with respect to dissemination of interventions.

Parent-Child Interaction Therapy (PCIT) and Triple P—Positive Parenting Program (Triple P) are two parenting interventions described as behavioral parent training interventions, but PCIT and Triple P rely on different delivery formats, and Triple P also uses a range of delivery formats, including using the media, individual, and group intervention. The current review summarizes what is known about the efficacy of PCIT and Triple P, and their subtypes, for three primary reasons. First, in a previous review, an emphasis was placed upon describing and attending to actual intervention content and delivery formats in future meta-analyses (Geeraert, Van den Noortgate, Grietens, & Onghena, 2004). Hence, we described PCIT and Triple P content and delivery formats here, and used meta-analytic techniques to assess what is currently known about the efficacy of these two interventions, which are both founded in social learning theory, but have different modes of service delivery. Although using meta-analysis to summarize the findings from trials of PCIT and trials of Triple P is not a direct comparison of delivery format, as other differences between the two types of interventions and each trial were not controlled, we elected to analyze trials of PCIT and Triple P because of the clear differences in delivery formats and the potential importance of such information for future research and intervention implementation.

Second, the number of trials of PCIT and Triple P has reached a level that makes them amenable to summarization separate from other parenting interventions. We also expected that these findings would be of importance to families, communities, and policymakers, because these two interventions have been (and are continuing to be) widely disseminated and funded in both the USA and Australia. PCIT and Triple P have become popular in Australia in recent years, and both have been used widely or are being dissem-

inated in the USA. Triple P has seen widespread growth in Australia, with recent dissemination to the USA (Prinz & Sanders, 2004). PCIT is used widely in the USA, with more recent dissemination to Australia (Nixon, 2001). PCIT and Triple P have attracted government funding in both countries to support their implementation.

Conclusions about the effectiveness of PCIT and Triple P for improving parenting behaviors and reducing child behavior problems have been made in some previous literature reviews. In one review, PCIT was described as a probably efficacious treatment for children with externalizing and antisocial behavior (Brestan & Eyberg, 1998), whereas Nixon (2002) cited both PCIT and Triple P as parent training interventions with demonstrated efficacy based upon numerous studies with rigorous methodologies. Webster-Stratton and Taylor (2001) cited both intervention types as empirically supported for the reduction of child antisocial behavior. However, a third reason for conducting the meta-analyses reported here was that none of these reviews were systematic reviews of all published evaluation studies of both PCIT and Triple P, and there was no use of quantitative meta-analytic methods to summarize findings from all studies. Based on the conclusions of these narrative reviews and others (e.g., Foote, Schuhmann, Jones, & Eyberg, 1998; Sanders, Cann, & Markie-Dadds, 2003) and the expectation that services provide interventions that are “evidence-based,” Triple P and PCIT have attracted significant funding in the USA and Australia. As such, the evidence of efficacy and impact of differing methodology of both Triple P and PCIT warranted a more detailed examination. In the current study, we completed a review and meta-analyses of all trials of Triple P and PCIT dated between 1980 and 2004.

Theoretical foundations of PCIT and Triple P

PCIT and Triple P have been derived from social and developmental theories. First, some PCIT and Triple P intervention components have a foundation in the propositions of social learning theory (Foote, Eyberg, & Schuhmann, 1998; Sanders et al., 2003). For example, a social learning framework has been employed to direct attention to the interactions between family members as the source of difficulty, rather than implying that the child or the parent is independently responsible for all problems. For minor child misbehavior differential reinforcement is the primary positive parenting strategy taught in both interventions. Differential reinforcement is a planned behavior management strategy of positively rewarding a child’s prosocial behavior whilst providing minimal attention to a child’s inappropriate behavior. In addition to differential reinforcement, both PCIT and Triple P focus on appropriate consequences for child misbehavior. Positive reinforcement of the parent during interactions

with the child is also used in PCIT, but this is not directly integrated into Triple P.

Second, applied behavior analysis, developmental models of social competence, and developmental psychopathology are theories described as foundations for Triple P (Sanders et al., 2003), and attachment theory has been described as a foundation of PCIT (Herschell, Calzada, Eyberg, & McNeil, 2002). Although PCIT and Triple P share theoretical underpinnings, there are often differences in intervention length, service delivery components, and how the intervention format met their specified goals.

Triple P—Positive Parenting Program

Triple P was designed to promote positive parenting and caring relationships between parents and children aged between 2 and 16 years (Sanders et al., 2003). Triple P has been described as a Behavioral Family Intervention (e.g., Sanders & McFarland, 2000) with a multi-tiered continuum of service intervention (Sanders et al., 2003). Families are offered information on parenting and behavior management strategies through a variety of intervention structures that reflect the differing needs of parents. The Triple P service modality is structured to enable parents to access information from a variety of sources including multi-media, professional consultations and self-directed modules. Distinct tiers of Triple P are available (http://www.triplep.net/files/pdf/TripleP_Model.Tabole.pdf). These include Standard, Group, Enhanced, Self-Directed, and Media (i.e., one version of Universal Triple P). A child mean age of five years was found in the Triple P evaluations included in the current study.

Many of the published articles on the efficacy of Triple P are based on Standard Triple P, Group Triple P or Enhanced Triple P. Standard Triple P works with single families, while group sessions are conducted within Group Triple P. Both of these types of Triple P emphasize the role of parents in the development and maintenance of child misbehavior by assisting parents to identify possible causes and establishing goals for behavior change (Turner, Markie-Dadds, & Sanders, 1998). Through didactic presentations, individual or small group activities and homework, parents use differential reinforcement, communication skills, effective consequences for misbehavior, and planned activity scheduling (Turner et al., 1998). Approximately 10 sessions are available for Standard Triple P, whereas, it is usual to offer five group sessions and three telephone consultations to Group Triple P participants.

In Enhanced Triple P, three adjunct modules, tailored to the parents' individual needs, are added to Standard Triple P. Modules listed on the Triple P web site (<http://www1.triplep.net/>) and in the Practitioner's Manual for Enhanced Triple P (Sanders, Markie-Dadds, & Turner,

1998) are Practice, Coping Skills, and Partner Support. In the Practice module, the goals are to identify and resolve problems with implementing new parenting strategies. In the Coping Skills module, parents are assisted with personal issues, such as depression and anxiety. In the Partner Support module, dual parent families who are experiencing difficulties in communication, relationships and/or co-parenting are offered support (Sanders et al., 1998).

The final two interventions of Triple P are Self-Directed and Media. Self-Directed Triple P participants are supplied with a text and a self-help manual and provided with 10 weeks of structured learning tasks and do not have contact with a professional (Sanders, Markie-Dadds, Tully, & Bor, 2000). Skills taught in Self-Directed Triple P are similar to those taught in Standard Triple P and weekly telephone consultations with a professional can be arranged. Media Triple P is described as 12 episodes of an "infotainment" television intervention (Sanders, Montgomery, & Brechman-Toussaint, 2000). Goals of the intervention are described as providing parents with parenting strategies, to normalize the challenge of parenting and to increase community awareness of family relationships (Sanders, Montgomery et al., 2000).

Parent-Child Interaction Therapy (PCIT)

PCIT is an individualized intervention developed for caregivers and their 4- to 7-year-old children with externalizing behavior (Hembree-Kigin & McNeil, 1995). Therapeutic outcomes of PCIT have been described as being guided by observed changes in parent-child interactions, rather than self-report changes in the child's behavior only. A parent is assisted to alter her/his behavior via direct coaching strategies. In turn, this change in parent behavior is expected to improve the child's problem behaviors, and increase positive interactions within the parent-child dyad. Although the above is a shared principle of behavioral parent training, the use of direct, in vivo coaching of parental behaviors differentiates PCIT from most other behavior parent training interventions, including Triple P. In the reviewed studies of PCIT, the average length of treatment was between 12 and 14 weeks.

There are two phases in PCIT, which are labeled Child Directed Interaction and Parent Directed Interaction. Progression from one phase to the next is predominantly dependent on attaining prescribed levels of specific skills known as mastery criteria (Hembree-Kigin & McNeil, 1995). However, in the studies reviewed here, PCIT was often limited to a specific number of sessions (e.g., 12 sessions). PCIT skills are taught via didactic presentations to parents, and direct coaching of parents while they are interacting with their children. In didactic sessions (usually two sessions), the focus is on teaching the parent specific skills related to each phase of the therapy and these sessions are conducted prior

to the direct coaching sessions. The remainder of PCIT (usually about 10–12 sessions in the studies reviewed) involves direct coaching sessions. These sessions are conducted with the parent and child in a play therapy room with the therapist in another room behind a one-way mirror. The therapist and the parent communicate through a “bug-in-the-ear” device. This device permits the therapist to provide direct coaching of parental communication and behavior management skills, immediate feedback and social reinforcement of the parent. Parents are expected to practice the skills at home.

As in Triple P, there are variants of the standard PCIT intervention. Although PCIT was described as a 12-week intervention in most studies, Nixon et al. (2003, 2004) used an abbreviated version of PCIT that followed the same guidelines as Standard PCIT (i.e., two phases of treatment), however the two didactic sessions were replaced with videos for home viewing. In addition, instead of the standard 10–12 PCIT direct coaching sessions, abbreviated PCIT offered five in-vivo coaching sessions, which alternated with five 30-minute telephone consultations.

Another variant of PCIT adds a 6-session motivation component (PCIT + motivation). This delivery format was used in a randomized controlled trial of PCIT with families who had a history of physically maltreating their children (Chaffin et al., 2004). The motivation sessions included videotaped testimonials from previous participants, psychoeducation, decision making exercises, self-motivational cognitions, self-efficacy, and an understanding of the consequences of child maltreatment. Before proceeding to Standard PCIT, participants submitted personal statements regarding parental beliefs, practices and goals for therapy. After PCIT, a 4-week group intervention was implemented in order to improve generalization and maintenance of skills. Also, Chaffin et al. (2004) implemented a second version of Enhanced PCIT. In this version, participants were provided with PCIT plus the motivational component, as well as individual counseling sessions for depression, marital issues or substance abuse, and home visits.

