

Case Study 1: An Evidence-Based Practice Review Report

How effective is Parent-Child Interaction Therapy at reducing disruptive externalising behaviours in young children?

Summary

This literature review evaluates the effectiveness of the evidence based intervention Parent-Child Interaction Therapy (Brinkmeyer & Eyberg, 2003) in reducing disruptive externalising behaviours in young children. It is a parent and child intervention run by a trained therapist following a manual. It is based on positive parenting skills that aim to develop the parent and child relationship and reduce disruptive behaviour through two phases; Child-Directed Interaction (CDI) and Parent-Directed Interaction (PDI).

A systematic literature review was carried out on four databases; Web of Science, PsycINFO, ERIC and Medline. Five studies met the inclusion criteria and these were coded using an adapted coding protocol from the Procedural Manual of the Task Force on Evidence-Based Interventions in School Psychology (Kratochwill, 2003) and then given an overall Weight of Evidence score using Gough's (2007) Weight of Evidence Framework. Largely the findings within this review suggest that PCIT has a positive impact on the reduction of disruptive externalising behaviours in young children who have significant behavioural difficulties. The findings and limitations are discussed along with recommendations for future research.

Introduction

What is it?

Parent-Child Interaction Therapy (PCIT) (Brinkmeyer & Eyberg, 2003) is an evidencebased programme which was developed to be used with children who are 2-7 years old. It is based on developing positive parenting skills which can be used with children who are displaying disruptive behaviour (Eyberg, Nelson & Boggs, 2008). PCIT can be completed usually within 12-20 1 hour weekly sessions however, it is not time limited and is dependent on when parents have mastered the set of skills that are taught. It is also dependent on whether parents rate their child's behaviour within the average range on a standardised rating scale (Brinkmeyer & Eyberg, 2003).

The therapist delivering the intervention communicates with the parent through a one way mirror so that it is just the child and the parent within the room, the parent wears an ear piece known as a "bug-in-the-ear" (McNeil & Hembree-Kigin, 2010). It is distinctive in its delivery as it is intensive and the sessions consist of direct parent-child interaction coaching (Funderburk & Eyberg, 2011).

PCIT involves working with parents and children through two phases; Child-Directed Interaction (CDI) and Parent-Directed Interaction (PDI), which are outlined in the intervention manual (McNeil & Hembree-Kigin, 2010). Parents are taught CDI and PDI skills in a single session before the child is present (Eyberg et al., 2008). The CDI phase consists of building the parent-child relationship and positive parenting skills through child led play based on play therapy skills (McNeil & Hembree-Kigin, 2010). There is an emphasis on

increasing positive attention rather than focusing on negative behaviours and parents are encouraged to actively ignore (Funderburk & Eyberg, 2011; Eyberg et al., 2008). This phase is intended for parents to develop PRIDE skills; praise, reflection, imitation, description and enthusiasm. Throughout the sessions parents are coached with an importance on avoiding leading negative behaviours such as; commands, questioning, criticism, sarcasm or negative physical behaviours (Funderburk & Eyberg, 2011). These skills are a foundation for the next phase, PDI. This intends to develop the parents' skills for improving child compliance and decreasing disruptive behaviours (McNeil & Hembree-Kigin, 2010). Parents are taught to praise compliance, use good commands, and also to use a time out chair for non-compliance and establish consistent rules within the home (McNeil & Hembree-Kigin, 2010). They are coached to give clear instructions and to follow through on the praise and time out rules calmly and consistently (Eyberg et al., 2008). PCIT also involves parents completing homework to imbed and practise the skills (Eyberg et al., 2008).

Psychological relevance

The theoretical underpinnings for PCIT come from the social learning theory (Bandura, 1977) in particular Patterson's Coercion Theory (1982). It is also based on principles from attachment theory (Bowlby, 1978) and operant conditioning (Skinner, 1938) particularly in the CDI phase.

Coercion theory (Patterson, 1982) describes a process where the child's negative behaviour is a result of cycles of reinforcement from the parents. It is where the parent positively reinforces through coercive behaviour which

causes the child to comply (Osofsky, Stepka & King, 2017). Therefore, the child's negative behaviour is reinforced by parents and becomes a learnt behaviour which could be relevant to children displaying disruptive behaviour. PCIT attempts to target these negative interactions and change them to positive reinforcement such as; praise with an emphasis on building a warm relationship between the parent and child. PCIT incorporates the basic principles of behaviour analysis (Skinner, 1938) where the consequences the child receives will influence their behaviour and some will strengthen the behaviour meaning that it is more likely to occur again (Shriver & Allen, 2008).

Although PCIT is centrally based on the social learning theory and operant conditioning it also incorporates attachment theory such as; the importance of a strong parent and child relationship and developing parent's positive responses to their child's behaviour in a supportive and non-coercive way. Another important aspect of PCIT is that it aims to change the child's environment by giving the parents homework and encouraging them to practice the learnt skills across contexts to increase consistency of their interactions (Allen, Timmer & Urquiza, 2014).

Rationale

It is important to work with the parents and child together as research has found that parents directly affect their child's behaviour and emotional development (Patterson & Fisher, 2002; McNeil & Hembree-Kigin, 2010). Research into responsiveness found that positive parent responses led to positive child behaviour and positive child socialisation later in life. The developmental theory also suggested that it was necessary for child

development to have parental responsiveness such as; warmth and sensitivity (Shriver & Allen, 2008).

It is necessary to highlight the importance of early intervention which is recommended in the Special Educational Needs and Disability Code of Practice: 0 to 25 years (SENDCoP) (DfE & DoH, 2015). When behaviour problems occur frequently and they are of high intensity then they are likely to continue to occur as the child gets older without the right support. It has been found that if there are problems within the parent and child relationship then these problems are likely to persist into school age therefore it is important to target these behaviours through early intervention (McNeil & Hembree-Kigin, 2010). This is also relevant in terms of the attachment theory (Bowlby, 1978) if children have a poor early attachment with their primary caregivers this can be a significant risk factor in later life (Allen et al., 2014).

The SENDCoP (DfE & DoH, 2015) emphasises the importance of Educational Psychologists (EPs) working with parents and a key part of their role is to recommend interventions to schools, other settings or parents directly. Therefore, it is necessary for EPs to be aware of the principles behind PCIT and the evidence base for its effectiveness in reducing disruptive behaviour for young children.

Review question

How effective is Parent-Child Interaction Therapy at reducing disruptive externalising behaviours in young children?

Critical Review of the Evidence Base for Parent-Child

Interaction Therapy

Literature Search

A literature search was conducted in January 2019 on four electronic databases. The databases were; Web of Science, PsycINFO, ERIC and Medline and the search terms can be seen in Table 1 below. These search terms were used to find potential studies that were likely to answer the research question and to include in this review.

Table 1

Search Terms

Databases searched	Search Terms
Web of Science	“Parent-Child Interaction Therapy” OR PCIT
PsycINFO	AND
ERIC	“conduct problems” OR conduct OR behaviour OR behavior
Medline	

This search generated 188 studies however 93 of these were duplicates. The studies were screened by title excluding 31 and then by abstract excluding 50. This meant that there were 14 studies which had full text screening using the inclusion and exclusion criteria which can be found in Table 2. Following full text screening 9 studies were excluded which left 5 which met the inclusion criteria and therefore are included in this review, these are listed in Table 3. Figure 1 is a flow diagram which shows the search process and the list of excluded studies with exclusion criteria can be found in Appendix 1.

