Case Study 1: An Evidence-based Practice Review Report

Theme: Interventions implemented by parents.

How effective are evidence based behavioural parent training programmes for reducing challenging behaviour in preschool children? Efficacy, variability and moderating factors

Section 1: Summary

Challenging behaviour in early childhood has been linked to significant long-term negative effects (Cheng et al., 2007). Behavioural Parent Training (BPT) teaches parents about behaviour management skills with the aims of reducing challenging behaviour and improving the parent-child relationship. BPT is used in the UK as a first line intervention for challenging behaviour in preschool children (NICE, 2013). Consequently, educational psychologists may work with parents who have participated in or will participate in such interventions.

This systematic literature review explored the efficacy of evidence based BPTs for reducing challenging behaviour in young children. A literature search identified six studies that were critically appraised using Gough's (2007) Weight of Evidence Framework. Five studies were given medium ratings and one study was given a low rating. A mixture of medium and small effect sizes were found, and non-significant effects were found for two studies.

Overall, findings indicate that BPTs do reduce challenging behaviour in preschool children in the short-term, but there is little follow up evidence. Variable factors such as attrition and benefits of the interventions are also discussed, as well as limitations and recommendations for future research.

Section 2: Introduction

Preschool child behaviour and development

Higher incidences of challenging behaviour in childhood are linked to life-long negative impacts in numerous areas, including peer relationships, school, job stability and success in later life (Cheng et al., 2007; McCulloch et al., 2000). Disruptive behaviour is one of the most common reasons for children to be referred to mental health support (Baumel et al., 2016). Disruptive, challenging and externalising behaviour are often synonymous and can include defiance, aggression, hyperactivity and impulsivity (Tarver et al., 2014).

There are numerous meta-analyses that focus on challenging behaviour in children with autism (Deb et al., 2020; Tarver et al., 2019), traumatic brain injury (Brown et al., 2013; Shen et al., 2022) and other developmental conditions (Skotarczak & Lee, 2015; Tellegen & Sanders, 2013). There is significantly less research focusing on young children showing externalising symptoms that may be associated with Attention Deficit (Hyperactivity) Disorder (ADHD), Oppositional Defiance Disorder (ODD) and Conduct Disorder (CD) (Tarver et al., 2014). BPTs are often considered to be the first line of treatment for significant challenging behaviour (Baumel et al., 2016; Veen-Mulders et al., 2017).

Behavioural parent training programmes

BPTs focus on changing parenting behaviour to create change in child behaviour (Forehand et al., 2013). They combine social learning theory, behavioural theories and attachment theory to create interventions focused on improving the quality of the parent-child relationship (Reitman & McMahon, 2013). BPTs are designed to enhance parenting skills in order to decrease challenging behaviour (Veen-Mulders et al., 2017). Parents are first taught differential reinforcement, to give attention to positive behaviour and ignore negative behaviour (Reitman & McMahon, 2013). They are then taught to give specific instructions, praise compliance, and use time-out for noncompliance (Hembre-Kigin & McNeil, 2013).

BPTs can differ significantly in length and intensity (Tucker et al., 1997). The evidenced-based BPTs included in this review are The Incredible Years (IY), Parent Child Interaction Therapy (PCIT), and The Triple P-Positive Parenting Programme (Triple P). All the programmes are highly standardised and include protocol for facilitator training, delivery and supervision (Brinkmeyer & Eyberg, 2003; Sanders, 2008; Webster-Stratton, 2001). Programmes encourage parents to practice at home with their children as homework (Brinkmeyer & Eyberg, 2003; Sanders, 2008; Webster-Stratton, 2001). In the UK, the National Institute for Health and Care Excellent (NICE) recommends individual and group BPTs for children aged three plus who are at risk of developing or have a diagnosis of AD(H)D, ODD, or CD. (2013). Programme descriptions are included in Appendix A.

Programme Specific Evidence

There are many meta-analyses focusing on IY. Several reviews have found IY improves challenging behaviour, but do not look at the preschool age specifically (Gardener & Leitjen, 2017; Menting et al., 2013). Gardner and Leijten (2017) argue the variation found between individual family and child outcomes suggest further research is needed to understand long-term effects of the programme, and how to improve programme accessibility. An independent review found that IY is effective in the short-term, but that there is a lack of long-term effectiveness evidence for different demographics and ages, and particularly the preschool age range (Pidano & Allen, 2015).

A large systematic review of 101 studies over 33 years concluded Triple P has significant short- and long-term effects on the behaviour of children aged between two and nine (Sanders et al., 2014). However, it did not provide data of specific ages. A smaller review comparing PCIT and Triple P found both programmes to be effective (Thomas & Zimmer-Gembeck, 2007). No meta-analysis specifically on Triple P for young children with challenging behaviour could be found, but Triple P was included in meta-analyses that look at multiple treatments for externalising child behaviour. Triple P was found to be effective, but both reviews concluded that there was a lack of available long-term data (Dretzke et al., 2009; Lundahl et al., 2006).

Similarly, several meta-analyses have found PCIT to be effective when looking at a wider age range (Philips & Mychailyszyn, 2022; Aguayo et al., 2021) but that there have been limitations around understanding of follow up data (Aguayo et al., 2021). One analysis that included 12 studies found

that PCIT was very effective for the target population, but studies lacked diversity in demographic (Ward et al., 2016).