The delivery formats and specific intervention strategies of Triple P and PCIT

According to Salas and Cannon-Bowers (2001), Triple P and PCIT employ strategies that have been found to be effective — relevant information is presented to parents and parenting concepts are demonstrated through role-play in both Triple P and PCIT. Nevertheless, there are at least three key intervention strategies that differ between PCIT and Triple P. Summarizing the effects of these programs was expected to provide important information about the relative efficacy of different intervention components. First, a collaborative learning model is inherent in Triple P group interventions where group interactions assist in the learning process. Col-

laborative learning has demonstrated efficacy and reduces therapist time and resources (Shebilske, Jordan, Goettl, & Paulus, 1998). PCIT does not have a group version. Second, although parents in Triple P are encouraged to conduct homework tasks and discuss these with the therapist, observed practice of parents interacting with children and managing challenging child behaviors does not usually occur as part of the intervention design. Standard Individual and Enhanced Triple P have the provisions for child involvement, but child involvement was not described in the reviewed studies. In contrast to Triple P, the use of coaching strategies that include parents interacting with their children in PCIT is a key intervention strategy and provides opportunities for parents to practice new skills and get immediate feedback about performance from interventionists. Because the parent might be provided with direct remediation of incorrect implementation of skills and practice with implementation that works for the parent and the child, and because the overlearning principles of PCIT evident in the ‘Mastery Criteria’ allow for repeated practice of skills and increases retention rates of newly acquired skills (Driskell, Willis, & Copper, 1992), the direct coaching strategies in PCIT might yield larger effects on both parent and child outcomes. Third, although restricted by a small number of published studies, comparing the effectiveness of standard PCIT and Triple P to interventions that have been enhanced with additional services and other components was expected to add important information for future intervention design and research.

The current study

In the current study, our aim was to provide a comprehensive assessment of the state of the evidence to guide decisions about the implementation and continued dissemination of Triple P and PCIT in Australia and the USA, and to determine whether summarizing the efficacy of PCIT and Triple P allowed for conclusions about effective intervention delivery format. To do this, we identified all randomized controlled trials and single group follow-up studies of Triple P and PCIT dated between 1980 and 2004. We abstracted data from all identified studies, and used standard review and meta-analysis techniques to draw conclusions about the capacity of each type of intervention to improve parent and child behaviors. We provided specific analyses of each subtype of Triple P and PCIT that has been investigated and summarized the efficacy of each type of intervention. We reported effect sizes from independent group pre-treatment/post-treatment (IGPP) studies and single group pre-treatment/post-treatment (SGPP), and effect sizes based upon different sources of information (e.g., parent-report, observation), and compared effect sizes found in Standard PCIT to effect sizes found for Triple P.

Method

Retrieval and selection of studies

We identified studies dated between 1980 and 2004 that examined the efficacy of Triple P or PCIT. Five literature search strategies were used. First, electronic databases were searched (Medline and Psych INFO). Keywords used to identify articles were Triple P, Positive Parenting Program, Parent-Child Interaction Therapy and PCIT. In addition to using the keyword function, searches were also conducted on known Triple P and PCIT researchers (Eyberg, Funderburk, McNeil, Nixon, Sanders and Turner). Second, the reference sections of all retrieved articles were searched for additional trials. Third, major reviews of parental behavioral training and child behavior problems and reference lists were inspected. Fourth, authors of identified trials were contacted and all assisted in the identification of additional publications. Fifth, our searches included book chapters and attempts to obtain unpublished trials from known authors of published trials of PCIT and Triple P, but no additional trials were identified and retrieved.

Included and excluded studies

Studies reported between 1980 and 2004 were included in the meta-analysis if they met three criteria: These criteria included a focus on the efficacy of either Triple P or PCIT, the inclusion of at least one parent or child *behavior problems* outcome measure, and the inclusion of empirical data needed for meta-analysis.

Thirty-two studies of PCIT were retrieved, and we included 13 studies from 8 cohorts and 3 research groups (Chaffin – 1 cohort, Eyberg/McNeil – 6 cohorts, Nixon – 1 cohort; see Table 1). Seven studies reported only means or did not include needed information for meta-analysis, nine studies were not effectiveness studies, one study considered the effectiveness of PCIT within the classroom setting with the teacher coached in PCIT skills (i.e., parents were not included in the intervention), and one study was a duplicate of results reported in an included study.¹ One further study met most of the inclusion criteria, however, it was not included because three treatment groups with different levels of father involvement were compared and total effects for all participants were not reported (Bagner & Eyberg, 2003).

We retrieved 28 studies of Triple P (or earlier versions of this intervention). Eleven studies based on 10 cohorts from 1 research group (Sanders) were included in meta-analyses (see Table 1). Twelve studies did not include enough information to be included in meta-analysis. Four studies

were excluded as they focused on child outcomes other than child behavior problems (e.g., bed-wetting or pain).¹ One study met most inclusion criteria, but results were reported by subgroups in order to examine marital discord as a moderator of treatment effectiveness, and total results for all participants were not reported (Dadds, Schwartz, & Sanders, 1987).

Data abstraction and coding

From each study, we abstracted study and intervention characteristics (see Table 1). Each author abstracted information independently and results were compared to correct inconsistencies. All discrepancies were due to misreading article content or data entry errors by one author. Hence, to correct disagreements, the original reference was consulted and the discrepancies were easily rectified. Intervention characteristics that were extracted included the length of time in treatment, the treatment components (e.g., additional individual sessions for depression, psychoeducational information on ADHD, marital relationship enhancement, etc.), and, for Triple P, whether the intervention was administered to groups or individual parents/families. Other abstracted data were characteristics of participants, study design, and outcomes. When available, parent and child characteristics included child gender, parent gender, race/ethnicity, and child age. Study design factors included an indicator of randomization versus matched/other comparison group, type of comparison group, time in treatment, time to follow-up assessments, size of treatment and comparison groups and attrition rate.

Outcome measures used in each study were classified as negative or positive measures. With the exception of positive and negative behavioral observations, scores on positive measures were reversed so that higher scores always reflected more negative parent and child behaviors.

PCIT study outcomes

The outcomes variables measured in studies of PCIT included child problem behaviors as measured by parent- or teacher-report, clinic or classroom observation, and parent stress and behavior as measured by parent-report. Child problem behavior was assessed with parent-report questionnaires in all 13 PCIT studies, while teacher reports were gathered in two published studies from a single cohort. The most commonly used measures were the Eyberg Child Behavior Inventory (ECBI, Eyberg & Pincus, 1999, mother report), the Parenting Stress Inventory (PSI, Abidin, 1990, mother report), and the Dyadic Parent-Child Interaction Coding System (DPICS, Robinson & Eyberg, 1981, observation).

We included more than one outcome measure from studies of PCIT efficacy, $M = 9$ measurement subscale scores per study with a range from 1 to 24 scores for each family; 10 studies from 6 cohorts included multiple measures of both

¹ A list of references to papers excluded from this review is available from the authors.

Table 1 Studies included in the review and meta-analyses

Cohort code	Authors, year	Design	Referral source	Sample size	Mean child age	% Boys	Comparison groups	Attrition, %	Measures of child behavior	Measures of parents
Parent-Child Interaction Therapy										
A1	Eyberg et al., 1995	RCT	clinic	27	4.5	80	1 T, 1 W	28	DPICS, ECBI	PLOC, PSI, DPICS
A2	Schulmann et al., 1998	RCT	clinic	64	4.9	81	1 T, 1 W	34	ECBI	DPICS, PLOC, PSI
A3	Hood & Eyberg, 2003	RCT	clinic	23	4.7	70	1T, 1W	NR	ECBI	PLOC
B	Eyberg et al., 2001	Single cohort study	clinic	20	4.7	100	1T	35	CBCL, ECBI, DPICS	DPICS, PSI
C1	McNeil et al., 1991	Nonrandomized trial	clinic	30	4.5	100	1 T, 2 MC	NR	ECBI, CTRS, SESBI, DPICS, classroom observation	ECBI, CTRS, SESBI, classroom observation
C2	Funderburk et al., 1998	Nonrandomized trial	clinic	84	4.8	100	1 T, 2 MC	NR	ECBI, CBCL, SESBI, classroom observation	ECBI, CTRS, SESBI, classroom observation
D1	Nixon, 2001	RCT	clinic & media	55	3.9	73	1 T, 1 W, 1 SV	NR	DSM - ODD, ECBI	DPICS, PS, PLOC, PSI, PSOC
D2	Nixon et al., 2003	RCT	clinic & media	75	3.9	71	2 T, 1 W, 1 SV	18	CBCL, DSM - ODD, DPICS, ECBI	DPICS, PS, PLOC, PSI, PSOC
D3	Nixon et al., 2004	RCT	clinic & media	75	3.9	71	2 T, 1 W, 1 SV	18	ECBI, DPICS	ECBI, DPICS
E	McNeil et al., 1999	RCT	clinic	32	5.2	75	1 T, 1 W	NR	ECBI, CBCL	PSI
F	Chaffin et al., 2004	RCT	clinic	110	4-12	NR	2 T, 1 CG	NR	BASC	CAPI, DPICS
G	Brestan et al., 1997	RCT	clinic	30	4.5	83	1 T, 1 W	NR	ECBI	ECBI
H	Eyberg & Robinson, 1982	Single cohort study	NR	7	4.9	85	1 T	NR	DPICS, ECBI	DPICS
Triple P—Positive Parenting Program										
A1	Sanders, Markie-Dadds et al., 2000	RCT	media	305	3.4	68	3 T, 1 W	NR	ECBI, PDR, FOS	PS, FOS, PSOC
A2	Bor et al., 2002	RCT	media	87	3.4	68	2 T, 1 W	28	ECBI, PDR, FOS	PS, FOS, PSOC
B	Ireland et al., 2003	RCT	clinic & media	37	3.7	65	2 T	3	ECBI	PS
C	Sanders, Montgomery et al., 2000	RCT	media	56	4.6	85	1 T, 1 W	NR	ECBI	PS, PSOC