Table 2

Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria	Justification
1. Language	The study must be published in the English language	The study is published in a language that is different to English	Reviewer doesn't have the means to translate studies that are not published in English
2. Intervention	The intervention must be Parent-Child Interaction Therapy	The intervention is different to Parent-Child Interaction Therapy or a variation of it	To ensure that the effects were from Parent-Child Interaction Therapy
3. Type of design	The study must be a randomised control trial	The study is not a randomised control trial	The review compared the effectiveness of the intervention using randomised control trials
4. Comparison group	The study must include Parent-Child Interaction Therapy in comparison to a wait list control group, treatment as usual or group Parent-Child Interaction Therapy	The study compares Parent-Child Interaction Therapy to another intervention	The reviewer does not have to extract the information and this review is looking at the impact of Parent-Child Interaction Therapy rather than another intervention
5. Participants	The study must include pre-school or primary age children who have disruptive	The study does not include participants who are of pre-school or primary age or	The reviewer aims to seek the effectiveness of this intervention in helping young children with

	Inclusion Criteria	Exclusion Criteria	Justification
	or challenging behaviour	that present disruptive or challenging behaviour	disruptive or challenging behaviour
6. Measures	The study must have pre and post intervention data available	The study does not have pre and post intervention data available	To ensure that the reviewer could effectively review the study
7. Type of publication	The study is in a peer reviewed journal	The study is not in a peer reviewed journal	Peer reviewed journals mean that the quality of the studies have been checked
8. Publication date	The study must have been published between 2007-2019	The study was published before 2007	The reviewer wanted to include the most recent studies of Parent-Child Interaction Therapy

Figure 1

A Flow Diagram of the Literature Search Process

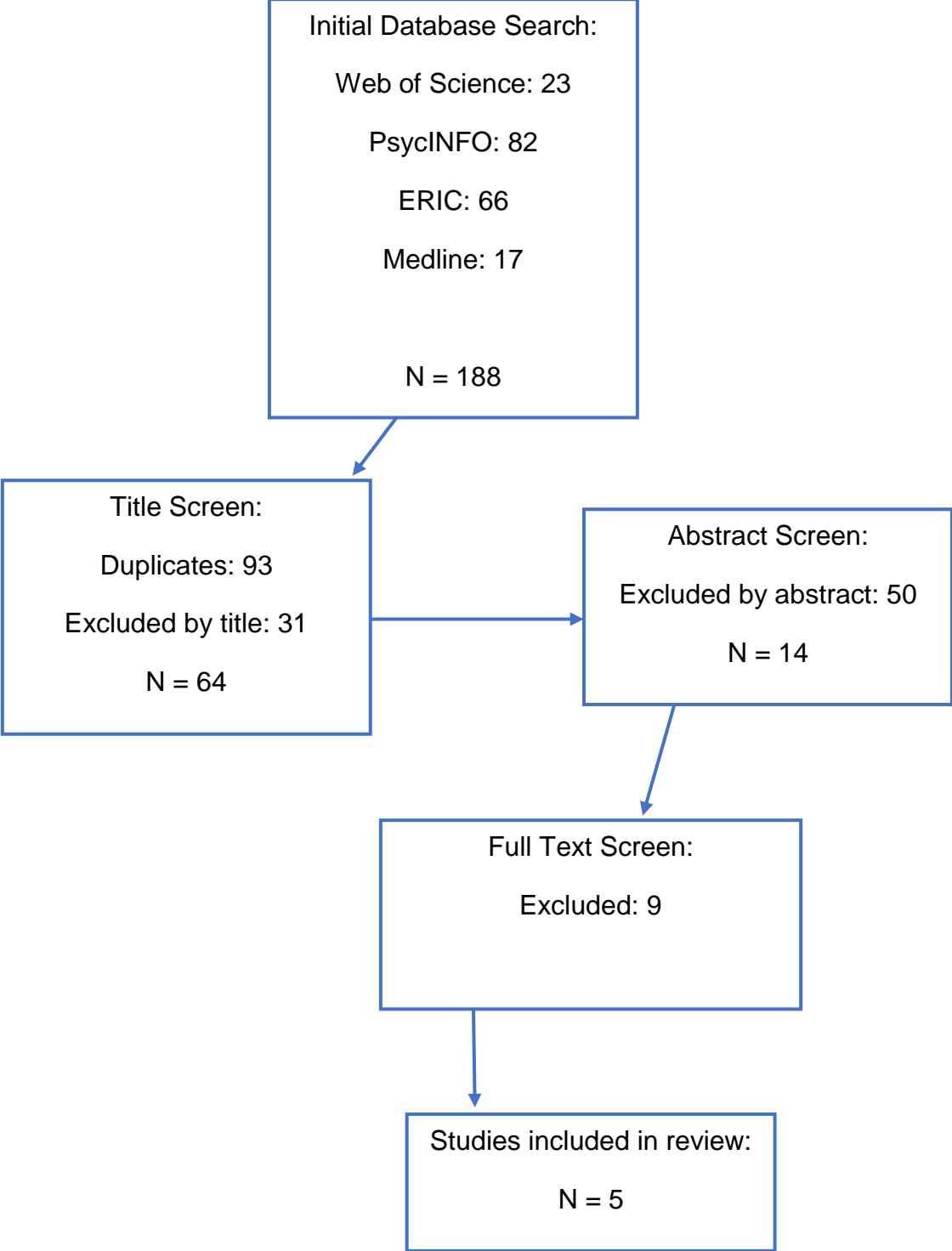


Table 3

List of studies included in review

Study References
1. Bjørseth, Å., & Wichstrøm, L. (2016). Effectiveness of Parent-Child Interaction Therapy (PCIT) in the Treatment of Young Children's Behavior Problems. A Randomized Controlled Study. <i>PLoS ONE</i> , 11(9), 1–19.
2. Bagner, D.M., & Eyberg, S.M. (2007). Parent-Child Interaction Therapy for Disruptive Behavior in Children with Mental Retardation: A Randomized Controlled Trial. <i>Journal of Clinical Child and Adolescent Psychology</i> , 36(3), 418–429.
3. Niec, L.N., Barnett, M.L., Prewett, M.S., & Shanley Chatham, J.R. (2016). Group Parent-Child Interaction Therapy: A Randomized Control Trial for the Treatment of Conduct Problems in Young Children. <i>Journal of Consulting and Clinical Psychology</i> , 84(8), 682–698.
4. Matos, M., Bauermeister, J., & Bernal, G. (2009). Parent-Child Interaction Therapy for Puerto Rican Preschool Children with ADHD and Behavior Problems: A Pilot Study. <i>Family Process</i> , 232.
5. Bagner, D. M., Sheinkopf, S. J., Vohr, B. R., & Lester, B. M. (2010). Parenting Intervention for Externalizing Behavior Problems in Children Born Premature: An Initial Examination. <i>Journal of Developmental & Behavioral Paediatric</i> , 31(3), 209–216.

Critical appraisal of the studies

The five studies included in this review were evaluated for quality and relevance using the Weight of Evidence (WoE) Framework (Gough, 2007).

This framework outlines three dimensions to work out the weighting to give to different studies. It includes WoE A which is the quality of methodology, WoE B which is the relevance of the methodology in relation to the review question and WoE C which is the relevance of the study in answering the review

question. These are then all combined and averaged together to give an overall WoE D of each study, which is the extent to which the study answers the review question (Table 4).

For WoE A an adapted version of the coding protocol for group designs from the Procedural Manual of the Task Force on Evidence-Based Interventions in School Psychology (Kratochwill, 2003) was used for all five studies. See Appendix 4 for a list of all items excluded from the coding protocol and the rationale for this. The completed coding protocols for each study are included in Appendix 3 and the criteria for WoE B, C and D can be found in Appendix 4.