There is significant evidence that IY, Triple P, and PCIT individually are effective interventions for challenging behaviour in children (Tully & Hunt, 2016). There is a lack of meta-analyses focusing on young children and the interventions Triple P and IY (Pidano & Allen, 2015). There are also questions around the long-term efficacy of all three programmes (Aguayo et al., 2013). Furthermore, despite a significant number of studies reviewing the programmes, there is still much variability in the outcome for individual families, and attrition rates are not well understood (Menting et al., 2013).

As all three BPTs are derived from the same theories, it is of interest to consider their efficacy as a whole, as it is likely more interventions will continue to be created and given on this same basis. It is also of interest to consider if differences between the programmes affect their efficacy, for example in hours of parent time the programme takes, or additional supporting benefits of the programmes.

Rationale and Relevance to EP practice

EPs in the UK often work with preschool aged children who are experiencing behavioural difficulties. As PCIT, IY, and Triple P are all used in the UK and recommended by NICE (2013) for children aged three and upwards, it is of importance that EPs understand these programmes and their effectiveness. It may help them to work more effectively with families who have already gone through the intervention, are currently, or are considering it. It also may be of interest to consider if elements of each programme are more suited to the needs of a particular family. As there is a

national shortage of EPs and parents and schools are potentially facing a long wait for EP input, it is of particular importance that families are able to access other services and EPs are able to work with these other services to support them as much as possible (Squires et al.,, 2007).

Review question

How effective are evidence based behavioural parent training programmes for reducing challenging behaviour in preschool aged children?

Section 3: Critical Review of the Evidence

Literature Search

Three electronic databases (Eric (EBSCO), PsycINFO, and Web of Science) were used for the literature search which was undertaken in January 2023.

Table 1 shows the search terms used. The initial search also included date limit of the last 10 years. The initial search produced 877 results.

Table 1

Database Search Terms

Search	Database Search Terms
1	'parent-child interaction therapy' OR 'PCIT' OR 'the incredible years' OR 'Triple P' OR 'Positive parenting program*' OR 'Oregon social learning' OR 'behavio*ral parent* intervention*' OR 'BPT' OR 'Parent Management Training - Oregon Model' OR 'parent management training'

2	'conduct disorder*' OR 'conduct problem*' OR 'behavio* disorder*' OR 'behavio* problem*' OR 'aggressive behavio*' OR 'emotional behavio* problem*' OR 'child* problem* behavio*r*' OR 'externalizing disorder*' OR 'externalizing problem*' OR 'disruptive behavio*r*' OR 'disruptive behavio*r* disorder*' OR 'Disruptive behavio*r* problem*' OR 'Oppositional Defiant Disorder' OR 'challenging behavio*r'
3	'young child' or 'toddler' or 'preschool*' or 'pre-school'
4	'double-blind' OR 'random* control trial' OR 'RCT' OR 'Group comparison'

*Note. * was used to account for multiple possible endings to a key word. 'AND' was used to combine search terms, and 'OR' was used to separate alternative words for the same concept.*

Figure 1 includes a flowchart illustrating the literature search process and how many studies were excluded at each stage. The titles and abstracts of 861 studies were reviewed using the inclusion and exclusion criteria shown in Table 2. Afterwards, the full text of 24 articles was assessed using the same criteria. Appendix 1 shows the studies that were excluded after at the full text review stage, including the reason for exclusion.

After completing the full text review, six studies met the inclusion criteria and can be seen in Table 3.

Table 2

Inclusion and Exclusion Criteria

Criteria	Inclusion Criteria	Exclusion Criteria	Rationale
Publication Date	Published after 2013.	Published before 2013.	Older research may use outdated versions of the interventions, or be less relevant to current practice.
Study Design	Randomised Control Trials (RCTs)	Not RCTs.	RCTs are the gold standard for effectiveness (Petticrew & Roberts, 2006).
Intervention	IY, Triple P, PCIT, Parent Management Training	Any other intervention.	The four interventions studied are behavioural parent trainings with a very strong evidence base and are most generalisable to use in the UK.
Child Age	Child age between three and six or over 50% of children included between three and six.	<50% of children included in study between three and six.	This review focuses on preschool age children.

Criteria	Inclusion Criteria	Exclusion Criteria	Rationale
Country	OECD country.	Non-OECD country.	OECD countries are more generalisable to the UK.
Study Language	English.	Any other language.	The reviewer can only read studies in English.
Outcome Measure	Measurement of challenging, defiant, aggressive or externalising behaviour.	Does not include a measurement of challenging, defiant, aggressive or externalising behaviour.	This review focuses on how the interventions support challenging behaviour.
Diagnosis	Children with a diagnosis of ADHD, behaviour disorder, or no diagnosis.	Studies that include children with developmental disabilities, acquired brain injuries, Autism Spectrum Disorder.	There are numerous studies focusing on these interventions for the excluded populations.
Pharmacological Intervention	Children not taking medication for hyperactivity or behaviour, or studies that measure and define strictly how participants take medication.	Studies that do not consider if children are taking medication for hyperactivity or behaviour or do not analyse this.	Medications can interfere with the results of the interventions.
Study data	The study data has only been used in one study included in the review.	The same study data has been used in two or more studies included in this review. When this happens, the original study that collected the data will be used.	Two or more studies with the same data will bias the literature review results towards their findings.