Table 1 Continued

Cohort code	Authors, year	Design	Referral source	Sample size	Mean child age	% Boys	Comparison groups	Attrition, %	Measures of child behavior	Measures of parents
D	Sanders et al., 2004	RCT	clinic & media	98	4.4	50	2 T	25	ECBI, PDR, HCPC, FOS	CAPI, POQ, Blame and Intentionality PAI, PS, PSOC
E	Sanders & McFarland, 2000	RCT	NR	47	4.4	75	2 T	8	CBCL, PDR, FOS	FOS, PSOC
F	Connell et al., 1997	RCT	media	40	4.3	44	1 T, 1 W	42	ECBI, PDR	PS, PSOC
G	Nicholson & Sanders, 1999	RCT	clinic & media	60	9.6	64	1 T, 1 W	30	CBCL, PDR	
H	Hoath & Sanders, 2002	RCT	clinic & media	21	7.7	80	1 T, 1 W	4	ECBI, SESBI, PSBC	PS
I	Leung et al., 2003	RCT	clinic	91	4.2	64	1 T, 1 W	24	ECBI, PDR, SDQ	PS, PSOC
J	Martin & Sanders, 2003	RCT	media	45	5.8	NR	1 T, 1 W	40	ECBI, PSBC	Parental Adjustment, Parental Practices

Note. NR—Not reported; T—Treatment group; W—Waitlist control; MC—Matched control; SV—Social validation; CG—Community Group; ECBI—Eyberg Child Behavior Inventory; PDR—Parent Daily Report; HCPC—Home and Community Problem Checklist; FOS—Family Observation Scale; CAPI—Child Abuse Potential Inventory; POQ—Parent Opinion Questionnaire; PAI—Parental Anger Inventory; PS—Parenting Scale; PSOC—Parenting Sense of Competency Scale; SESBI—Sutter-Eyberg Student Behavior Inventory; PSBC—Problem Setting and Behavior Checklist; CBCL—Child Behavior Checklist; SDQ—Strength and Difficulties Questionnaire; PLOC—Parenting Locus of Control; DPICS—Dyadic Parent-Child Interaction Coding System; CTRS—Conner’s Teacher Rating Scales; ADHD—Attention Deficit Hyperactivity Disorder; DSM-ODD—Diagnostic Statistical Manual - Oppositional Defiant Disorder; PSI—Parenting Stress Index; BASC—Behavior Assessment System for Children.

child and parent outcomes, while 3 studies from 2 cohorts included only measures of child outcomes.

Triple P study outcomes

The outcome measures in studies of Triple P included child problem behaviors as measured by parent- or teacher-report and clinic observation, while parenting stress and behavior was measured by parent-report. Child and parent behaviors were measured by parent-report questionnaires in all 11 studies of Triple P. The most common parent-report measures for child behavior outcomes included the ECBI and Parent Daily Report (PDR, Chamberlain, & Reid, 1987). Clinic observation of child behavior occurred in four studies using a version of the Family Observation Schedule (FOS-R-III, Sanders, Waugh, Tully, & Hynes, 1996, cited in Sanders, Markie-Dadds et al., 2000). One study considered teacher perceptions of child behavior problems measured by the Sutter, Eyberg Student Behavior Inventory (SESBI; Eyberg & Pincus, 1999).

We included more than one outcome measure from studies of Triple P efficacy, $M = 8$ measurement subscale scores per study with a range from 3 to 22 scores per family. Seven studies from 6 cohorts included multiple measures of both child and parent outcomes, while 4 studies from 4 cohorts included only measures of child outcomes.

Descriptions of trial participants

In all included PCIT and Triple P trials, pre-treatment data indicated that the average child was within the clinical or borderline range of externalizing behaviors, as measured by either the ECBI (Intensity Scale), CBCL externalizing T score or the Behavior Assessment System for Children (BASC) externalizing T score. In 12 of the 13 studies of PCIT, the mean score for child externalizing behavior was reported within the clinical range. Seven of 11 studies of Triple P reported pre-treatment child externalizing mean scores within the clinical range, whereas the other four studies reported pre-treatment means within the borderline range. Two studies, one of PCIT (Chaffin et al., 2004) and one of Triple P (Sanders et al., 2004), tested the efficacy of intervention with families at risk of or engaged in maltreatment. Both of these studies reported that the average child had a pre-treatment mean externalizing behavior score within the borderline range.

Some studies restricted child participants by excluding children with developmental disorders and/or intellectual impairment (Brestan, Eyberg, Boggs, & Algina, 1997; Connell, Sanders, & Markie-Dadds, 1997; Eyberg et al., 2001; Hoath & Sanders, 2002; Leung, Sanders, Leung, & Lau, 2003; Sanders & McFarland, 2000; Schuhmann, Foote, Eyberg,

Boggs, & Algina, 1998), whereas exclusionary characteristics of parents were more varied. Parents' intellectual impairment was an exclusionary criterion in two studies of Triple P (Hoath & Sanders 2002; Sanders, Markie-Dadds, Tully, & Bor, 2000) and four studies of PCIT (Funderburk et al., 1993; Hood & Eyberg, 2003; McNeil, Eyberg, Eisenstadt, Newcomb, & Funderburk, 1991; Schuhmann et al., 1998). In two other Triple P studies parents' major psychiatric diagnosis was an exclusionary criterion (Sanders & McFarland, 2000; Leung et al., 2003).

Some studies specified inclusionary criteria. In studies of Triple P, these included parental depression (Sanders, Markie-Dadds et al., 2000), marital conflict (Ireland, Sanders, & Markie-Dadds, 2003; Sanders Markie-Dadds et al., 2000), minimum 20 h/week employment (Martin & Sanders, 2003), and parental notification of child maltreatment (not necessarily confirmed) or elevated scores on three subscales of the State-Trait Anger Expression Inventory (Sanders et al., 2004). One study of PCIT specified the requirements of a confirmed notification of child physical maltreatment (Chaffin et al., 2004). The other 12 PCIT studies did not specify inclusionary criteria for parents.

Computation and analyses of effect sizes

We computed single group effects from pre-treatment to post-treatment (SGPP) and pre-treatment to follow-up (4 months to one year after treatment completion; SGPF). We also analyzed independent groups pre-treatment/post-treatment (IGPP) and independent groups pre-treatment/follow-up (IGPF) effects. In the following sections, we refer to treatment versus "comparison" groups in the IGPP and IGPF analyses. Comparison groups included randomized waitlists, matched control groups, and alternative community intervention or social validation ("normal community") groups (see Table 1).

SGPP and SGPF effect sizes

The formula used to calculate a single group effect size (SGPP or SGPF) was the following:

$$d_{sg} = (M_{post} - M_{pre})/SD_{pre}$$

where,

d_{sg} was the single group effect size (SGPP or SGPF) reported in Tables,

M_{post} was the mean value at post-treatment or at follow-up,

M_{pre} was the mean value at pre-treatment, and,

SD_{pre} was the standard deviation at pre-treatment.

In one single group study, means and standard deviations were not reported, so paired *t*-test values were used to calculate effect sizes.

IGPP and IGPF effect sizes

As suggested by Becker (1988) and Morris & DeShon (2002), the final IGPP and IGPF effect sizes used in analyses were calculated as the difference between the treatment and comparison single group effects. The three formulas used to calculate IGPP effect sizes and IGPF effect sizes were:

$$\begin{aligned}d_t &= (M_{t\text{-post}} - M_{t\text{-pre}}) / SD_{t\text{-pre}} \\d_c &= (M_{c\text{-post}} - M_{c\text{-pre}}) / SD_{c\text{-pre}} \\d_{ig} &= d_t - d_c\end{aligned}$$

where,

d_t was the effect size for the treatment group,
 d_c was the effect size for the comparison group,
 d_{ig} was the final IGPP or IGPF effect size reported in the Tables and Figures,
 $M_{t\text{-post}}$ and $M_{c\text{-post}}$ were the mean values for the treatment group (t-post) or the comparison group (c-post) at post-test or at follow-up,
 $M_{t\text{-pre}}$ and $M_{c\text{-pre}}$ were the mean values for the treatment group (t-pre) or the comparison group (c-pre) at pre-test,
 $SD_{c\text{-pre}}$ and $SD_{t\text{-pre}}$ were the standard deviations at pre-test.

Hence, d_{ig} was the treatment effect size net of the comparison group effect size. When d_{ig} was positive this indicated a larger effect in the treatment group than in the comparison group. When d_{ig} was negative this indicated a larger effect in the comparison group than in the treatment group.

We used pre-test *SDs* for all calculations, as suggested by Becker (1988), because pre-test *SDs* are often more consistent across studies than post-test *SDs*. The use of pre-test *SDs* to estimate effect sizes, therefore, results in estimates that are more comparable across different experimental manipulations (in this case, between studies of PCIT and Triple P). Post-test *SDs* have been found to be less consistent than pre-test *SDs* as a result of different experimental manipulations. We kept single group effect sizes separate from independent group effect sizes, because there is continuing debate about when these can be combined, and because single group effect sizes are often larger than independent group effect sizes (see Carlson & Schmidt, 1999; Morris & DeShon, 2002). Although there are many reasons that single group and independent group effect sizes might differ and both study designs can have their own biases, one reason single group effect sizes often are larger is because they do not take into account the time effect (e.g., history or maturation effect) that can be partly accounted for by a comparison group (Morris & DeShon, 2002).

tion effect) that can be partly accounted for by a comparison group (Morris & DeShon, 2002).

PCIT effect sizes

Thirteen studies were included in calculations of SGPP effect sizes; eight studies were included in calculations of IGPP effect sizes. Four studies were available for calculations of SGPF and IGPF effects. We calculated the following for child behavioral outcomes: 65 SGPP effects, 30 SGPF (4 month to 1 year after treatment) effects, 49 IGPP effect sizes, and 14 IGPF effect sizes. For parenting behaviors, we calculated 47 SGPP effects, 14 SGPF (4 month to 1 year after treatment) effects, 54 IGPP effects, and 16 IGPF effects.

Triple P effect sizes

Eleven studies were included in the calculations of SGPP; whereas seven studies were included in calculations of IGPP. There were no studies with comparison groups that included a follow-up assessment. We calculated the following for child behavioral outcomes: 78 SGPP effects, 27 SGPF (4 month to 1 year after treatment) effects, and 53 IGPP effects. For parenting behaviors, we calculated 77 SGPP effects, 55 SGPF (4 month to 1 year after treatment) effects, and 50 IGPP effects.

Reduction of effect sizes and final analysis

Effect sizes were categorized by purpose of the measure (to assess child or parent), method (questionnaire versus observation) and reporter (mother, father, teacher). In some cases, effect sizes in the same category within a study were averaged after accounting for repetition from multiple studies from the same sample. This averaging was done to reduce bias that might be introduced from studies that used more measures, and findings reported in more than one publication were included only once. In sum, we report one effect size per study sample for (a) child behavior by method and reporter and (b) parenting by method and reporter. Effect sizes were analyzed with DSTAT (Johnson, 1989) to compute bias-corrected summary effect sizes, confidence intervals, *r* values, to examine homogeneity of effect sizes and outliers, and to compare effect sizes for PCIT to those for Triple P. Since there was homogeneity of effect sizes in all analyses, we did not examine correlates of effect sizes.