Table 4

Weight of Evidence for the Studies

Study	WoE A	WoE B	WoE C	WoE D
Bjørseth & Wichstrøm, 2016.	2	2	3	2.33 Medium
Bagner & Eyberg, 2007.	1.5	1	2	1.5 Medium
Niec et al., 2016	2	2	1	1.66 Medium
Matos et al., 2009	1.75	2	1	1.58 Medium

Study	WoE A	WoE B	WoE C	WoE D
Bagner et al., 2010	2	2	2	2
Medium				

Weak 0 – 1.4 Medium 1.5 – 2.4 Strong 2.5 – 3

Participants

All studies were published between 2007- 2016, three were carried out in the USA (Bagner & Eyberg, 2007; Niec et al., 2016; Bagner et al., 2010) one in Norway (Bjorseth & Wichstrom, 2016) and one in Puerto Rico (Matos et al., 2009). Four of the studies reported the demographics of the participants however, although Matos et al. (2009) reported comparisons between the two groups in terms of demographics they did not provide specific details which was reflected in their WoE C score. There were a total of 252 participants across the five studies with sample sizes ranging from 28 – 81. The majority of the participants were boys who represented 158 out of 220 participants, excluding the 32 participants from Matos et al. (2009) as their demographics were unclear. The participants' ages ranged from 1.5 (Bagner et al., 2010) – 7 years old (Bjorseth & Wichstrom, 2016).

There was inconsistency between the studies regarding the involvement of both mothers and fathers. Bjorseth and Wichstrom (2016) included both mothers and fathers within their study in all but 10 cases which included just mothers' reports. However, when they reported the findings it was unclear of the numbers in each group. Niec et al. (2016) included findings from 81 mothers and 46 fathers within their study. Bagner et al. (2010) included just the mothers of the children within the study, similarly, Bagner and Eyberg

(2007) only reported the mother's views although they did have 6 fathers participating. Matos et al. (2009) had a mixture of mothers, fathers, step fathers and single mothers however, they just reported the findings from the family so the reviewer wasn't able to make a comparison between mothers' and fathers' ratings. Previously, research has found that the involvement of a father figure within an intervention leads to better maintenance long term or treatment gains (Bagner & Eyberg, 2003).

All studies took place in clinics and participants had all been referred for significant behavioural problems which is relevant to the current review. Bjorseth and Wichstrom (2016)'s participants had consecutively been referred to two outpatient CAMHS clinics for behavioural problems and scored ≤ 120 on the Eyberg Child Behaviour Inventory (ECBI). Niec et al.'s (2016) participants were seeking support from a mental health clinic and had a diagnosis of oppositional defiant disorder (ODD) or conduct disorder. Similarly, Bagner and Eyberg's (2007) participants had a diagnosis of ODD and mild to moderate "mental retardation", they were referred through health care professionals, teachers or self-referral. The participants in Matos et al.'s (2009) study had a diagnosis of ADHD specifically the hyperactivity type rather than inattentive and they had significant behavioural problems. They were all attending a pre-school and accessed the study through pre-school's referral or self-referral. Although Matos et al. (2009) reported these demographics of their participants they failed to give further details which was why their WoE C score was low. Finally, Bagner et al. (2010) included participants who were born ≤ 37 weeks gestation and had significant behavioural problems. They were referred

through neonatal follow up clinic, health professionals, an early intervention programme or again they were self-referred.

Intervention

Niec et al. (2016) and Matos et al. (2009) both scored low on WoE C due to making adaptations to the PCIT intervention. Niec et al. (2016) put time restraints on their intervention, keeping the time constant for both the individual PCIT group and the group PCIT comparison group, which meant that the reviewer couldn't conclude whether there was a difference in completion for both groups and this reduced the fidelity. It also meant that families may not have been ready to finish therapy or may not have achieved the termination criteria set out in the manual (Brinkmeyer & Eyberg, 2003). Matos et al. (2009) made adaptations regarding mastery criteria and treatment extension, they also did not use ECBI to terminate the treatment condition. The ECBI is usually used to assess whether the children's behaviour problems have become normalised according to this scale thus being a termination criterion for PCIT. Bagner et al. (2010) did not report whether their therapists had received training, however, they did highlight that the manual was followed and this is reflected in the medium WoE C score.

Design

In terms of implementation fidelity, the measures taken within the studies varied. In three of the studies therapists received weekly supervision (Niec et al., 2016; Matos et al., 2009; Bagner & Eyberg, 2007). However, this was either with an author of the study or a co-therapist rather than someone who was

uninvolved which could have influenced discussions and the objectivity of these supervisions. In Bjorseth and Wichstrom's (2016) study they had weekly supervisions for the first two cases and then this went down to monthly, however, they did receive feedback from a co-therapist who observed them during sessions but this was only when the co-therapist was available, which they reported to be 45% of the time. This was reflected in their medium score for implementation fidelity within WoE A. Niec et al.'s (2016) study was the only one where supervisors carried out live observations and coding of videoed sessions was completed by advanced PCIT therapists who were not involved in the study, getting 88% treatment fidelity resulting in a high score for implementation fidelity within WoE A. Bagner et al. (2010) did not report any supervision for their therapists throughout the intervention but a research assistant randomly coded videoed sessions who again, was uninvolved and scored 94% adherence to the manual which was reflected in their medium WoE A score. Bagner and Eyberg (2007) also randomly coded their videotapes but they did not report who coded these scoring 97% adherence to the manual. However, they did report that there was weekly supervision, use of the manual and formal training which resulted in a high implementation fidelity score within WoE A. Furthermore, Matos et al. (2009) used a research assistant to code their videotapes and scored 98% but it was unclear whether they were involved within the study. It would have been useful if observers who were not related to the study could have checked if the therapists were using the manualised approach although all studies did report the use of the manual, which means they should be replicable.

Between the five studies there is variation in the length of time between pre and post treatment assessments and follow up. Three completed follow up assessments including both groups used in the study however none of the studies completed multiple follow up assessments which would have resulted in a high WoE A score for follow up assessment or two follow up assessments which would have resulted in a high WoE B score. Bjorseth and Wichstrom (2016) carried out post treatment at 6 months and a follow up assessment at 18 months after the start of treatment which was reflected in their medium WoE A score. Niec et al. (2016) completed a post treatment assessment. However, they did not specify when this was completed. They also completed a six-month follow up. Bagner et al. (2010) completed a post treatment assessment at 4 months and a follow up assessment at 8 months which was reflected in their medium WoE A score for follow up assessment. Matos et al. (2009) carried out a post treatment assessment at 3.5 months and then a follow up another 3.5 months later but this just included the PCIT group as the waitlist control group were then receiving the intervention, making it impossible to compare the findings against the control group which resulted in a low WoE A score for follow up assessment. However, Bagner and Eyberg (2007) did not complete any follow up assessments and at their 4 month post treatment assessment 4 families had not completed PCIT so were not able to take part, this is reflected in their lower WoE A and B score. Therefore, this makes it difficult to discuss the long term effects of PCIT. Additionally, there was a high rate of attrition within Bagner and Eyberg's (2007) study of 47% which is reflected in the lower medium WoE A score. Matos et al. (2009) had a low rate of attrition of 9% at the follow up assessment however, this was only including

the PCIT group as the waitlist control group did not complete the assessment, which again was reflected in the lower medium WoE A score.