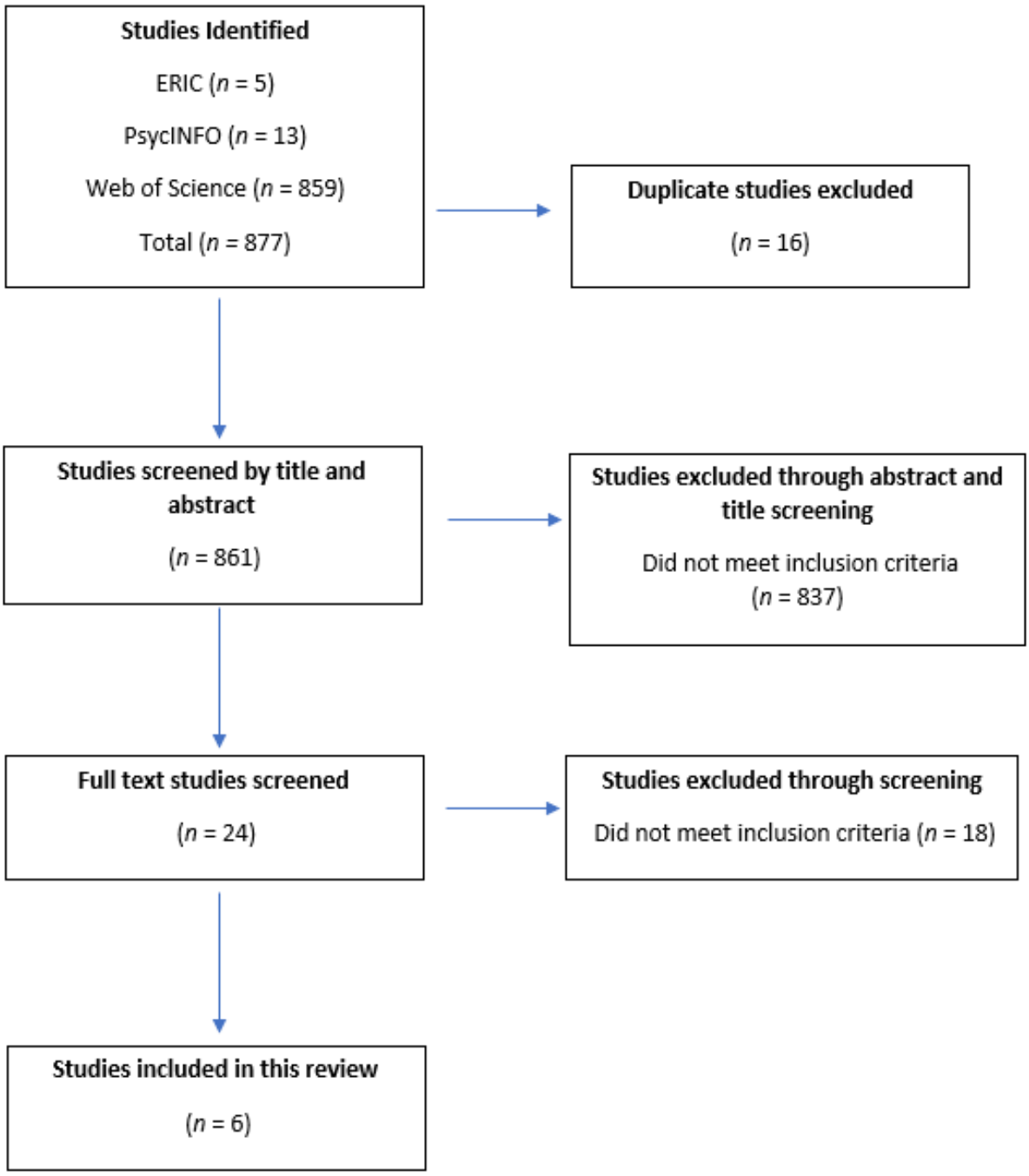
Table 3

Final Six Studies

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. *PloS One*, *11*(9), e0159845.
2. Franke, N., Keown, L. J., & Sanders, M. R. (2020). An RCT of an online parenting program for parents of preschool-aged children with ADHD symptoms. *Journal of Attention Disorders*, *24*(12), 1716-1726
3. Leckey, Y., McGilloway, S., Hickey, G., Bracken-Scally, M., Kelly, P., & Furlong, M. (2019). A randomised control trial of parent and child training programmes (versus wait list control) for children with ADHD-type behaviours: A pilot study. *Child Care in Practice*, *25*(4), 419-438
4. Healey, D., & Healey, M. (2019). Randomized controlled trial comparing the effectiveness of structured-play (ENGAGE) and behavior management (TRIPLE P) in reducing problem behaviors in preschoolers. *Scientific Reports*, *9*(1), 1-9.
5. Seabra-Santos, M. J., Gaspar, M. F., Azevedo, A. F., Homem, T. C., Guerra, J., Martins, V., & Moura-Ramos, M. (2016). Incredible Years parent training: What changes, for whom, how, for how long?. *Journal of Applied Developmental Psychology*, *44*, 93-104
6. Sonuga-Barke, E. J., Barton, J., Daley, D., Hutchings, J., Maishman, T., Raftery, J., & Thompson, M. J. (2018). A comparison of the clinical effectiveness and cost of specialised individually delivered parent training for preschool attention-deficit/hyperactivity disorder and a generic, group-based programme: a multi-centre, randomised controlled trial of the New Forest Parenting Programme versus Incredible Years. *European Child & Adolescent Psychiatry*, *27*(6), 797-809.