Most observational and survey measures used within trials of PCIT and Triple P were for the assessment of negative child or negative parenting behaviors. Hence, except in the cases of *observations* of positive child or parent behavior, negative effect sizes were expected as they would indicate declines in problem behaviors from pre-treatment to later assessments in studies with single group repeated

measurement designs, and greater declines in the treatment group compared to the comparison group in studies with independent groups repeated measurement designs. With the exception of one PCIT study (Eyberg & Robinson, 1982), coding of observational measures in all studies was reported to have been conducted by coders who were independent of the study and blind to participant status.

Results: PCIT

Child behavior change during treatment: Single group treatment effects

In PCIT, medium to large effect sizes were observed for single group child behavior change from pre-to post-treatment (see Table 2). Improvements were found in clinic observations of both negative and positive child behavior, $d = -.54$ and $.94$, respectively, and large effects were found when outcomes were based on mother and father reports of child negative behavior, $d = -1.31$ and $-.83$, respectively. In one study of 10 children, no significant changes were found in classroom observations of positive child behavior and teacher report of negative child behavior from pre to post treatment.

Effect sizes for child behavior change from pre-to follow-up, however, were less likely to be significantly different from 0. The only statistically significant effect size was for mother (or mixed mother/father) report of negative child behavior, $d = -1.10$. No significant change in child behaviors from pre-treatment to follow-up was found when effect sizes were based on clinic observation, father report (when isolated from mother report), teacher report, or classroom observation.

Child behavior: Treatment versus comparison groups

Standard PCIT

When compared to waitlist, medium and large effects ranging from absolute values of $.61$ – 1.45 were found favoring PCIT for mother (or mixed mother/father) reports of negative child behavior, and father reported negative child behavior (see Table 2). However, no significant effect was found for observed negative child behaviors. When PCIT children were compared to normal (i.e., nonclinical) comparison groups, mothers in PCIT reported greater declines in their children's behavior problems than other mothers and large effects were also found for positive behaviors as observed in the classroom, and teacher reports of negative behaviors of children in PCIT compared to a normal comparison group. Effect sizes ranged from an absolute value of 1.21 – 1.57 . When compared to a deviant community group, teachers reported

greater improvements in negative behaviors of children in PCIT $d = -1.16$ however there were no significant effects for observations of positive behavior in the classroom.

Abbreviated PCIT

No significant effects of Abbreviated PCIT were found when child behaviors were compared to waitlist comparison groups (see Table 2). When Abbreviated PCIT participants were compared to a normal community comparison group, there was no significant effect on observed positive child behavior, but there was significant improvement for PCIT participants when child behaviors were assessed with mother report, $d = -1.57$.

Enhanced PCIT

No comparisons of Enhanced PCIT (PCIT + motivation or PCIT + motivation + individual) versus waitlist comparison groups were conducted. Instead, in one study, parents with a history of maltreating their children were assigned to PCIT + motivation or a community group didactic intervention. In this study, parent-reports of child behavior were collected, and the effect of PCIT was large, $d = -.83$ (see Table 2). A large effect on mother-reported child behavior problems also was found for a PCIT + motivation + individual intervention when compared to a community group didactic intervention, $d = -2.16$.

Parent change during treatment: Single group treatment effects

There were significant changes in parenting outcomes pre-treatment to post-treatment, with effect sizes for negative and positive parenting ranging from an absolute value of 1.11 – 3.11 (see Table 3). The only exception was a nonsignificant effect size found for father-reported negative behaviors in one small study of 12 fathers, $d = -.68$.

Medium to large effects were found for clinic observations of negative parent behaviors and parent-report measures of negative parent factors from pre-treatment to follow-up. Effect sizes ranged from an absolute value of $.61$ to $.94$.

Parenting: Treatment versus comparison groups

Standard PCIT

There was more improvement in parent behavior and functioning among parents in PCIT than those in a waitlist. All effect sizes except one (observations of fathers' negative behaviors in one small study of 22 fathers) were significant and usually large in magnitude, d ranged from an absolute value of $.76$ to 5.67 (see Table 3). Similar large effects of PCIT were found for observations of mothers, $d = -1.03$,

Table 2 PCIT child outcomes with treatment group change and comparison effects

Measure	Cohorts/Studies Included ^a	N	Effect Size (Es)	Es Lower 95% CI	ES upper 95% CI	r	Conclusion
PCIT: Single Group Pre- to Post-Treatment							
Parent report	A,B,C,D,E, F,G,H	146	-1.31	-1.56	-1.06	-.55	Improve
Father report	A,G,H	24	-0.83	-1.42	-0.24	-.38	Improve
Teacher report	C	10	-0.97	-1.90	0.04	-.44	
Clinic observation	B,C,D,G,H	60	-0.54	-0.90	-0.13	-.26	Improve
Clinic observation (positive)	B,C,D,G,H	60	0.94	0.55	1.34	.43	Improve
Father Clinic observation	I	3	-0.90	-2.58	0.78	-.40	
Father Clinic observation (positive)	I	3	1.22	-0.52	2.96	.52	
Classroom observation (positive)	C	10	0.97	-0.03	1.80	.43	
PCIT: Single Group Pre-Treatment to Follow-up							
Parent report	A,B,C,D	60	-1.10	-1.48	-0.72	-.48	Improve
Father report	D	13	-0.58	-1.36	0.20	-.28	
Teacher report	C	9	-0.97	-1.95	0.01	-.43	
Clinic observation	B,D	30	-0.43	-0.94	0.08	-.21	
Clinic observation (positive)	B,D	30	0.30	-0.21	0.81	.14	
Classroom observation (positive)	C	9	0.84	-0.12	1.80	.38	
Independent Groups Comparisons							
Standard PCIT vs. Waitlist							
Parent report	A,D,E,G	137	-1.45	-1.82	-1.07	-.59	*
Father report	A,G	38	-1.23	-1.93	-0.54	-.52	*
Clinic observation	D,G	47	0.11	-0.46	0.68	.06	
Clinic observation (positive)	D,G	47	0.61	0.02	1.20	.29	*
Standard PCIT vs. Normal Community Group							
Parent report	D	38	-1.57	-2.29	-0.84	-.62	*
Teacher report	D	19	-1.23	-2.22	-0.25	-.53	*
Clinic observation (positive)	D	38	0.11	-0.53	0.75	.05	
Classroom observation (positive)	D	19	1.21	0.23	2.19	.52	*
Standard PCIT vs. Deviant Community Group							
Teacher report	C	18	-1.16	-2.16	-0.15	-.50	*
Classroom observation (positive)	C	18	0.80	-0.57	1.31	.37	
Abbreviated PCIT vs. Waitlist							
Parent report	D	37	-0.40	-1.05	0.25	-.20	
Clinic observation	D	37	0.40	-0.25	1.05	.20	
Clinic observation (positive)	D	37	0.00	-0.65	0.65	.00	
Abbreviated PCIT vs. Normal Community Group							
Parent report	D	37	-1.57	-2.31	-0.83	-.62	*
Clinic observation (positive)	D	37	-0.08	-0.73	0.57	-.04	
PCIT + Motivation Enhance vs. Community Group Didactic Program							
Parent report	F	77	-0.83	-1.29	-0.36	-.38	*
PCIT + Motivation + Individual vs. Community Group Didactic Program							
Parent report	F	68	-2.16	-2.76	-1.56	-.73	*

*Indicates that there was a significantly larger pre-treatment to post-treatment change (effect) within the Treatment group compared to the Comparison group. CI = Confidence interval.

^aSee Table 1 for a list of study codes.

and mother-reports of their parenting, $d = -1.59$, when compared to mothers in a normal community comparison group.

Abbreviated PCIT

Compared to waitlist, medium or large effects of Abbreviated PCIT were found for observed changes in positive, $d = .92$,

but not negative, parenting behavior and mother-reports of their parenting, $d = -.74$ (see Table 3). When negative parenting behaviors of participants in Abbreviated PCIT were compared to a normal community group, there were moderate to large effects on clinic observation of negative parental behavior, $d = -.82$, and mother reports of their behaviors, $d = -.75$.

Table 3 PCIT parenting outcomes change and comparison effects

Measure	Cohorts/Studies Included ^a	N	Effect size (ES)	ES lower 95% CI	ES upper 95% CI	r	Conclusion
PCIT: Single Group Pre- to Post-Treatment							
Parent report	A,B,D,E,F,G	122	-1.11	-1.38	-0.84	-.49	Improve
Father report	A	12	-0.68	-1.50	0.14	-.32	
Clinic observation	A,B,D,F,G,H	111	-1.46	-1.76	-1.16	-.59	Improve
Clinic observation (positive)	A,B,D,F,G,H	111	1.15	0.87	1.43	.50	Improve
Father clinic observation	A,H	15	-1.53	-2.34	-0.72	-.61	Improve
Father clinic observation (positive)	A,H	15	3.11	2.05	4.17	.84	Improve
PCIT: Single Group Pre-Treatment to Follow-up							
Parent report	A,B,D	53	-0.83	-1.23	-0.43	-.38	Improve
Clinic observation	B,D	30	-0.94	-1.47	-0.41	-.42	Improve
Clinic observation (positive)	B,D	30	0.61	0.09	1.13	.28	Improve
Independent Groups Comparisons							
Standard PCIT vs. Waitlist							
Parent report	A,D,E,G	125	-1.16	-1.55	-0.78	-.50	*
Father report	A	22	-1.02	-1.91	-0.13	-.45	*
Clinic observation	A,D,G	92	-0.76	-1.19	-0.34	-.36	*
Clinic observation (positive)	A,D,G	92	3.66	2.79	4.52	.88	*
Father clinic observation	A	22	-0.67	-1.53	0.20	-.32	
Father clinic observation (positive)	A	22	5.67	3.79	7.54	.94	*
Standard PCIT vs. Normal Community Group							
Parent report	D	38	-1.59	-2.32	-0.85	-.62	*
Clinic observation	D	38	-1.03	-1.71	-0.35	-.46	*
Abbreviated PCIT vs. Waitlist							
Parent report	D	37	-0.74	-1.41	-0.08	-.35	*
Clinic observation	D	37	-0.54	-1.20	0.12	-.26	
Clinic observation (positive)	D	37	0.92	0.24	1.60	.42	*
Abbreviated PCIT vs. Normal Community Group							
Parent report	D	41	-0.75	-1.39	-0.12	-.35	*
Clinic observation	D	41	-0.82	-1.46	-0.18	-.38	*
PCIT + Motivation vs. Community Group Didactic Program							
Parent report	F	77	-1.15	-1.64	-0.67	-.50	*
Clinic observation	F	77	-1.15	-1.64	-0.67	-.50	*
Clinic observation (positive)	F	77	1.65	1.13	2.16	.64	*
PCIT + Motivation + Individual vs. Community Group Didactic Program							
Parent report	F	68	-0.86	-1.36	-0.36	-.40	*
Clinic observation	F	68	-4.79	-5.70	-3.84	-.92	*
Clinic observation (positive)	F	68	2.32	1.71	2.94	.76	*

*Indicates that there was a significantly larger pre-treatment to post-treatment change (effect) within the Treatment group compared to the Comparison group. CI = Confidence interval.