Measures

All studies reported the measures that they used and most of these were consistent. PCIT should be evaluated at post treatment using the same measures that were used at pre-treatment (McNeil & Hembree-Kigin, 2010). ECBI (Robinson, Eyberg & Ross, 1980) was used in all five studies, this is a parent rating scale looking at children's behaviour problems using two scales – the intensity scale (the intensity and frequency of problems) and the problem scale (whether parents see that behaviour as being problematic). Three studies used both the intensity and problem scales (Bagner & Eyberg, 2007; Bagner et al., 2010; Matos et al., 2009) whereas, two only used the intensity scale (Niec et al., 2016; Bjorseth & Wichstrom, 2016). Three studies also used the Child Behaviour Checklist (Achenbach, 1991) which is used to measure internalising and externalising behaviours, two of these studies focused particularly on externalising behaviours which is relevant to this review (Bagner & Eyberg, 2007; Bjorseth & Wichstrom, 2016). Another relevant outcome measure used was the Behaviour Assessment System for Children – 2 (BASC-2) (Reynolds & Kamphaus, 2004) and subscales of this focused on hyperactivity and aggression in Matos et al. (2009). All of the studies got a medium WoE A score for measurement as they all used multi measures to report their findings however to score a high WoE A score for measurement there needed to be evidence of multiple sources and reliability and validity reported which none of the studies did. All of the studies were based on parent

report which could be limiting to the studies. It also means that they are subjective and there may be variation in how parents interpret their children's behaviours.

Findings

All five of the studies were group-design and reported the effect sizes within their results section. Cohen's *d* (Cohen, 1992) was reported for Bagner et al. (2010), Bagner and Eyberg (2007) and Bjorseth and Wichstrom (2016). Hedge's *g* (Hedges & Olkin, 1985) was reported for Matos et al. (2009) and partial eta squared was reported for Niec et al. (2016). These were all converted to Hedge's *g* using Ellis (2009) for consistency and due to the sample sizes being <20 in three of the studies (Matos et al., 2009; Bagner et al., 2010; Bagner & Eyberg, 2007). The difference between Cohen's *d* and Hedge's *g* for samples that are over 20 is very small, so it is preferable to report Hedge's *g* (Lakens, 2013).

The effect sizes reported in Table 5 were calculated using the time 2 post treatment measures mean and standard deviation of both conditions as not all of the studies included follow up assessments. When considering the effect sizes found within the studies it is important to reflect on the overall WoE rating that they received however, in this review they were all given a medium rating.

It is key to note that the effect sizes differed between and within studies. The largest effect sizes were found by Bagner et al. (2010) finding positive effects from the ECBI intensity measure and CBCL externalising measure suggesting that PCIT had a positive effect on reducing externalising disruptive behaviours

in children who were born prematurely and had significant behavioural problems. Matos et al. (2009) also found consistently large effect sizes across BASC-PRS and ECBI scales however, it is important to consider that this could be due to the PCIT group receiving 1.5 hour weekly sessions where the three other studies kept the weekly sessions to 1 hour. This is not including Bjorseth and Wichstrom (2016) who did not provide information regarding the length of the weekly sessions.

Niec et al. (2016) found that on the mother's report of ECBI intensity and BASC-2 scales there was no effect between the individual PCIT group and the group PCIT. However, on the father's report there was a medium negative effect on ECBI intensity and BASC-2 externalising suggesting that the group PCIT had more of an effect on reducing externalising behaviours in children with ODD or conduct disorder. Conversely the BASC-2 adaptability measure found a medium effect size suggesting that the individual PCIT had positive effects on increasing adaptive behaviour. They also reported that post treatment gains were maintained at the follow up assessment. It is important to highlight that the group PCIT received 2 hour weekly sessions whereas the individual PCIT received 1 hour weekly sessions which could have had an impact on the findings. Bagner and Eyberg (2007) who also included participants with ODD found large effect sizes on CBCL externalising and ECBI intensity, again highlighting that there was a positive effect on children's externalising behaviour.

Bjorseth and Wichstrom (2016) found small effect sizes on mother's reports on CBCL externalising and ECBI intensity, similarly the father reports found small

effect sizes on CBCL externalising and medium on ECBI intensity. However, it is important to note that in the TAU group the therapists were able to determine the time, frequency and duration of treatment although they excluded interventions that were too similar to PCIT, this was down to the therapist's discretion. The average length of time for the PCIT group was 21.4 sessions and 18.84 sessions for the TAU group. It appears that when the comparison group was receiving TAU or Group PCIT then there was less of an effect size between the two groups.

Table 5

Effect sizes for primary outcomes

Study	Sample size	Measures	Effect size	Effect size descriptor	Overall WoE
Bjørseth & Wichstrøm, 2016.	40 in PCIT group	Mother report:			Medium
	41 in TAU group	CBCL – externalising	g = 0.28	Small	
		ECBI – intensity	g = 0.39	Small	
		Father report:			
		CBCL – externalising	g = 0.24	Small	
		ECBI – intensity	g = 0.51	Medium	
Bagner & Eyberg, 2007.	10 PCIT group	Mother report:			Medium
	12 wait list control group	CBCL – externalising	g = 1.04	Large	
		ECBI – intensity	g = 1.43	Large	
		ECBI – problem	g = 0.64	Medium	

Study	Sample size	Measures	Effect size	Effect size descriptor	Overall WoE
Niec et al., 2016	Mother report:	Mother report:			Medium
	42 in PCIT group	ECBI – intensity	$g = -0.13$	No effect	
	39 in Group PCIT comparison group	BASC-2 – externalising	$g = -0.18$	No effect	
		BASC-2 – internalising	$g = 0.08$	No effect	
	Father report:				
	24 in PCIT group	BASC-2 – adaptability	$g = 0.06$	No effect	
	22 in Group PCIT comparison group	Father report:			
		ECBI – intensity	$g = -0.71$	Medium	
		BASC-2 – externalising	$g = -0.69$	Medium	
BASC-2 – internalising		$g = -0.18$	No effect		
	BASC-2 - adaptability	$g = 0.54$	Medium		
Matos et al., 2009	20 in PCIT group	BASC-PRS – hyperactivity	$g = 1.78$	Large	Medium
	12 in waitlist control group	BASC-PRS - aggression	$g = 1.43$	Large	
		ECBI – intensity	$g = 1.60$	Large	

Study	Sample size	Measures	Effect size	Effect size descriptor	Overall WoE
		ECBI- problem	g = 1.99	Large	
Bagner et al., 2010	11 in PCIT group 14 in waitlist control group	Mother report: CBCL – attention CBCL – aggression CBCL – externalising CBCL – internalising ECBI – intensity ECBI – problem	g = 1.20 g = 2.07 g = 2.45 g = 1.44 g = 2.72 g = 1.68	Large Large Large Large Large Large	Medium

Table 6

Effect size descriptors

Effect size descriptor	Hedge's g
Small	0.2
Medium	0.5
Large	0.8

Conclusion and Recommendations

This review evaluated five studies looking at the efficacy of Parent-Child Interaction Therapy with children with significant behaviour difficulties. They were each given overall WoE ratings of `medium` based on Gough's WoE Framework (2007). The majority of findings were positive with Matos et al. (2009) and Bagner et al. (2010) finding large effect sizes, Bagner and Eyberg (2007) finding large and medium effect sizes, Bjorseth and Wichstrom (2016) finding medium and small effect sizes. However, there was largely no effect found by Niec et al. (2016) but they did find three medium effect sizes on the father's reports. As Bagner et al. (2010) found the largest effect sizes in their study it appears that PCIT is effective in treating behavioural problems for children who are 1.5 – 5 years old and were born prematurely. They followed the manual and ensured the intervention was not time limited and continued until the parents had mastered the interaction skills, these conditions seem to have been effective. These findings suggest that PCIT was effective in reducing externalising behavioural problems in children but there were some limitations to the studies reviewed. These limitations can be found within the WoE appraisal in Table 4.

Only three studies carried out follow up assessments after post treatment measures so this review chose to focus on the post treatment effects and was not able to compare the long term effects of PCIT. The sample sizes used were quite small and they were from specific populations such as; children who were born prematurely, children with ADHD and children with ODD or conduct disorder although all were displaying significant behaviour difficulties.

Therefore, these are not necessarily representative of the wider community and subsequently could limit the generalisability of their findings.