Figure 1

Flowchart of database search of studies



Mapping the Field

The studies chosen all included measuring challenging behaviour. The BPT used varied, as did the purpose, outcome measures, size of study, and location of study. The key features of each study are listed in Appendix B.

Weight of Evidence

Gough's (2007) Weight of Evidence (WoE) Framework was used to evaluate the quality and relevance of the six studies. The three parts comprise of methodological quality (WoE A), methodological relevance (WoE B), and topic relevance (WoE C).

Methodological quality (WoE A) was completed using the coding protocol devised by Gersten et al. (2005), which was chosen as it is for group experimental designs.

In order to assess methodological relevance (WoE B), the author developed a coding protocol with criteria specific to RCTs. The questions focus on randomisation, follow up measures, and control measures.

Topic relevance (WoE C) was assessed using a protocol designed by the author that focused on report data, location, intervention format, participant age, and study focus.

WoE A, B, and C were weighted equally and the average of these scores provided the overall WoE D. A summary of each study's WoE ratings is given in Table 4. Full criteria and coding protocol for WoE A, B and C is available in Appendix C.

Table 4

WoE A, B, C, and D

Study	WoE A	WoE B	WoE C	WoE D
Bjorseth & Wichstrom, (2016)	3	2.33	1.6	2.31 (medium)
Franke et al., (2020)	2	1.67	1.8	1.82 (medium)
Healey & Healey, (2019)	1	3	2.6	2.2 (medium)
Leckey et al., (2019)	2	2.67	1.8	2.16 (medium)
Seabra-Santos et al., (2016)	1	1.33	1.6	1.31 (low)
Sonuga-Barke et al., (2018)	2	2.33	2.6	2.31 (medium)

Note. Low rating = 1-1.6, medium rating = 1.7-2.3, high rating = 2.4-3

Participants

The total participants across the six studies was 669. Two studies focused on children aged between 3 and 4, (Franke et al., 2020; Healey & Healey, 2020), one study on children aged between two and seven, (Bjorseth & Wichstrom, 2020), one study children aged between two and four, (Sonuga-Barke et al.,2014), one study children aged between three and six (Seabra-Santos et al., 2016) and one study children aged between three and seven (Leckey et al., 2019). Studies that focused on children aged between three and five were given a higher rating in WoE C because this is the usual age of preschool children in the UK. Participants were recruited from a wide variety of sources.

Studies were excluded if they were not from an OECD country, as this is more generalisable to the UK context. Two studies were from New Zealand

(Franke et al., 2020; Healey & Healey 2019), and one from each respectively of Norway (Bjorseth & Wichstrom, 2020), Ireland (Leckey et al., 2019), Portugal (Seabra-Santos et al., 2016), and the UK (Sonuga-Barke et al., 2018). Studies were given a higher rating in WoE C if the country spoke English, and given the highest rating being from the UK, as this is the focus of this review.

Three studies focused on the use of a BPT for ADHD/ADD behaviours in young children (Franke et al., 2020; Leckey et al., 2019; Sonuga-Barke et al., 2018). Two studies focus specifically on behaviour problems (Bjorseth & Winchstrom, 2016; Healey & Healey, 2019) and one study considers the effectiveness of the intervention generally but measures challenging behaviour (Seabra-Santos et al., 2016). As the focus of this review is challenging behaviour, studies were given the highest rating In WoE C, if they focused on this.

Most studies reported an even split of demographic information across treatment arms (Bjorseth & Wichstrom, 2016; Healey & Healey, 2016; Leckey et al., 2019; Seabra-Santos et al., 2016; Sonuga-Barke et al., 2018). One study did not include treatment arm demographic information (Franke et al., 2020). No study focused on demographic differences.

In line with inclusion criteria, one study included children who were taking psychotropic medication, but carefully controlled and recorded results in analysis and therefore not excluded (Bjorseth & Wichstrom, 2016).

Table 5

Extra information about the variability of the interventions

Study	Intervention	Hours of parent training	Participants and providers	Attrition	Additional benefits and support	Relevant outcome measures
Bjorseth & Wichstrom (2020)	PCIT	Average 21.4	1 or 2 therapists to 1 parent	Pre-test to 6 months: 28.4% 18 months: 19.8%	No information included	ECBI intensity CBCL externalising Child non-compliance
Franke et al. (2020)	TPOL	8 online modules completed at own pace, access to 16 weeks	Training online autonomous, 1:1 for phone calls	45% of participants did not complete the programme	Audio-visual, interactive, two telephone consultations with facilitator.	Conners EC-BEH Defiance
Healey & Healey (2019)	Triple P	10	2 therapists, to 4-6 families	Not included	Childcare as children taught games by graduate students while parents trained.	BASC- 2 aggression parent & teacher
Leckey et al. (2019)	IY	28	2 therapists, 9-12 parents	Not included	DVDs, weekly phone calls with facilitators, free transportation, childcare and refreshments.	SDQ Conduct Conners CPRS-SF

Study	Intervention	Hours of parent training	Participants and providers	Attrition	Additional benefits and support	Relevant outcome measures
Seabra-Santos et al. (2016)	IY	28	2 therapists, 8-12 parents	8% at 6 months	Childcare and snacks provided.	PKBS-2 external
Sonuga-Barke et al., (2018)	IY	24-36	2 therapists, group size not clear	Not included	CDs, books and gifts AND parents received weekly phone calls from therapists. Where possible if missed a session received a home visit.	ECBI intensity ECBI problem SNAP-IV parent SNAP-IV teacher

Research Design

All studies included in this review are RCTs; the gold standard and therefore provide the clearest evidence for analysis of efficacy (Petticrew & Roberts, 2006). However, some distinctions can be made between the quality of the studies.