^a See Table 1 for a list of study codes.

Enhanced PCIT

There was a large effect when Enhanced PCIT + motivation was compared to a community group didactic intervention favoring Enhanced PCIT on all parent measures, *d* ranged from an absolute value of 1.15 to 1.65 (see Table 3). Enhanced PCIT + motivation + individual also was compared to a community group didactic intervention on all parent measures. Large effect sizes ranging from an absolute value of .86 to 4.79 favoring Enhanced PCIT + motivation + individual also were found for observed negative and positive parent behavior and negative parenting based on self-report.

Results: Triple P

Child behavior change during treatment: Single group treatment effects

For Triple P, small and medium effects were found for single group pre- to post-treatment child behavior (see Table 4). Clinic observed negative child behavior, negative child behavior measured by both mother and father reports and stepparent reports of negative child behavior improved. Effect sizes ranged from an absolute value of .31 to .73. However, a small study of 9 children with teacher reports of

Table 4 Triple P child outcomes change and comparison effects

Measure	Cohorts/Studies Included ^a	N	Effect size (ES)	ES lower 95% CI	ES upper 95% CI	r	Conclusion
Triple P: Single Group Pre- to Post-Treatment							
Parent report	A,B,C,D,E,F,G,H,I,J	449	-0.73	-0.86	-0.59	-.34	Improve
Father report	A,B,F	182	-0.52	-0.73	-0.31	-.25	Improve
Teacher report	H	9	-0.04	-0.96	0.88	-.02	
Stepparent report	G	26	-0.65	-1.21	-0.09	-.31	Improve
Clinic observation	A,D,E	274	-0.31	-0.47	-0.14	-.15	Improve
Clinic observation (positive)	D	74	.41	-.06	.83	.20	
Triple P: Single Group Pre-Treatment to Follow-up							
Parent report	A,B,C,D,E,F,H,J	357	-0.70	-0.85	-0.54	-.33	Improve
Father report	A,B	154	-0.45	-0.67	-0.22	-.22	Improve
Clinic observation	A,D,E	269	-0.61	-0.78	-0.43	-.29	Improve
Clinic observation (positive)	D	74	0.36	0.03	0.68	.18	Improve
Independent Groups Comparison, Triple P vs. Waitlist							
Standard, individual							
Parent report	A,G	178	-0.69	-0.99	-0.38	-.32	*
Father report	A	100	-0.60	-1.01	-0.20	-.29	*
Stepparent report	G	42	-0.57	-1.20	0.07	-.27	
Clinic observation	A	136	-0.22	-0.55	0.12	-.11	
Group							
Parent report	I,J	96	-0.67	-1.08	-0.26	-.32	*
Enhanced							
Parent report	A,H	149	-0.96	-1.30	-0.61	-.43	*
Father report	A	101	-0.56	-0.96	-0.16	-.27	*
Teacher report	H	20	-0.03	-0.91	0.85	-.02	
Clinic observation	A	129	-0.46	-0.82	-0.11	-.23	*
Self-Directed							
Parent report	A,F	155	-0.51	-0.84	-0.19	-.25	*
Father report	A,F	128	-1.26	-1.64	-0.88	-.53	*
Clinic observation	A	132	-0.02	-0.36	0.32	-.01	
Media							
Parent report	C	46	-0.79	-1.39	-0.19	-.37	*

*Indicates that there was a significantly larger pre-treatment to post-treatment change (effect) within the Treatment group compared to the Comparison group. CI = Confidence interval.

^aSee Table 1 for a list of study codes.

child behavior did not yield a significant effect size. Analyses of single group pre-to follow-up data on child behavior resulted in small and medium effect sizes for all measures, *d* ranged from an absolute value of .36 to .70 (see Table 4).

Child behavior: Treatment versus comparison groups

Standard Individual Triple P

Medium effects were found in favor of Standard Individual Triple P when compared to waitlist comparison group when child negative behavior was reported by mothers (or mixed mothers/fathers), *d* = -.69 or by fathers, *d* = -.60 (see Table 4). This did not hold for clinic observations of negative child behavior and one study including reports from 42 stepparents.

Group Triple P

When compared to a waitlist control, a medium effect was found for Standard Group Triple P when based on mother (or mixed mother/father) report, *d* = -.67 (see Table 4). No studies used observational methods or collected reports from individuals other than parents.

Enhanced Triple P

When Enhanced Triple P was compared to waitlist, medium and large effect sizes favoring Enhanced Triple P were found for clinic observations of negative child behavior and father and mother reports of negative child behavior. Effect sizes ranged from an absolute value of .46 to .96 (see Table 4). In a small study of 20 teachers, there was no significant effect

of Triple P when negative child behavior was compared to a waitlist comparison group.

Self-directed Triple P

Mothers in Self-Directed Triple P reported moderate improvements, $d = -.51$, while fathers reported large improvements in their children’s behavior, $d = -1.26$, compared to waitlist mothers and fathers (see Table 4). No significant effect on clinic observations of child negative behavior was found.

Media Triple P

Compared to waitlist, Media Triple P had a large effect on child behaviors as reported by parents, $d = -.79$.

Parent change during treatment: Single group treatment effects

Small and medium effects on parenting behaviors were found in analyses of single group pre-to post-treatment data, $d = -.70$ for mother report and $d = -.38$ for father report

(see Table 5). There was no significant effect for observations of negative parent behavior. Small and medium effects were found for pre-to follow-up changes in parenting self-report measures for mother report and father report and observed negative parent behavior ranging from an absolute value of .28 to .69.

Parenting: Treatment versus comparison groups

Standard Individual Triple P

When compared to waitlist, Standard Individual Triple P had a large effect favoring the Triple P intervention when based on mother (or mixed mother/father) reports of parenting, $d = -1.07$, and a medium effect based on father reports of parenting, $d = -.40$ (see Table 5). However, this did not hold for clinic observed negative parent behaviors.

Group Triple P

Compared to a waitlist comparison group, a medium effect size was found in favor of Standard Group Triple P when

Table 5 Triple P parenting outcomes change and comparison effects

Measure	Cohorts/Studies included ^a	N	Effect size (ES)	ES lower 95% CI	ES upper 95% CI	r	Conclusion
Triple P: Single Group Pre- to Post-Treatment							
Parent report	A,B,C,D,E,F,H,I,J	420	-0.70	-0.88	-0.55	-.33	Improve
Father report	A,B,E,F	194	-0.38	-0.58	-0.18	-.12	Improve
Clinic observation	A,E	217	-0.19	-0.38	0.00	-.05	
Triple P: Single Group Pre-Treatment to Follow-up							
Parent report	A,B,C,D,E,F,H,J	354	-0.69	-0.85	-0.54	-.33	Improve
Father report	A,B,E	159	-0.28	-0.50	-0.06	-.14	Improve
Clinic observation	A,E	195	-0.45	-0.65	-0.24	-.22	Improve
Independent Groups Comparison, Triple P vs. Waitlist							
Standard, Individual							
Parent report	A	136	-1.07	-1.43	-0.71	-.47	*
Father report	A	100	-0.40	-0.80	0.00	-.20	$p = .05$
Clinic observation	A	136	0.14	-0.20	0.48	.07	
Group							
Parent report	I,J	96	-0.69	-1.11	-0.27	-.33	*
Enhanced							
Parent report	A,H	149	-0.98	-1.32	-0.63	-.44	*
Father report	A	101	-0.46	-0.86	-0.06	-.23	*
Clinic observation	A	129	-0.11	-0.46	0.24	-.06	
Self-Directed							
Parent report	A,F	155	-1.44	-1.80	-1.09	-.59	*
Father report	A,F	128	-0.36	-0.71	-0.01	-.18	$p = .05$
Clinic observation	A	132	0.07	-0.28	0.41	.03	
Media							
Parent report	C	46	-0.45	-1.03	0.14	-.22	

Note.*indicates that there was a significantly larger pre-treatment to post-treatment change (effect) within the Treatment group compared to the Comparison group. CI = Confidence interval. ^aSee Table 1 for a list of study codes.

based on mother (or mixed mother/father) reports of their parenting, $d = -.69$ (see Table 5).

Enhanced Triple P

When comparing Enhanced Triple P with waitlist, there were statistically significant effects favoring Enhanced Triple P for measures of negative parent factors as reported by mothers (or mixed mother/father; large effect), $d = -.98$ and fathers, $d = -.46$ (medium effect, see Table 5). This did not hold for clinic observations of negative parent behaviors.

Self-directed Triple P

When we analyzed findings comparing Self-Directed Triple P to waitlist, large and small effects favoring Self-Directed Triple P were found based on mother (or mixed mother/father) report, $d = -1.44$, and father report, $d = -.36$, respectively (see Table 5). This did not hold for observed negative parent behavior.

Media

There was no significant effect of Media Triple P on parenting based on self-report questionnaire.

Results: Comparison of PCIT and Triple P

We used categorical modeling to compare the effect sizes for PCIT and Triple P (Johnson, 1989). We conducted analyses to compare SGPP effect sizes, and compared IGPP for Standard PCIT (compared to waitlist) to the multiple forms of Triple P (compared to waitlist).