Although the therapists did code (Eyberg, Nelson, Duke & Boggs, 2004) some of the sessions and the studies used largely consistent methods these were principally reliant on parents' reports rather than multiple sources. While it is important to get the parents views as this intervention emphasises the importance of building a positive relationship it would have been useful to gain a range of information in order to triangulate the findings. It could have been that the parents' perceptions of the behaviour had changed rather than the behaviour itself, additional qualitative data may have been helpful to gain some further insight.

With all of this in mind, future research could benefit from having a larger sample size with different populations and active comparison groups. Although within the inclusion criteria of this review these were excluded, the studies with Group PCIT or TAU found less of an effect between groups. Furthermore, due to the small negative effect sizes found in the fathers' reports of Niec et al.'s (2016) study there could be some value added to having group PCIT rather than the traditional individual PCIT but this would require further research. It would also be interesting to look further into the differing views of mothers and fathers within this intervention. Finally, this review has looked at young children in a clinic based setting but an increase in studies exploring different age groups, particularly older children would be useful to determine whether PCIT is effective across age groups and in different settings. In relation to the practice of EPs I feel that this intervention should be recommended by EPs for

children who are presenting with challenging behaviour and for children whose parents may be finding it hard to manage this challenging behaviour. Especially where the negative behaviour of children may be being reinforced by the parents (Patterson, 1982). It is important for EPs to emphasise the collaborative working of parents and children together and PCIT is based on these principles.

References

- Achenbach, T. (1991). *Manual for the teacher's report form and teacher version of the child behaviour profile*. Burlington: University of Vermont, Department of Psychiatry.
- Allen, B., Timmer, S., & Urquiza, A. (2014). Parent-Child Interaction Therapy as an attachment-based intervention: Theoretical rationale and pilot data with adopted children. *Children And Youth Services Review, 47*(3), 334-341.
- 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.
- Bagner, D.M., & Eyberg, S.M. (2007). Parent-Child Interaction Therapy for Disruptive Behavior in Children with Mental Retardation: A Randomized Controlled Trial. *Journal of Clinical Child and Adolescent Psychology, 36*(3), 418–429.
- Bagner, D. M., Sheinkopf, S. J., Vohr, B. R., & Lester, B. M. (2010). Parenting Intervention for Externalizing Behavior Problems in Children Born Premature: An Initial Examination. *Journal of Developmental & Behavioral Paediatric, 31*(3), 209–216.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bjørseth, Å., & Wichstrøm, L. (2016). Effectiveness of Parent-Child Interaction Therapy (PCIT) in the Treatment of Young Children's Behavior Problems. A Randomized Controlled Study. *PLoS ONE, 11*(9), 1–19.
- Bowlby, J. (1978). Attachment theory and its therapeutic implications. *Adolescent Psychiatry, 6*, 5-33.
- Brinkmeyer, M. Y., & Eyberg, S. M. (2003). Parent-child interaction therapy for oppositional children. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidenced-based psychotherapies for children and adolescents* (pp. 204–223). New York: Guilford.
- Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155.
- Department for Education and Department of Health (2015) *Special educational needs and disability code of practice: 0 to 25 years*. Available at: <https://www.gov.uk/government/publications/send-code-of-practice-0-to-25> (Accessed: 8 February 2019).

- Ellis, P.D. (2009), "Effect size calculators," website
<https://www.polyu.edu.hk/mm/effectsizafaqs/calculator/calculator.html>
 accessed on 11th February 2019.
- Eyberg, S. M., Nelson, M. M., Duke, M., & Boggs, S. R. (2004). *Manual for the Dyadic Parent-Child Interaction Coding System*. 3rd ed. Gainsville.
- Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-Based Psychosocial Treatments for Children and Adolescents with Disruptive Behavior. *Journal of Clinical Child & Adolescent Psychology*, 37:1, 215-237.
- Funderburk, B. W., & Eyberg, S. (2011). Parent-Child Interaction Therapy. In J. C. Norcross & K. D. Freedheim (Eds.), *History of psychotherapy: Continuity and change (2nd ed.)* (pp. 415-420). Washington, DC, US: American Psychological Association
- Gough, D., & Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22:2, 213-228.
- Guyatt, G. H., Sackett, D. L., Sinclair, J. C., Hayward, R., Cook, D. J., & Cook, R. J. (1995). Users' Guides to the Medical Literature. *Journal of the American Medical Association*, (2745), 1800-1804.
- Hedges, L. V., & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.
- Kratochwill, T. R. (2003). Task Force on Evidence-Based Interventions in School Psychology. *American Psychology Association*.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t -tests and ANOVAs. *Frontiers in Psychology*, 4(November). 1–12.
- Matos, M., Bauermeister, J., & Bernal, G. (2009). Parent-Child Interaction Therapy for Puerto Rican Preschool Children with ADHD and Behavior Problems: A Pilot Study. *Family Process*, 232.
- McNeil, C. B., & Hembree-Kigin, T. L. (2010). *Parent-child interaction therapy (2nd ed., Issues in Clinical Child Psychology)*. New York: Springer.
- Niec, L.N., Barnett, M.L., Prewett, M.S., & Shanley Chatham, J.R. (2016). Group Parent-Child Interaction Therapy: A Randomized Control Trial for the Treatment of Conduct Problems in Young Children. *Journal of Consulting and Clinical Psychology*, 84(8), 682–698.
- Osofsky, J., Stepka, P., & King, L. (2017). *Parent-child interaction therapy*. (pp. 75-92). American Psychological Association.

- Patterson, G. (1982). *Coercive family process. (A social learning approach; vol.3)*. Eugene, Or: Castalia.
- Patterson, G. R., & Fisher, P. A. (2002). Recent developments in our understanding of parenting: Bidirectional effects, causal models, and the search for parsimony. In M. H. Bornstein (Ed.), *Handbook of parenting: Practical issues in parenting* (pp. 59-88). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Reynolds, C., & Kamphaus, R. (2004). *Behavior assessment system for children* (2nd ed.). Circle Pines, MN: American Guidance Service Inc.
- Robinson, E.A., Eyberg, S.M. & Ross, A.W. (1980). The standardisation of an inventory of child conduct problem behaviours. *Journal of Clinical Child Psychology*, 9, 22-28.
- Shriver, M, D., & Allen, K. D. (2008). *Working with parents of noncompliant children: A guide to evidence-based parent training for practitioners and students*. Washington, DC, US: American Psychological Association.
- Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. Oxford, England: Appleton-Century.

Appendices

Appendix 1: List of excluded studies

Table 7

Excluded studies

Reference	Criteria for exclusion
Eisenstadt, T. H., Eyberg, S., McNeil, C. B., Funderburk, B & Newcomb, K. (1993). Parent-Child Interaction Therapy With Behavior Problem Children : Relative Effectiveness of Two Stages and Overall Treatment Outcome. <i>Journal of Clinical Child Psychology</i> , 22:1, 42-51.	Criteria 8

Reference	Criteria for exclusion
<p>McNeil, C. B., Capage, L. C., Bahl, A., Blanc, H., Capage, L. C. (1999). Importance of Early Intervention for Disruptive Behavior Problems: Comparison of Treatment and Waitlist-Control Groups. <i>Early Education and Development, 10:4</i>, 445-454.</p>	<p>Criteria 8</p>
<p>Rothenberg, W. A., Weinstein, A., Dandes, E. A., & Jent, J. F. (2018). Improving Child Emotion Regulation: Effects of Parent– Child Interaction-therapy and Emotion Socialization Strategies. <i>Journal of Child and Family Studies</i>.</p>	<p>Criteria 3</p>
<p>Shawler, P. M., Bard, M. E., Taylor, E. K., Wilsie, C., Funderburk, B., & Silovsky, J. F. (2018). Children and Youth Services Review Parent-Child Interaction Therapy and young children with Problematic Sexual Behavior: A conceptual overview and treatment considerations. <i>Children and Youth Services Review, 84(August 2017)</i>, 206–214.</p>	<p>Criteria 2</p>
<p>Pade, H., Taube, D.O., Aalborg, A.E., & Reiser, P.J. (2006). An Immediate and Long-Term Study of a Temperament and Parent-Child Interaction Therapy Based Community Program for Pre-schoolers with Behavior Problems. <i>Child & Family Behavior Therapy, 28:3</i>, 1-28.</p>	<p>Criteria 2</p>
<p>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. <i>Child & Behaviour Therapy, 23:4</i>, 1-20.</p>	<p>Criteria 8</p>