To be replicable and reliable, RCTs should include clear randomisation process information. Therefore, in WoE B studies were given the highest ranking for highly detailed randomisation information (Healey & Healey, 2016; Leckey et al., 2019; Sonuga-Barke et al., 2018) and the lowest for limited information (Franke et al., 2020; Seabra-Santos et al., 2016).

A prolonged follow up period can provide further information on the efficacy and long-term effects of an intervention, and studies were therefore ranked in WoE B on this. Studies were given the highest ranking if they included a follow up longer than six months that included data useable in this review (Bjorseth & Wichstrom, 2016; Healey & Healey, 2019).

A key part of the rigour of RCT studies is the use of control. All studies in this review had a control group, but studies that used a treatment as usual (TAU) active control were given a higher rating because an active control helps to control the possibility of placebo effect from the intervention (Bkorseth & Wichstrom, 2016; Franke et al., 2020; Seabra-Santos et al., 2016). The highest rating was given to the studies that had an active wait-list control (Healey & Healey, 2019; Leckey et al., 2019; Sonuga-Barke et al., 2018).

Intervention Content and Delivery

While all the interventions are BPTs, there is variation in content and delivery. The hours of each intervention, additional parent benefits, and outcomes measures used, are included table 5.

Interventions involved include clear protocols which were referenced by all studies. A variety of evidence based BPT interventions were used across the studies. Three used IY, (Leckey et al., 2019; Seabra-Santos et al., 2016; Sonuga-Barke, 2018), one used PCIT (Bjorseth & Wichstrom., 2016), one used standard Triple P, (Healey & Healey, 2019), one online Triple P (TPOL) (Franke et al., 2020).

Several of the studies interventions took place in community or locally based settings (Leckey et al., 2019; Seabra-Santos et al., 2016; Sonuga-Barke et al., 2018) one study took place in mental health clinic (Bjorseth & Wichstrom, 2016), TPOL took place from home (Franke et al., 2020), and one study took place at a research centre (Healey & Healey, 2019). There is no evidence indicating families had to travel significant distances to access interventions.

All studies included facilitators or therapists who had significant relevant experience and training, apart from Sonuga-Barke et al., (2018) who trained all therapists specifically for the programme, and the Franke et al., (2020) main intervention was online, although phone calls for this programme were made by experienced facilitators. This is also part of the protocol of all of the interventions.

Several of the factors recorded in Table 5 are thought to be of particular pertinence to this study because they may demonstrate or have

affected the utility of the sessions to the parents. For example, a study that shows high attrition may indicate that the intervention did not remain a priority for the parents. Studies were given a higher rating in WoE A if they included attrition data and if overall attrition was below 30%.

IY protocol states that parents must be offered childcare and where necessary transportation support to access the programme. This is most likely why IY is the only intervention to include such benefits. Two of the IY studies did not include data on attrition, and so it is not possible to see if there is correlation between these advantages and study attrition (Leckey et al., 2019; Sonuga-Barke et al., 2018).

Fidelity of Intervention

IY, PCIT and Triple P all include strict rules and protocols for the facilitators to follow, and therefore by following this they should have fair or high fidelity. Two studies used video-taping to ensure fidelity (Bjorseth & Wichstrom, 2016; Sonuga-Barke et al., 2018) and were rated higher in WoE A.

Two studies used a modified version of the intervention (Franke et al., 2020; Leckey et al., 2019). Studies that used a protocol altered version of the intervention received a lower rating, and the lowest went to the online version of Triple-P (Franke et al., 2020) as there is a smaller evidence base than group Triple P.

Measures

All studies used an array of standardised outcome measures that were considered reliable, and relevant outcome measures can be viewed in table 5. More information about the purpose and reliability of the measures can be

found in appendix E. Studies were excluded in initial screening if they did not include a pre and post measure of externalising behaviour, or specifically defiance or aggression. In WoE C, the measures relevant to this review were considered (available in Table 6). Studies were given the lowest rating if they only used parent report measures relevant to this review, as such measures are the most likely to be biased (Ringoot et al., 2015). Studies that used two types of report from parent/observer/teacher were given a medium rating (Healey & Healey, 2019; Sonuga-Barke et al., 2018) and no studies used all types of report for measures relevant to this review.