Single group pre-treatment/post-treatment (SGPP) effect sizes

For SGPP, PCIT had larger effects of parent report of child negative behaviors, but not when effect sizes were based on observed child behaviors. More specifically, there was a larger improvement for study families in PCIT than in Triple P for parent-report of child behavior problems, -1.31 vs. $-.73$, $p < .001$ (see Tables 2–5). However, there was no difference in SGPP effect size for observed child negative behavior, $-.54$ vs. $-.31$, $p = .27$.

Findings when comparing SGPP for parenting problems showed that PCIT had larger effects for both parent report and observed parent negative behaviors. There was a larger improvement for study families in PCIT than in Triple P when outcomes were based on parent report of parenting problems, -1.11 vs. $-.70$, $p < .001$. There also was a larger

improvement in observed parent negative behavior for PCIT than Triple P, -1.46 vs. $-.19$, $p < .001$.

Independent groups pre-treatment/post-treatment (IGPP) effect sizes

Child negative behaviors

Figure 1 shows IGPP effect sizes and confidence intervals for the four forms of PCIT and five forms of Triple P when compared to waitlist. The effects in Fig. 1 are based on parent-reports of child negative behavior. As shown, there were large effect sizes for child negative behavior in all forms of PCIT, except the Abbreviated version, which did not have a significant effect on child negative behavior. There were medium to large effects for all forms of Triple P.

As can be seen by comparing the confidence intervals in Fig. 1, results of our analyses that compared Standard PCIT to multiple forms of Triple P showed differences in PCIT and some forms of Triple P for parent report of child negative behaviors. The effect size for PCIT, -1.45 , was significantly larger than the effect sizes for Self-directed, $-.51$, $p < .001$, Group, $-.67$, $p < .01$, and Individual Triple P, $-.69$, $p < .01$, but the PCIT effect size was not larger than the effect size for Enhanced, $-.96$, or Media Triple P, $-.79$. However, in contrast to these findings, for IGPP based on observed child negative behaviors, there were no differences in the effect size for Standard PCIT, $.11$, compared to Triple P in the Self-directed, $-.02$, Individual, $-.22$, and Enhanced, $-.46$, forms (figure not provided).

Parent negative behaviors

Figure 2 summarizes IGPP effect sizes based on parent reports of parent negative behavior, and illustrates that all forms of PCIT and Triple P, except Media Triple P, had medium or large effects. Media Triple P did not have a significant effect on self-reported negative parent behavior. All other interventions reduced negative parenting of those in treatment when compared to waitlist. Of these effective interventions, Abbreviated PCIT and Group Triple P had the most modest effects on parenting, but effect sizes were still medium in size and quite similar to other, often more intensive and longer forms of these interventions (i.e., all confidence intervals had a range of values in common).

When we statistically compared Standard PCIT to forms of Triple P, PCIT had a larger effect when compared to Media Triple P, but not when compared to other forms of Triple P; the effect size for Standard PCIT, -1.16 , was not different when compared to Self-directed, -1.44 , Group, $-.69$, Individual, -1.07 , and Enhanced Triple P, $-.98$. Yet, PCIT did have a larger effect on parent report of parenting problems than media Triple P, $-.45$, $p < .05$. However, for

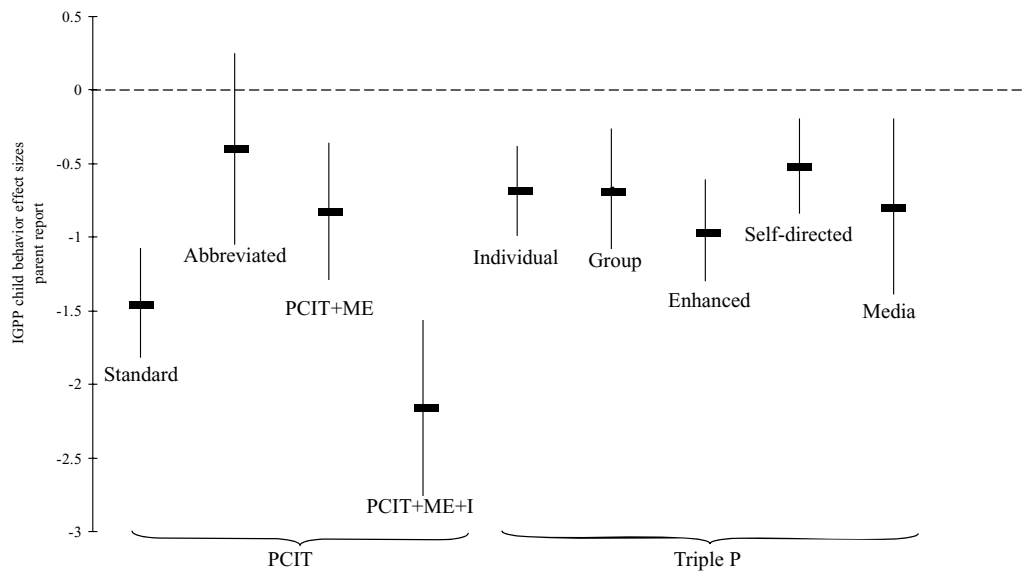


Fig. 1 Independent Groups Pre-treatment/Post-treatment (IGPP) effects for parent report of child behavior problems. *Note:* Lower effect sizes indicate greater declines in the treatment group as compared to

waitlist comparison group. Confidence intervals that cross 0 indicate that the effect size was not significantly different from 0

IGPP of observed parent negative behaviors, Standard PCIT had a larger effect size, $-.76$, than Triple P in the Self-directed, $.07$, $p < .01$, Individual, $.14$, $p < .01$, and Enhanced, $-.11$, $p < .05$, forms (figure not provided).

and has been described as efficacious in previous meta-analytic and review articles (Barlow & Stewart-Brown, 2000; Serketic & Dumas 1996). In an attempt to go beyond statements pertaining to the efficacy of behavioral parent training as a broad category, our review and meta-analyses described the reported efficacy of two parenting interventions with known behavioral parent training origins. The interventions were chosen for this review due to their similar theoretical foundations, and their wide-spread dissemination, popularity and significant level of government funding in both the USA and Australia, but their differing modes of delivery.

Discussion

Behavioral parent training is a popular form of parenting intervention (McMahon & Forehand 2003; Nixon, 2002; Sanders et al., 2003; Webster-Stratton & Taylor 2001)

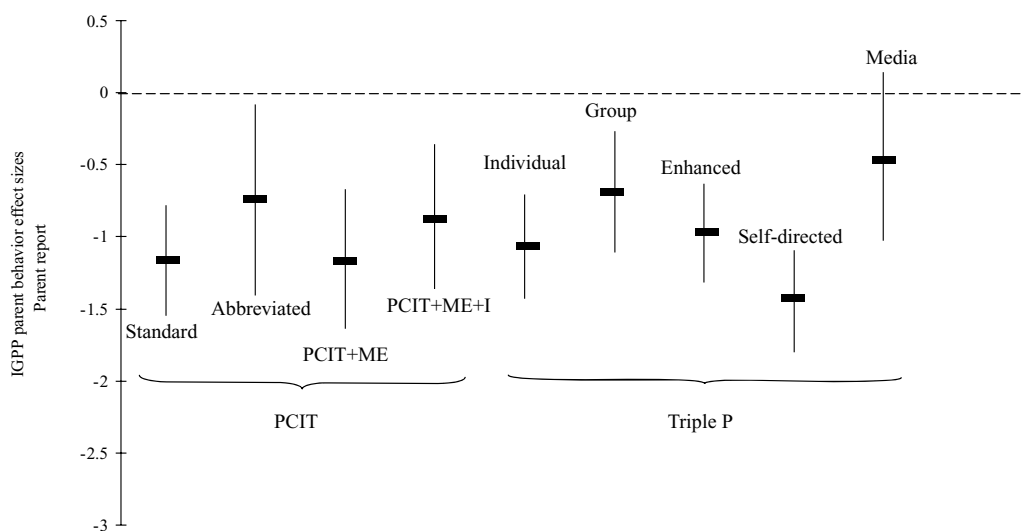


Fig. 2 Independent Groups Pre-treatment/Post-treatment (IGPP) effects for self-reported parenting problems. *Note:* Lower effect sizes indicate greater declines in the treatment group as compared to

waitlist comparison group. Confidence intervals that cross 0 indicate that the effect size was not significantly different from 0

We analyzed 24 evaluations of Parent-Child Interaction Therapy (PCIT) and Triple P-Positive Parenting Program. Findings revealed that these interventions improve parenting, such as improving parental warmth, decreasing parental hostility, increasing parental self-efficacy, and reducing parental stress. Most versions of these interventions also reduce negative child behaviors, such as aggression and extreme tantrums and opposition. These results put numerical weight behind previous narrative reviews in which authors have concluded that interventions are effective when they include behavioral parent training (Nixon, 2002; Webster-Stratton & Taylor, 2001; Weisz, Hawley, & Jensen Doss, 2004).

Participation in PCIT and Triple P results in improvements in child behavior and parenting from pre- to post-treatment (i.e., short-term improvements in behavior). In addition, although few studies included follow-up beyond immediate post-treatment, there was some support from prospective assessment of treatment participants for continued positive effects of these interventions up to 3 months after intervention completion. However, conclusions regarding long-term effectiveness of either intervention must remain tentative. Follow-up data collected from treatment participants were never compared to waitlist or alternative treatment groups. More specifically, follow-up data were not collected from waitlist participants for any of the Triple P studies and follow-up data only were collected from a social validation group in only one series of PCIT studies in Australia (Nixon et al., 2003, 2004).

Findings show that the strength of the effects of PCIT and Triple P, and conclusions about which intervention may result in greater improvements in parenting and family functioning, depended on the measures used to assess outcomes and the subtype of each intervention. Although findings were not without exception, Standard PCIT tended to have larger effects than Triple P when compared to waitlist and when outcomes were based on parent report of child negative behaviors and observed parent negative behaviors. In contrast, there was no effect size difference when findings were based on observed child behaviors and only one difference (Standard PCIT vs. Media Triple P) for parent report of parenting. In addition, Standard PCIT did not have a larger effect than Enhanced Triple P, except when comparing observed parent negative behavior.