Reference	Criteria for exclusion
<p>Abrahamse, M. E., Junger, M., Chavannes, E. L., Coelman, F. J. G., Boer, F., & Lindauer, R. J. L. (2012). Parent – child interaction therapy for preschool children with disruptive behaviour problems in the Netherlands. <i>Child and Adolescent Psychiatry and Mental Health</i>, 6(1), 1.</p>	<p>Criteria 3</p>
<p>McCabe, K., & May, Y. (2009). Parent–Child Interaction Therapy for Mexican Americans: A Randomized Clinical Trial. <i>Journal of Clinical Child & Adolescent Psychology</i>, 38:5, 753-759.</p>	<p>Criteria 4</p>
<p>Veen-Mulders, L., Hoofdaker, B.J., Nauta, M.H., Emmelkamp, P., & Hoekstra, P.J. (2018). Methylphenidate Has Superior Efficacy Over Parent–Child Interaction Therapy for Preschool Children with Disruptive Behaviors. <i>Journal of Child and Adolescent Psychopharmacology</i>, 28(1), 66–73.</p>	<p>Criteria 4</p>

Appendix 2: Mapping the field

Table 8

Summary of studies

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
Bjørseth, Å., & Wichstrøm, L. (2016).	81 Norwegian families with 2-7 year old children who had been consecutively referred to two outpatient CAMHS clinics for behaviour problems and scored ≤ 120 on the Eyberg Child Behavior Inventory (ECBI).	Parent-Child Interaction Therapy two phases: Child-Directed Interaction (CDI) and Parent-Directed Interaction (PDI). Sessions were weekly and both parents were encouraged to attend therapy. Intervention ran	RCT with two groups. PCIT (40) or Treatment as Usual (TAU) (41).	41 families allocated to TAU. 39 completed allocated intervention. Therapists determined the type, duration and frequency of treatment. Consisting of individual therapy with the child, parent	ECBI – intensity Child Behavior Checklist - externalising (CBCL) - externalising Dyadic Parent-Child Interaction Coding System (DPICS). Preschool Age Psychiatric Assessment (PAPA).	Small effect sizes were found on mother's reports on CBCL externalising and ECBI, similarly the father reports found small effect sizes on CBCL externalising and medium on ECBI intensity. So the PCIT was more

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
	Attrition rate from pre-test to 6 and 18 month follow up of 28% and 20%.	<p>until parents had mastered the skills in CDI and PDI at pre-specified levels.</p> <p>Average length of treatment was 21.14 sessions.</p> <p>Norwegian translation of the PCIT protocol.</p> <p>40 allocated to PCIT. 35 completed allocated intervention.</p>		counselling, play therapy, family therapy or Marte Meo, a video-based parent intervention.		effective in reducing disruptive behaviour in the participants.
Bagner, D.M., & Eyberg, S.M. (2007)	30 families, female primary caregivers with children aged 3-6 years old.	<p>PCIT two phases: CDI and PDI.</p> <p>Sessions were weekly lasting 1 hour. The</p>	RCT with two groups. PCIT (15) immediate treatment or waitlist control group (15).	15 families were allocated to a waitlist control group.	<p>CBCL - externalising</p> <p>ECBI – intensity</p> <p>ECBI - problem</p>	Found large effect sizes on CBCL externalising and ECBI intensity,

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
	<p>Participants had “mental retardation” and comorbid oppositional defiant disorder.</p> <p>22 families completed the treatment. At the 4 month follow up there were 8 dropouts. However, 4 receiving PCIT had not yet finished treatment and 2 of these subsequently dropped out.</p>	<p>treatment manual was followed.</p> <p>Average length of treatment for PCIT was 12 sessions.</p>			<p>Parenting Stress Index- Short Form (PSI-SF)</p> <p>DPICS</p> <p>Therapy Attitude Inventory (TAI)</p>	<p>highlighting that there was a positive effect on children’s externalising behaviour in the PCIT group.</p>
Niec, L.N., Barnett, M.L., Prewett, M.S., & Shanley	81 families with 3-6 year old children who had a diagnosis	PCIT two phases: CDI and PDI.	RCT with two groups. PCIT (42) or a novel	Group PCIT. Groups consisted of 3 to	Behavioral Assessment System for Children – 2;	Mother’s report of ECBI intensity and BASC-2 scales found

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
Chatham, J.R. (2016).	<p>of oppositional defiant or conduct disorder.</p> <p>The families were seeking support from a mental health clinic.</p> <p>28 completed the PCIT intervention and 29 completed the group PCIT intervention.</p>	Sessions were weekly for 1 hour.	group PCIT (39).	<p>7 parents and their child.</p> <p>Sessions were weekly for 2 hours to allow for the larger number receiving treatment.</p>	<p>Parent Rating Scale 2-5 year olds and Parent Rating Scale 5-11 year olds (BASC-2) – externalising</p> <p>BASC-2 – internalising</p> <p>BASC-2 - adaptability</p> <p>ECBI - intensity</p> <p>DPICS-III</p> <p>Multidimensional Scale of Perceived Social Support-Child Behavior (MSPSS-C)</p> <p>PSI-SF</p>	<p>there was no effect between the individual PCIT group and the group PCIT. However, on the father's report there was a medium negative effect on ECBI intensity and BASC-2 externalising suggesting that the group PCIT had more of an effect on reducing externalising behaviours in children with ODD or conduct disorder.</p>

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
Matos, M., Bauermeister, J., & Bernal, G. (2009).	32 Puerto Rican families with children aged 4-6 years old with a diagnosis of ADHD and significant behavioural problems. One family dropped out of the PCIT group immediately.	PCIT two phases: CDI and PDI. Sessions were weekly and lasted 1.5 hours. Intervention was delivered in Spanish following adapted treatment manual.	RCT with two groups. PCIT (20) and a 3.5 month wait list control group (12).	12 families were allocated to a 3.5 month waitlist control group.	TAI Disruptive Behaviour Scale for Children – Spanish (DBRS) Hyperactivity and Aggression Subscales of the Behavioral Assessment System for Children-Parent Rating Scale (BASC-PRS) ECBI – intensity ECBI - problem Family Experiences Inventory (FEI)	Mothers reported that there was a reduction in hyperactivity and there were less aggressive and defiant behaviours. Found consistently large effect sizes across BASC-PRS and ECBI scales for the reduction in disruptive behaviours.

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
					Parent Practices Inventory (PPI)	
					Beck Depression Inventory-Spanish (BDI-S)	
					Treatment Evaluation Scale (TES)	
					TAI	
Bagner, D. M., Sheinkopf, S. J., Vohr, B. R., & Lester, B. M. (2010).	28 children aged between 1.5 and 5 years old who were born prematurely and presented with challenging behaviour. 3 families dropped out of	PCIT two phases: CDI and PDI. Sessions were weekly for 1 hour.	RCT with two groups. PCIT (14) immediate treatment and a waitlist control group (14).	14 families were allocated the waitlist control group.	CBCL – attention CBCL – aggression CBCL – externalising	Found positive effects from the ECBI intensity measure and CBCL externalising measure suggesting that PCIT had a positive effect on reducing

Study	Participants	Intervention	Design	Comparison Group	Measures	Outcomes
	the PCIT group during treatment.	Average length of treatment was 13 sessions.			CBCL – internalising ECBI – intensity ECBI - problem DPICS PDI-SF Parenting Scale	externalising disruptive behaviours in children who were born prematurely and had significant behavioural problems.