Table 6

Effect sizes and outcomes

Study	Outcome measure and significance	Effect size immediate post test	Effect size 6 months	Effect size 12 months +	Effect size descriptor	WoE D
Bjorseth & Wichstrom (2020)	ECBI		0.64 (medium)	0.52 (18 months; medium)	Within group, effect of time	2.31
	CBCL externalising		0.52* (medium)	0.22* (18 months; small)		
Franke et al. (2020)	Conners EC-BEH defiance/ Aggression	0.45 (medium)	0.09 (small)		Between group, effect of time	1.82
Healey & Healey (2019)	BASC-2 aggression	-1.96 (large)	-1.79 (large)	-1.79 (large)	Within group, effect of time	2.2
	Parent BASC-2 Aggression	-0.96 (large)	-0.65 (medium)	-0.49 (medium)		
	Teacher					

Study	Outcome measure and significance	Effect size immediate post test	Effect size 6 months	Effect size 12 months +	Effect size descriptor	Wo E D
Leckey et al. (2019)	SDQ – conduct problems		-0.33* (small)		Between group, effect of time	2.16
	Conners CPRS-FS Oppositional		-0.32* (small)			
Seabra-Santos et al. (2016)	PKBS-2 EXTernal	0.53		.	within group, effect of time x condition	1.31
	SDQ Conduct	0.49 (0.65) (medium)				
Sonuga-Barke et al., (2018)	ECBI intensity	-0.01* (no effect)			Between group, effect of time	2.31
	ECBI problem	-0.16* (small)				
	SNAP-IV parent	-0.09* (no effect)				
	SNAP-IV teacher	0.2 * (small)				

*Notes: * results are non-significant, Effect size descriptors: small = .2 or below, medium = .5, large = .8 or above*.

Findings and Effect sizes

All studies reported effect sizes (Table 6). Three studies reported Cohen's d (Bjorseth & WIchstrom, 2016; Franke et al., 2020; Leckey et al., 2019). For Healey & Healey (2019), Seabra-Santos et al., (2016) and Sonuga-Barke et al., (2018) effect sizes were calculated using mean and

standard deviation. Table 6 shows effect sizes with other relevant characteristics of the studies. Effect sizes were converted to Cohen's d for ease of comparison using the Campbell Collaboration effect size calculator (Wilson, 2020) and using Psychometrica calculator to convert eta squared (Lenard & Lenard, 2016). Only measures relevant to challenging behaviour were included in Table 6. The descriptions of all relevant measures can be found in Appendix E.

Effect sizes are an indication of the magnitude of an interventions impact (Ladkens, 2013). Most of the studies found medium effect sizes for at least one measure (Bjorseth & Wichstrom, 2020; Franke et al., 2020; Healey & Healey; 2019; Seabra-Santos et al., 2016)). Several studies reported negative effect sizes as the measure they use indicated a lower result when the child showed less externalising behaviour. All studies included a measure that showed that the child's challenging behaviour decreased as a result of the intervention.

One study found a very large effect size through one measure (Healey & Healey, 2019), and this study had a medium WoE D weighting (2.2). However, these results should be interpreted with caution as the sample size was small and the intervention and control samples were of unequal sizes.

One study's effect sizes could only be calculated for the immediate post-test data, despite having a long follow up period (Seabra-Santos, 2016). This was because the study did not provide clear data for each time period. This contributed to the low WoE B rating. Although a medium effect size was found, this study had the lowest WoE D rating overall and results should therefore be treated with caution.

Three studies (Bjorseth & Wichstrom, 2016; Leckey et al., 2019; Sonuga-Barke et al., 2018) reported insignificant findings for the outcome measure of interest. This means that although effect sizes are included, the results and effect sizes of some or all of the measures of these studies cannot be given weight.

Bjorseth & Wichstrom (2016) included two relevant outcome measures, and one found significant results with a medium effect size. This study has the highest WoE D but it was still within the medium weighting range. The study also includes the longest follow up data, at 18 months, and is therefore helpful in filling in the gaps in knowledge in this area. The findings of this study should therefore be given some weight, and suggests that BPTs are effective in reducing challenging behaviour, while the effect decreases from medium to small with time.

Healey and Healey (2019) and Sonuga-Barke et al., (2018) included relevant teacher rated outcomes as well as parent outcomes. Healey and Healey (2019) found that teachers rated the Triple P group as significantly improving during the follow up period, but they also reported significant improvements between the waitlist and baseline phases, and so the results do not clearly show that the improvement was as a result of the Triple P intervention. Sonuga-Barke et al.,s (2018) results were not significant. While this was a large RCT and was ranked highly in WoE D, the study suggests that a possible limitation of results may be the high attrition rate, which was marked down accordingly in WoE A. The results of this study therefore add caution to the positive interpretation of the other studies.

At the immediate post-test, all studies with significant results had medium or large effect sizes. This suggests that the results confirm what is already known, which is that BPTs have good short-term efficacy for challenging behaviour in young children. All studies that include 6-18 month follow up results show decreasing effect size as the test point is further from the intervention. However, they show that the benefits of the interventions were generally maintained.

Section 4: Conclusions and Recommendations

This review aimed to evaluate the effectiveness of evidence based behavioural parenting interventions for preschool aged children with challenging behaviour. It also aimed to consider how additional benefits of an intervention, or attrition in the studies may have correlated with the efficacy of an intervention. It examined this by reviewing RCTs of appropriate interventions that followed NICE (2013) guidelines.

Of the six studies reviewed, five received a medium WoE D rating and one received a low rating (Seabra-Santos et al., 2016). The similarity in findings between the studies is perhaps indicative of the generally high quality of RCT studies, and of the strict protocols of the interventions included.