For some intervention types and when observational measures were used, effects were not consistently found to be different from 0 (i.e., effects were not significant). Significant treatment effects on *children's behaviors* were found for both Triple P and PCIT interventions when outcomes were assessed via female caregiver (or combined female/male caregiver) reports or observation. Effects were generally medium for Triple P and large for PCIT. The exceptions were a large effect size for Enhanced Triple P and a small and nonsignificant effect size for Abbreviated PCIT. Effect sizes based

on teacher reports were calculated on single studies with small samples, and were rarely as large as those based on female caregiver report or observation and often did not show significant improvements in children's behaviors, but improvements among PCIT participants were greater than waitlist comparison groups (Funderburk et al., 1998; Hoath & Sanders, 2002; McNeil et al., 1991). Large effect sizes were found for father reports of child behavior and parenting when PCIT was compared to a waitlist (Brestan et al., 1997; Nixon et al., 2003, 2004; Schuhmann et al., 1998) and small to large effects were found with the same measures for Enhanced Triple P and Self-Directed Triple P (Connell et al., 1997; Sanders, Markie-Dadds et al., 2000). Parenting during Self-Directed Triple P did not significantly improve when father reported behaviors.

Considerations for the generalization of study findings

There are three key issues to consider when generalizing the results of the current meta-analyses. All issues come from the sampling details included in studies and the measurement strategies used. First, the demographic characteristics of families included in many of the studies were unclear or may have been limited to moderate or higher income families. Only two of the Triple P studies had participants within the lower socio-economic status (SES) group and with low parent education (Hoath & Sanders, 2002; Sanders et al., 2004). Sanders and McFarland (2000) reported participants in the low SES range, however failed to report parental education. Although sometimes unclear, all other participants in Triple P studies appeared to have been from middle class or higher SES groups and, with the exception of one study (Bor, Sanders, & Markie-Dadds, 2002), middle or higher levels of parental education. Due to the high number of Triple P studies in the meta-analysis with middle or higher SES, it is not certain that findings can be generalized to low income or high risk groups at this time.

With the exception of three studies from one study group (Nixon, 2001; Nixon et al., 2003; Nixon et al., 2004), participant demographic data were not consistently reported in studies of PCIT. When income or educational information was reported, participants were low-to-mid SES (Eyberg et al., 2001; Hood & Eyberg, 2003; Nixon, 2001; Nixon et al., 2003, 2004; McNeil et al., 1991; Schuhmann et al., 1998). In the cases when education was reported, PCIT participants had moderate levels of education (Brestan et al., 1997; Eyberg & Robinson, 1982; Nixon, 2001; Nixon et al., 2003, 2004). Given the inconsistency in reporting of demographic data, it also is difficult to draw conclusions about generalization of PCIT to low SES and high-risk families.

Second, the methods of recruiting families differed. Triple P studies most often used media outlets to advertise and recruit families via self-referral (9 of the 11 Triple P studies

used media as a recruitment strategy), while PCIT participants were most often clinic-referred (only 3 of the 13 PCIT studies used media as a recruitment strategy). Only one Triple P study used clinic referred participants (Leung et al., 2003). It remains unclear how much of the difference between PCIT effect sizes and Triple P effect sizes may be due to differing recruitment methods. Previous researchers have suggested that there will be selection bias effects when participants are recruited via the media and self-referred as compared to relying on referrals from clinics or other professional sources (Berlin, O'Neal, & Brooks-Gunn, 1998). Reyno and McGrath (2006) found a moderate effect size when comparing source of referral and treatment outcome in favor of self-referral over referral from professionals. However, we found no evidence of this in the current study. Although, there was some variability in recruitment strategies across PCIT studies, and across Triple P studies, homogeneity of effect sizes was found within studies of PCIT and within studies of Triple P. This suggests that recruitment methods may not have greatly influenced effect size.

Third, parent-reported child behavior problems were assessed with either the Eyberg Child Behavior Inventory (Eyberg & Pincus, 1999) or the Child Behavior Checklist (Achenbach, 1991) in all but one study included in this review. Of studies of PCIT, the ECBI was used in nine studies, the ECBI and the CBCL were used in three studies, and the BASC was used in one study. Similarly, in studies of Triple P, the ECBI was used in nine studies, and the CBCL was used in two studies. This resulted in more meaningful comparisons of the effect sizes for children's behaviors found in studies of PCIT to those found in studies of Triple P.

We also compared intervention types on effect sizes for observed child and parent behaviors and parent-reports of their own parenting behaviors. In contrast to parent reports of child behaviors, there were some differences in the observational measures of children and parents, and measures that assessed parenting, and we cannot rule out the possibility that differences are due to the use of different measures. For example, DPICS (Robinson & Eyberg, 1981) was the observational coding system used in studies of PCIT. DPICS coding resembles and follows directly from treatment activity (i.e., parents are coached in specific communication skills such as praises, descriptions and reflections and DPICS is an observed frequency count of these verbalizations). Progression through treatment often depends on seeing improvements in DPICS scores. In Triple P, a different observational coding system (the Family Observation Schedule, Sanders et al., 1996, cited in Sanders, Markie-Dadds et al., 2000) was used to assess observed behaviors of children and parents. Although the FOS was developed to assess the aims of Triple P it does not directly follow treatment activity nor is progression of treatment relative to this measure, whereas coaching of PCIT skills is directly linked to the DPICS. Further, fam-

ilies are repeatedly videotaped and coded for observed behaviors (using DPICS results throughout PCIT sessions to guide intervention strategies) and this is not done in Triple P. These differences may have made it more likely to find observed improvements in child and parent behaviors among families in PCIT (Weisz, Weiss, Han, Granger, & Morton, 1995). Overall, including observational measures in addition to parent report measures is recommended for future studies, and the finding of more significant effect sizes in PCIT than in Triple P when using observational measures may show that it is advantageous, both for intervention development and outcome studies, to develop observational measures that are designed to test specific intervention aims (as has been done in PCIT).

Implications for clinical practice and intervention dissemination

Triple P and PCIT both have been designed to provide participants with a range of intervention options. For example, therapist contact time and length of service delivery varies for Triple P - from Media Triple P and Self-Directed Triple P with minimal therapist involvement, to Group Triple P and Standard Individual Triple P. Enhanced Triple P also is available, which is administered individually and includes additional sessions on coping skills and partner support. Due to the limited number of comparison studies of various intervention types, it is difficult to determine the benefits of the enhanced intervention versions. Currently the findings do not clearly support the additional benefits of concurrent treatment options for families, such as those provided in Enhanced Triple P and Enhanced PCIT. When compared to a waitlist control, Self-Directed Triple P (with minimal therapist contact) produced similar effect sizes when compared to Enhanced Triple P. Larger effects were found for PCIT with the addition of a motivational enhancement and individual services (PCIT + ME + I) when compared to Abbreviated PCIT. However, these large effect sizes for PCIT + ME + I were likely influenced by the structure of the intervention. Chaffin et al. (2004) reported that participants were required to succeed in the motivational component prior to commencing PCIT. This requirement may have inflated effect sizes, because participants who entered and completed PCIT had already completed a 6-week motivational course and, although retention rates for the motivation course were reported, the overall attrition rate was unclear. These mixed findings are consistent with Reyno and McGrath's (2006) assertion that further study is required to determine whether the addition of concurrent treatment in interventions, which have already been shown to be efficacious, provides any additional positive outcomes for participants. It is possible that participants may not be able to maintain commitment to a variety of intervention approaches simultaneously, and treatment

providers may not have the resources to develop and maintain high standards across all intervention components.

Although PCIT and Triple P have overlapping content, a further implication for clinical practice may be embedded in the differences in design and intensity of each intervention. The somewhat larger, and sometimes significantly larger, effect sizes for Standard PCIT when considering child behavior change may illustrate that intensity and design matter to the level of positive outcomes for parents and children. With respect to child behavior change, it appears that the direct involvement of both the parent and the child in joint therapeutic sessions that provide parents opportunities to practice new skills may produce somewhat larger improvements, especially in parent reports of child behavior problems and observed parent negative behavior. However, there was less evidence that there were greater improvements in observed child negative behavior in PCIT than in Triple P, and parents did not tend to report greater improvements in PCIT than in Triple P. In sum, a more firm determination of differences in behavior change and parenting between interventions awaits further studies to add to the meta-analyses conducted here. Nevertheless, these findings provide some initial support for the usefulness of the intervention methodology of PCIT of enhanced parental practice of skills under direct observation and coaching by a therapist.

Of further interest to clinicians is the notion of dissemination and transportability of interventions, which are designed and tested in controlled clinic environments. Both PCIT and Triple P have been disseminated to community agencies in the USA and Australia, however, there have been no published, independent dissemination or transportability studies from either intervention. This means there is no current evidence for effectiveness of either Triple P or PCIT in a community setting. Triple P and PCIT could enhance evaluations of intervention effectiveness by evaluating the respective interventions in a diversity of settings. As dissemination of each intervention continues, independent evaluations by the community services now implementing PCIT and/or Triple P would increase the knowledge of intervention effectiveness and the effects of dissemination.

Conclusions

According to the current guidelines for evidence-based practice in psychology, which are limited in a number of ways by not including criteria for length of follow-up or study sample size (American Psychological Association, 1995; Chambless & Hollon, 1998), we tentatively conclude that PCIT meets the criteria for a “well-established treatment” and Triple P meets the criteria for a “probably efficacious treatment.” We were precluded from describing Triple P as a well-established treatment, because our search revealed that

Triple P evaluations have not yet been conducted by two independent investigators or investigatory teams. Given the need to make choices about how resources are allocated for a range of interventions, independent trials of PCIT and Triple P with longer-term follow-up and cost-effectiveness analyses should be conducted.

Researchers are charged with the responsibility to provide the community with information that assists effective decision-making. Multiple trials of both PCIT and Triple P have greatly assisted in developing two parenting interventions to assist parents to have better relationships with their children and to reduce child behavior problems. This provides the community with much needed resources. Some of the remaining critical future questions are which intervention and what components are most beneficial for which population and in what context. To determine this, further studies specifically examining participant demographics, comparisons to alternative treatment options and independent evaluations of Triple P will assist clinicians in determining which intervention would best meet the needs of their client and in doing so could possibly alleviate future years of family dysfunction and parenting stress, and alter pathways of delinquent and antisocial behavior among a large number of children and adolescents.