Appendix 3: Coding Protocol

[Adapted from the Procedural Manual of the Task Force on Evidence-Based Interventions in School Psychology, American Psychology Association, Kratochwill, T.R. (2003)]

Coding Protocol

Name of Coder: X

Date: 05/02/2019

Full Study Reference in proper format:

Niec, L.N., Barnett, M.L., Prewett, M.S., & Shanley Chatham, J.R. (2016). Group Parent-Child Interaction Therapy: A Randomized Control Trial for the Treatment of Conduct Problems in Young Children. *Journal of Consulting and Clinical Psychology, 84*(8), 682–698.

Intervention Name (description of study): Parent-Child Interaction Therapy

Type of Publication:

Book/Monograph

Journal Article

Book Chapter

Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)

Completely randomized design

Randomized block design (between participants, e.g., matched classrooms)

Randomized block design (within participants)

Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)

Nonrandomized design

Nonrandomized block design (between participants)

Nonrandomized block design (within participants)

Nonrandomized hierarchical design

Optional coding for Quasi-experimental designs

A3. Overall confidence of judgment on how participants were assigned (select one of the following)

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- N/A
- Unknown/unable to code

B. Participants

Total size of sample (start of study): 81

Intervention group sample size: 42

Control group sample size: 39

C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment
- Unknown

D. Stage of Program

- Model/demonstration programs
- Early stage programs
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure
- Prior exposure
- Unknown

II. Key Features for Coding Studies and Rating Level of Evidence/ Support

A. Measurement (Estimating the quality of the measures used to establish effects)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes

Yes

No

Unknown/unable to code

A2 Multi-method (at least two assessment methods used)

Yes

No

N/A

Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

Yes

No

N/A

Unknown/unable to code

A4. Validity of measures reported (select one of the following)

Yes validated with specific target group

In part, validated for general population only

No

Unknown/unable to code

Overall Rating for measurement: 2

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

B. Comparison Group

B1 Type of Comparison Group (Select one of the following)

Typical contact

Typical contact (other) specify:

- Attention placebo
- Intervention element placebo
- Alternative intervention
- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
- Unable to identify type of comparison

B2 Overall confidence of judgment on type of comparison group

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc. were counter-balanced across intervention)

- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

B4 Group equivalence established (select one of the following)

- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality

- Low attrition (less than 20 % for post) – 26% (PCIT 26.3% and Group PCIT 25.6%)

- Low attrition (less than 30% for follow-up) – 33% follow up
- Intent to intervene analysis carried out?

Findings: Says this was conducted but doesn't report the results. The phi coefficient testing the association between treatment condition and attrition was not significant

Overall rating for Comparison group: 2

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

F. Implementation Fidelity

F1. Evidence of Acceptable Adherence

- Ongoing supervision/consultation
- Coding intervention sessions/lessons or procedures
- Audio/video tape implementation (select F1.2.1 or F1.3.2):
 - F1.2.1 Entire intervention
 - F1.3.2 Part of the intervention

F2. Manualization (select all that apply)

- Written material involving a detailed account of the exact procedures and the sequence in which they are to be used
- Formal training session that includes a detailed account of the exact procedures and the sequence in which they are to be used
- Written material involving an overview of broad principles and a description of the intervention phases
- Formal or informal training session involving an overview of broad principles and a description of the intervention phases

F3. Adaptation procedures are specified (select one) yes no unknown

Overall rating for Implementation Fidelity: 3

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

I. Follow Up Assessment

- Timing of follow up assessment: 6 months
- Number of participants included in the follow up assessment: 30
- Consistency of assessment method used: Yes

Overall rating for Follow Up Assessment: 1

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

III. Other Descriptive or Supplemental Criteria to Consider

A. External Validity Indicators

Sampling procedures described in detail yes no

Inclusion/exclusion criteria specified yes no

Specified criteria related to concern yes no

Summary of Evidence

Indicator	Overall evidence rating 0-3	Description of evidence Strong Promising Weak No/limited evidence Or Descriptive ratings
General Characteristics		
Design		Completely randomized design
Participants		81

Type of programme		Intervention/treatment
Stage of programme		Established program
Concurrent/ historical intervention exposure		Unknown
Key features		
Measurement	2	Reliability is mostly good. It uses multi-method and sources.
Comparison group	2	There was an alternative intervention comparison group. The same therapists worked with both the individual and group PCIT as they followed the same procedures. High attrition at post treatment (26%) and follow up (33%).
Implementation Fidelity	3	Weekly supervision, coding of intervention sessions showed it was adhering to the intervention protocol. Trainers had a written manual with procedures and a sequence and they had received formal training.
Follow Up Assessment	1	A 6 month follow up assessment was conducted but there was a high attrition rate.

Other Descriptive or Supplemental Criteria to Consider		
External Validity		The sampling procedures are detailed. The inclusion and exclusion criteria and the criteria of entry into the study is in line with the goal of the intervention.

Appendix 4: Weight of Evidence Criteria

Table 9

Amendments to the WoE A coding protocol

Sections not included in the coding protocol	Rationale
I. B1 Appropriate unit of analysis, B2 Familywise error rate controlled and B3 Sufficiently large N	Discussed elsewhere in the review.
I. B7 Coding and B8 Interactive process	Not relevant to current review as there are not qualitative research methods.
II. C1 Appropriate Statistical Analysis for primary outcomes	Discussed throughout review. Findings and effect sizes are also discussed.
II. C2 Percentage of primary outcomes that are statistically significant	Discussed elsewhere in the review.
II. C3 Evidence of appropriate statistical analysis for secondary outcome; C4 Percentage of secondary outcomes that are statistically significant	Not relevant to current review as it focuses on the primary outcomes.
II. C5 Overall summary of questions investigated	Discussed elsewhere in the review.
II. D Educational/clinical significance	Not relevant to current review, the outcomes are discussed elsewhere in the review.
II. E Identifiable components	Not relevant to current review as using Parent-Child Interaction Therapy as a whole intervention.
II. G Replication	Not relevant to current review.
II. H Site of Implementation	Not relevant to current review.
III. A1.2 Inclusion/exclusion criteria similar to school practice	Not relevant to current review.

Sections not included in the coding protocol	Rationale
III. A2 Participant characteristics specified for treatment and control group; A3 Details are provided regarding demographic variables; A4 Receptivity/acceptance by target participant population	Discussed elsewhere in the review.
III. A5 Generalization of effects	Discussed elsewhere in the review.
III. B Length of intervention	Discussed elsewhere in the review.
III. C Intensity of intervention	Discussed elsewhere in the review.
III. D Dosage/response	Not relevant to current review.
III. E Program implementer	Discussed elsewhere in the review.
III. F Characteristics of the intervener	Discussed elsewhere in the review.
III. G Intervention style or orientation	Discussed elsewhere in the review.
III. H Cost analysis data	Not relevant to current review.
III. I Training and support resources	Not relevant to current review.
III. J Feasibility	Not relevant to current review.