Overall, this review supports previous studies finding that the BPTs are overall effective for preschool aged children. Several studies included follow-up effects, but only one study that can be given weight included effect sizes for appropriate measures and showed that effects were maintained long-term (Bjorseth & Wichstrom, 2016). This indicates that BPTs may be an

effective long-term treatment for challenging behaviour in young children, but more evidence is needed.

This size of this review may have limited its utility. In the future, a larger review could expand the limitations beyond just RCTs. This will then include a bigger pool of studies and more data that could be of use. For example, more data may lead to a better understanding of correlation between attrition and other factors. Additionally, a larger review may be able to provide more clarity on follow up data and the long-term efficacy of the BPT interventions.

Additionally, only one study in the current review was conducted in the UK (Sonuga-Barke et al., 2018) and so it would be beneficial to conduct a wider review with more UK studies in to be the most generalisable to this context.

No studies included in the search fitted the inclusion criteria for Parent Management Training, which is often considered the fourth evidence-based intervention from the same root of Social Learning Theory (Mabe & Turner, 2001).

Furthermore, it would be highly beneficial if further studies also included independent observer measures to minimise reporter bias. Only two of the studies here included relevant pre-school teacher report methods, one found insignificant results (Sonuga-Barke et al., 2018) and one had very unequal sample sizes (Healey & Healey, 2019). It is therefore not clear from this review if a teacher measure is helpful.

In terms of EP practice, an awareness and understanding of the usage and implications of evidence based BPTs for challenging behaviour is

important. EPs should be able to effectively work with parents who have been through a BPT programme of children throughout the age range, as the skills they learn in the programme may be used throughout childhood. As the shortage of EPs time is one of the biggest challenges facing the profession, interventions such as BPT may be used to support parents access of other avenues of support.

In summary, this review of BPTs did provide some promising evidence for the efficacy of the approach as a whole with preschool aged children with challenging behaviour. There is also tentative exploration of the long-term effects of these interventions with follow up results. However, because of the limitations of this review already discussed, it should be recognised that further review may be able to more solidly link the long-term effects of these interventions with the effects of attrition and other variables. In the future, it would be beneficial to complete further review if further relevant studies are published from a UK context, and particularly if they provide further long-term data on the efficacy of BPTs.

References

Azevedo, A. F., Seabra-Santos, M. J., Gaspar, M. F., & Homem, T. C. (2015). Do Portuguese Preschoolers With High Hyperactive Behaviors Make More Progress Than Those With Low Hyperactivity After Parental Intervention?. *Journal of Early Intervention*, *37*(2), 119-137. http://dx.doi.org/10.1177/1053815115598006

Azevedo, Andreia Fernandes, et al. "The incredible years basic parent training for Portuguese Preschoolers with AD/HD behaviors: Does it make a difference?." *Child & Youth Care Forum*. Vol. 42. No. 5. Springer US, 2013. http://dx.doi.org/10.1007/s10566-013-9207

Baker, S., Sanders, M. R., Turner, K. M., & Morawska, A. (2017). A randomized controlled trial evaluating a low-intensity interactive online parenting intervention, Triple P Online Brief, with parents of children with early onset conduct problems. *Behaviour Research and Therapy*, *91*, 78-90. http://dx.doi.org/10.1016/j.brat.2017.01.016

Barlow, J., Bergman, H., Kornør, H., Wei, Y., & Bennett, C. (2016). Group-based parent training programmes for improving emotional and behavioural adjustment in young children. *Cochrane Database of Systematic Reviews*, *8*. https://doi.org/10.1002/14651858.CD003680.pub3

Baumel, A., Pawar, A., Kane, J. M., & Correll, C. U. (2016). Digital Parent Training for Children with Disruptive Behaviors: Systematic Review and Meta-Analysis of Randomized Trials. *Journal of Child and Adolescent Psychopharmacology*, *26*(8), 740–749. https://doi.org/10.1089/cap.2016.0048

Bjorseth, A., & Wichstrom, 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). https://doi.org/10.1371/journal.pone.0159845

Brinkmeyer, M. Y., & Eyberg, S. M. (2003). Parent-child interaction therapy for oppositional children. In *Evidence-based psychotherapies for children and adolescents* (pp. 204–223). The Guilford Press. https://doi.org/10.1177/0145445508317134

Brown, F. L., Whittingham, K., Boyd, R., & Sofronoff, K. (2013). A Systematic Review of Parenting Interventions for Traumatic Brain Injury: Child and Parent Outcomes. *The Journal of Head Trauma Rehabilitation*, *28*(5), 349. https://doi.org/10.1097/HTR.0b013e318245fed5

Chan, S., Leung, C., & Sanders, M. (2016). A randomised controlled trial comparing the effects of directive and non-directive parenting programmes as a

universal prevention programme. *Journal of Children's Services*, *11*(1), 38-53. http://dx.doi.org/10.1108/JCS-08-2014-0038

Cheng, S., Kondo, N., Aoki, Y., Kitamura, Y., Takeda, Y., & Yamagata, Z. (2007). The effectiveness of early intervention and the factors related to child behavioural problems at age 2: A randomized controlled trial. *Early Human Development*, *83*(10), 683–691. https://doi.org/10.1016/j.earlhumdev.2007.01.008

Chung, S., Leung, C., & Sanders, M. (2015). The Triple P–Positive Parenting Programme: The effectiveness of group Triple P and brief parent discussion group in school settings in Hong Kong. *Journal of Children's Services*, *10*(4), 339-352. http://dx.doi.org/10.1108/JCS-08-2014-0039