Acknowledgements This project was greatly improved through our conversations with honours and postgraduate students working in the Family Interaction Research Program at Griffith University, including Elbina Avdagic, Angela Anthonysamy, Michelle Hanisch, Kate McCarthy, Leanne McGregor, Rhiarne Pronk, Mark Scholes, Anne Stuksrud, and Judith Warner. We also thank members of the Griffith Psychological Health Research Centre at Griffith University for their willingness to discuss some of the ideas presented in this article.

References

- Abidin, R. R. (1990). *Parenting stress index* (3rd Ed.). Virginia: Paediatric Psychology Press.
- Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*. Burlington, VT: University of Vermont Department of Psychiatry.
- American Psychological Association. (1995). Task Force on Promotion and Dissemination of Psychological Procedures, Division of Clinical Psychology, Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, 48, 3–23.
- Bagner, D. M., & Eyberg, S. M. (2003). Father involvement in parent training: When does it matter? *Journal of Clinical Child and Adolescent Psychology*, 32, 599–605.
- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257–278.
- Barlow, J., & Stewart-Brown, S. (2000). Behavior problem and group-based parent education programs. *Journal of Developmental and Behavioral Pediatrics*, 21, 356–370.
- Berlin, L. J., O’Neal, C. R., & Brooks-Gunn, J. U. (1998). What makes early intervention programs work? The program, its participants, and their interaction. *Zero To Three, February-March*, 4–15.

- *Bor, W., Sanders, M. R., & Markie-Dadds, C. (2002). The effects of the Triple P- Positive Parenting Program on preschool children with co-occurring disruptive behavior and attentional/hyperactive difficulties. *Journal of Abnormal Child Psychology*, *30*, 571–587.
- Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology*, *27*, 180–189.
- *Brestan, E. V., Eyberg, S. M., Boggs, S. R., & Algina, J. (1997). Parent-Child Interaction Therapy: Parents' perceptions of untreated siblings. *Child & Family Behavior Therapy*, *19*, 13–28.
- *Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., Jackson, S., Lensgraf, J., & 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.
- Chamberlain, P., & Reid, J. (1987). Parent observation and report of child symptoms. *Behavioral Assessment*, *9*, 97–109.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7–18.
- *Connell, S., Sanders, M. R., & Markie-Dadds, C. (1997). Self-directed behavioral family intervention for parents of oppositional children in rural and remote areas. *Behavior Modification*, *21*, 379–408.
- Dadds, M. R., Schwartz, S., & Sanders, M. R. (1987). Marital discord and treatment outcome in behavioral treatment of child conduct disorders. *Journal of Consulting and Clinical Psychology*, *55*, 396–403.
- Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effects of over-learning on retention. *Journal of Applied Psychology*, *77*, 615–622.
- *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., Funderburk, B. W., Hembree-Kigin, T. L., McNeil, C. B., Querido, J. G., & Hood, K. K. (2001). Parent-Child Interaction Therapy with behavior problem children: One and two year maintenance of treatment effects in the family. *Child and Family Behavior Therapy*, *23*, 1–19.
- Eyberg, S., & Pincus, D. (1999). *Eyberg child behavior inventory and sutter-eyberg student behavior inventory-revised: Professional manual*. Florida: Psychological Assessment Resources Inc.
- *Eyberg, S. M., & Robinson, E. A. (1982). Parent-Child interaction training: Effects on family functioning. *Journal of Clinical Child Psychology*, *11*, 130–137.
- Foote, R. C., Eyberg, S. M., & Schuhmann, E. M. (1998). Parent-Child Interaction Approaches to the treatment of child behavior problems. In T. H. Ollendick, & Prinz, R. J. (Eds.), *Advances in Clinical Child Psychology* (pp. 125–151). New York and London: Plenum Press.
- Foote, R. C., Schuhmann, E. M., Jones, M. L., & Eyberg, S. M. (1998). Parent-child interaction therapy: A guide for clinicians. *Clinical Child Psychology and Psychiatry*, *3*, 361–373.
- *Funderburk, B. W., Eyberg, S. M., Newcomb, K., McNeil, C. B., Hembree-Kigin, T., & Capage, L. (1993). Parent-Child Interaction Therapy with behavior problem children: Maintenance of treatment effects in the school setting. *Child and Family Behavior Therapy*, *20*, 17–38.
- Geeraert, L., Van Den Noortgate, W., Grietens, H., & Onghena, P. (2004). The effects of early prevention programs for families with young children at risk for physical child abuse and neglect: A meta-analysis. *Child Maltreatment*, *9*, 277–291.
- Hembree-Kigin, T. L., & McNeil, C. B. (1995). *Parent-child interaction therapy*. New York: Plenum Press.
- Herschell, A. D., Calzada, E. J., Eyberg, S. M., & McNeil, C. B. (2002). Parent-child interaction therapy: New directions in research. *Cognitive and Behavioral Practice*, *9*, 9–16.
- *Hoath, F. E., & Sanders, M. R. (2002). A feasibility study of enhanced group Triple P- Positive Parenting Program for parents of children with attention-deficit/hyperactivity disorder. *Behavior Change*, *19*, 191–206.
- Hollenstein, T., Granic, I., Stoolmiller, M., & Snyder, J. (2004). Rigidity in parent-child interactions and the development of externalizing and internalizing behavior in early childhood. *Journal of Abnormal Child Psychology*, *32*, 595–607.
- *Hood, K. K., & Eyberg, S. M. (2003). Outcomes of Parent-Child interaction therapy: Mother's reports on maintenance three to six years after treatment. *Journal of Clinical Child and Adolescent Psychology*, *32*, 419–429.
- *Ireland, J. L., Sanders, M. R., & Markie-Dadds, C. (2003). The impact of parent training on marital functioning: A comparison of two group versions of the Triple P-Positive Parenting Program for parents of children with early-onset conduct problems. *Behavioral and Cognitive Psychotherapy*, *31*, 127–142.
- Johnson, B. T. (1989). *Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Erlbaum.
- Kumpfer, K. L., & Alvarado, R. (2003). Family strengthening approaches for the prevention of youth problem behaviors. *American Psychologist*, *58*, 457–465.
- *Leung, C., Sanders, M. R., Leung, S., & Lau, J. (2003). An outcome evaluation of the implementation of the Triple P-Positive Parenting Program in Hong Kong. *Family Process*, *42*, 531–544.
- *Martin, A. J., & Sanders, M. R. (2003). Balancing work and family: A controlled evaluation of the Triple P-Positive Parenting Program as a work-site intervention. *Child and Adolescent Mental Health*, *8*, 161–169.
- McMahon, R. J., & Forehand, R. L. (2003). *Helping the noncompliant child: Family-based treatment for oppositional behavior*. New York: Guilford Press.
- *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., 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 Psychology*, *20*, 140–151.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105–125.
- Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., & Davino, K. (2003). What works in prevention. *American Psychologist*, *58*, 449–456.
- *Nicholson, J. M., & Sanders, M. R. (1999). Randomized controlled trial of behavioral family intervention for the treatment of child behavior problems in stepfamilies. *Journal of Divorce and Remarriage*, *30*, 1–23.
- *Nixon, R. D. V. (2001). Changes in hyperactivity and temperament in behaviorally disturbed pre-schoolers after Parent-Child Interaction Therapy (PCIT). *Behavior Change*, *18*, 168–176.
- Nixon, R. D. V. (2002). Treatment of behavior problems in preschoolers: A review of parent training programs. *Clinical Psychology Review*, *22*, 525–546.
- *Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent-Child Interaction Therapy: A comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *Journal of Consulting and Clinical Psychology*, *71*, 251–260.

- *Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2004). Parent-Child Interaction Therapy: One and two year follow-up of standard and abbreviated treatments for oppositional preschoolers. *Journal of Abnormal Child Psychology*, *32*, 263–271.
- Prinz, R., & Sanders, M. R. (2004, September). *Preventing child abuse at a broad level: The U.S. Triple P system population-based trial*. Paper presented at ISPCAN 15th International Congress on Child Abuse and Neglect, Brisbane Australia.
- Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems—a meta-analytic review. *Journal of Child Psychology and Psychiatry*, *47*, 99–111.
- Robinson, E. A., & Eyberg, S. M. (1981). The dyadic parent-child interaction coding system: Standardisation and validation. *Journal of Consulting and Clinical Psychology*, *49*, 245–250.
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, *52*, 471–499.
- 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.
- *Sanders, M. R., Markie-Dadds, C., Tully, L. A., & Bor, W. (2000). The Triple P-Positive Parenting Program: A comparison of enhanced, standard and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology*, *68*, 624–640.
- Sanders, M. R., Markie-Dadds, C., & Turner, K. M. T. (1998). *Practitioner's Manual for Enhanced Triple P*. Milton, QLD: Families International Publishing Pty. Ltd.
- *Sanders, M. R., & McFarland, M. (2000). Treatment of depressed mothers with disruptive children: A controlled evaluation of cognitive behavioral family intervention. *Behavior Therapy*, *31*, 89–112.
- *Sanders, M. R., Montgomery, D. T., & Brechman-Toussaint, M. L. (2000). The mass media and the prevention of child behavior problems: The evaluation of a television series to promote positive outcomes for parents and their children. *Journal of Child Psychology and Psychiatry*, *41*, 939–948.
- *Sanders, M. R., Pidgeon, A. M., Gravestock, F., Connors, M. D., Brown, S., & Young, R. W. (2004). Does parental attributional retraining and anger management enhance the effects of the Triple P—Positive Parenting Program with parents at risk of child maltreatment? *Behavior Therapy*, *35*, 513–535.
- *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.
- Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A metaanalysis. *Behavior Therapy*, *27*, 171–186.
- Shebilske, W. L., Jordan, J. A., Goettl, B. P., & Paulus, L. E. (1998). Observation versus hands-on practice of complex skills in dyadic, triadic, and tetradic training-teams. *Human Factors*, *40*, 525–540.
- Turner, K. M. T., Markie-Dadds, C., & Sanders, M. R. (1998). *Facilitator's manual for group triple P*. Milton QLD: Families International Publishing Pty. Ltd.
- Webster-Stratton, C., & Taylor, T. (2001). Nipping early risk factors in the bud: Preventing substance abuse, delinquency, and violence in adolescence through interventions targeted at young children (0–8 years). *Prevention Science*, *2*, 165–192.
- Weisz, J. R., Hawley, K. M., & Jensen Doss, A. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America*, *13*, 729–815.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, *117*, 450–468.

References marked with an were included in the review and meta-analyses.