Table 10

Criteria for WoE A Measures

Weighting	Description
Strong evidence (3)	<ul style="list-style-type: none"> - Measures produce reliability coefficient of .85 or higher for majority of primary outcomes - Data collected using multiple methods - Data collected from multiple sources - Validity of measures reported to assess primary outcomes or measures are

Weighting	Description
Promising evidence (2)	<p>standardised or norm referenced</p> <ul style="list-style-type: none"> - Measures produce reliability coefficient of .70 - Data collected using multiple methods and/or multiple sources - If multiple primary outcome measures used above criteria must be met for 75% of primary outcomes
Weak evidence (1)	<ul style="list-style-type: none"> - Measures produce reliability coefficient of at least 0.50 - If multiple primary outcome measures used above criteria must be met for at least 50% of primary outcomes
No Evidence (0)	<ul style="list-style-type: none"> - Measures did not produce reliability coefficient scores or these were below 0.50 - Data was not collected using either multiple methods or sources

Table 11

Criteria for WoE A Comparison Group

Weighting	Description
Strong evidence (3)	<ul style="list-style-type: none"> - One other `active` control group - Group equivalence established – through random assignment to conditions - Evidence of counterbalancing of change agent

Weighting	Description
Promising evidence (2)	<ul style="list-style-type: none"> - Equivalent mortality and low attrition evidenced at follow up - One other `no intervention` control group <p>And 2 of the following:</p> <ul style="list-style-type: none"> - Group equivalence established - Evidence of counterbalancing of change agent - Equivalent mortality (intent to intervene analysis) with low attrition
Weak evidence (1)	<ul style="list-style-type: none"> - A comparison group <p>And 1 of the following:</p> <ul style="list-style-type: none"> - Group equivalence established - Evidence of counterbalancing of change agent - Equivalent mortality (intent to intervene analysis) with low attrition
No Evidence (0)	<ul style="list-style-type: none"> - No group equivalence

Table 12

Criteria for WoE A Implementation Fidelity

Weighting	Description
Strong evidence (3)	<ul style="list-style-type: none"> - Strong evidence of acceptable adherence <p>And 2 of the following:</p> <ul style="list-style-type: none"> - Ongoing supervision/consultation,

Weighting	Description
Promising evidence (2)	<p>coding sessions or audio/video tapes</p> <ul style="list-style-type: none"> - Use of a manual which includes 1. Written material with exact procedures and the sequence, 2. Formal training session with exact procedures and sequence - If administered in sessions/lessons then information must be provided for session to session/lesson to lesson basis - Description of adaptation if it occurs <p>- Evidence of acceptable adherence</p> <p>And 2 of the following:</p> <ul style="list-style-type: none"> - Ongoing supervision/consultation, coding sessions or audio/video tapes - Use of a manual which includes 1. Written materials with an overview of the principles and intervention phases, 2. Formal or informal training session with an overview of the principles and intervention phases
Weak evidence (1)	<ul style="list-style-type: none"> - Evidence of acceptable adherence measured through: - Ongoing supervision/consultation, coding sessions or audio/video tapes <p>OR</p> <ul style="list-style-type: none"> - Use of a manual
No Evidence (0)	<ul style="list-style-type: none"> - No evidence of implementation fidelity or

Weighting	Description
	evidence of unacceptable adherence

Table 13

Criteria for WoE A Follow Up Assessment

Weighting	Description
Strong evidence (3)	<ul style="list-style-type: none"> - Must have follow up assessments over multiple intervals - Including all participants from original sample using similar measures used to analyse primary outcomes
Promising evidence (2)	<ul style="list-style-type: none"> - Must have follow up assessments at least once - Including the majority of participants from original sample using similar measures used to analyse primary outcomes
Weak evidence (1)	<ul style="list-style-type: none"> - Must have a follow up assessment at least once - Including some participants from original sample
No Evidence (0)	<ul style="list-style-type: none"> - No follow up assessment

Table 14

Scores for WoE A

Study	Measurement (0-3)	Comparison Group (0-3)	Implementation Fidelity (0-3)	Follow up Assessment (0-3)	Overall score for WoE A

Bjørseth & Wichstrøm, 2016.	2	2	2	2	2	Medium
Bagner & Eyberg, 2007.	2	1	3	0	1.5	Medium
Niec et al., 2016	2	2	3	1	2	Medium
Matos et al., 2009	2	2	2	1	1.75	Medium
Bagner et al., 2010	2	2	2	2	2	Medium

Weak 0 – 1.4 Medium 1.5 – 2.4 Strong 2.5 – 3

WoE B looks at the methodological relevance of the study to the review question and how suitable it is in answering the question (Gough, 2007).

Therefore, WoE B will appraise how appropriate the research design is for evaluating the effectiveness of Parent-Child Interaction Therapy at reducing disruptive externalising behaviours in young children. WoE B is based on the evidence hierarchies (Guyatt et al., 1995) which originated from evidence based medicine.

Table 15

Criteria of WoE B

Weighting	Criteria for WoE B
High (3)	<ul style="list-style-type: none"> - The study must have an `active control` group - Participants should be randomly assigned to a group with equivalences reported - Pre and post measures must be available for all groups

Weighting	Criteria for WoE B
	<ul style="list-style-type: none"> - Measures must report findings from more than one source - There should be two follow up assessments
Medium (2)	<ul style="list-style-type: none"> - The study must have an `no intervention` or waitlist` control group - Participants should be randomly assigned to a group with equivalences reported - Pre and post measures must be available for both groups - There should be one follow up assessment
Low (1)	<ul style="list-style-type: none"> - The study does not include a control group - Pre and post measures should be available for the intervention group
Zero (0)	<ul style="list-style-type: none"> - The study does not meet the criteria outlined above

Table 16

Scores for WoE B

Study	Overall score for WoE B
Bjørseth & Wichstrøm, 2016.	2 Medium
Bagner & Eyberg, 2007.	1 Low
Niec et al., 2016	2 Medium
Matos et al., 2009	2 Medium
Bagner et al., 2010	2 Medium

WoE C looks at the topic relevance of the study and how suitable it is for answering the review question (Gough, 2007).

Table 17

Criteria of WoE C

Weighting	Criteria for WoE
High (3)	<ul style="list-style-type: none"> - The study must use the original version of Parent-Child Interaction Therapy with no adaptations (other than translation) - The study provides evidence that the primary outcome measures have a high reliability of $r = .85$ or higher and measure children's externalising behaviours - The intervention must be delivered by someone who has been trained in Parent-Child Interaction Therapy and is following the manual - Demographics of the sample are documented (age, gender, ethnicity, socio-economic status)
Medium (2)	<ul style="list-style-type: none"> - The study must use Parent-Child Interaction Therapy - The study provides evidence that the primary outcome measures had a reliability of $r = .70$ or higher and must measure children's externalising behaviour - The intervention must be delivered by someone who is trained in Parent-Child Interaction Therapy or is following the manual - Some demographics of the sample are documented
Low (1)	<ul style="list-style-type: none"> - The study uses an adapted version of Parent-Child Interaction Therapy - The study provides evidence that the primary outcome measures had a reliability of $r = .50$ and must use outcome measures which measure children's externalising behaviour - It isn't clear if the person delivering the intervention has received training or is following the manual
Zero (0)	<ul style="list-style-type: none"> - The study does not meet any of the criteria outlined above

Table 18

Scores for WoE C

Study	Overall Score for WoE C
Bjørseth & Wichstrøm, 2016.	3 High
Bagner & Eyberg, 2007.	2 Medium
Niec et al., 2016	1 Low
Matos et al., 2009	1 Low
Bagner et al., 2010	2 Medium

Table 19

Overall WoE ratings (WoE D)

Study	WoE A	WoE B	WoE C	WoE D
Bjørseth & Wichstrøm, 2016.	2	2	3	2.33 Medium
Bagner & Eyberg, 2007.	1.5	1	2	1.5 Medium
Niec et al., 2016	2	2	1	1.66 Medium

Study	WoE A	WoE B	WoE C	WoE D
Matos et al., 2009	1.75	2	1	1.58 Medium
Bagner et al., 2010	2	2	2	2 Medium

Weak 0 – 1.4 Medium 1.5 – 2.4 Strong 2.5 – 3