Cova, F., Rincón, P., Bustos, C., Streiner, D., King, M., Saldivia, S., Inostroza, C., Grandon, P., & Novoa, C. (2020). Randomized cluster trial of a parenting program in Chile: key mediators in the decrease in behavior problems in preschool children. *Clinical child psychology and psychiatry*, *25*(2), 320-332. http://dx.doi.org/10.1177/1359104519864124

Deb, S. S., Retzer, A., Roy, M., Acharya, R., Limbu, B., & Roy, A. (2020). The effectiveness of parent training for children with autism spectrum disorder: a systematic review and meta-analyses. *BMC psychiatry*, *20*(1), 1-24.

de Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008). Effectiveness of The Triple P Positive Parenting Program on Behavioral Problems in Children: A Meta-Analysis. *Behavior Modification*, *32*(5), 714–735. https://doi.org/10.1177/0145445508317134

Dittman, C. K., Farruggia, S. P., Keown, L. J., & Sanders, M. R. (2016). Dealing with disobedience: An evaluation of a brief parenting intervention for young children showing noncompliant behavior problems. *Child Psychiatry & Human Development*, *47*(1), 102-112. http://dx.doi.org/10.1007/s10578-015-0548-9

Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., Taylor, R. S., Sandercock, J., & Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomised controlled trials. *Child and Adolescent Psychiatry and Mental Health*, *3*(1), 7. https://doi.org/10.1186/1753-2000-3-7

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 313–337. https://doi.org/10.1111/j.1469-7610.2006.01618.x

Forehand, R., Jones, D., & Parent, J. (2013). Behavioral parenting interventions for child disruptive behaviors and anxiety: What's different and what's the same. *Clinical Psychology Review*, *33*(1), 133–145.

https://doi.org/10.1016/j.cpr.2012.10.010

Franke, N., Keown, L., & Sanders, M. (2020). An RCT of an Online Parenting Program for Parents of Preschool-Aged Children With ADHD Symptoms. *Journal of Attention Disorders*, *24*(12), 1716–1726. https://doi.org/10.1177/1087054716667598

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.

Gardner, F., & Leijten, P. (2017). Incredible Years parenting interventions: Current effectiveness research and future directions. *Current opinion in psychology*, *15*, 99-104.

Gross, D., Belcher, H. M., Budhathoki, C., Ofonedu, M. E., Dutrow, D., Uveges, M. K., & Slade, E. (2019). Reducing preschool behavior problems in an urban mental health clinic: A pragmatic, non-inferiority trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(6), 572-581. http://dx.doi.org/10.1016/j.jaac.2018.08.013

Healey, D., & Healey, M. (2019). Randomized Controlled Trial comparing the effectiveness of structured-play (ENGAGE) and behavior management (TRIPLE P) in reducing problem behaviors in preschoolers. *Scientific reports*, *9*. https://doi.org/10.1038/s41598-019-40234-0

Heinrichs, N., Kliem, S., & Hahlweg, K. (2014). Four-year follow-up of a randomized controlled trial of Triple P group for parent and child outcomes. *Prevention Science*, *15*(2), 233-245. http://dx.doi.org/10.1007/s11121-012-0358-2

Hembree-Kigin, T. L., & McNeil, C. B. (2013). *Parent—Child Interaction Therapy*. Springer Science & Business Media.
http://dx.doi.org/10.1007/s10826-021-02073-4

Highlander, A., Quetsch, L., Girard, E., & McNeil, C. B. (2021). Preliminary Outcomes of an Incentive-based Parent-training Intervention. *Journal of Child and Family Studies*, *30*(11), 2845-2859
https://doi.org/10.1016/j.copsyc.2017.02.023

Jent, J. F., Rothenberg, W. A., Weinstein, A., Stokes, J., Barnett, M., Srivatsa, N., Dandes, E., & Garcia, D. (2021). Comparing traditional and Ebook-augmented Parent-Child Interaction Therapy (PCIT): a randomized control trial of pocket PCIT. *Behavior therapy*, *52*(6), 1311-1324.
http://dx.doi.org/10.1016/j.beth.2021.02.013

Jimerson, S. R., Oakland, T. D., & Farrell, P. T. (2006). *The Handbook of International School Psychology*. SAGE Publications.
http://dx.doi.org/10.4135/9781412976138

Kohlhoff, J., Cibralic, S., Wallace, N., Morgan, S., McMahon, C., Hawkins, E. Eapen, V., Briggs, N., Huber, A., & McNeil, C. (2020). A randomized controlled trial comparing parent child interaction therapy-toddler, circle of

security–parenting™ and waitlist controls in the treatment of disruptive behaviors for children aged 14–24 months: study protocol. *BMC psychology*, *8*(1), 1-14. http://dx.doi.org/10.1186/s40359-020-00457-7

Leckey, Y., McGilloway, S., Hickey, G., Bracken-Scally, M., Kelly, P., & Furlong, M. (2019). A Randomised Control Trial of Parent and Child Training Programmes (Versus Wait List Control) for Children with ADHD-Type Behaviours: A Pilot Study. *Child Care in Practice*, *25*(4), 419–438. http://dx.doi.org/10.1080/13575279.2018.1481368

Leijten, P., Gardner, F., Landau, S., Harris, V., Mann, J., Hutchings, J., Beecham, J., Bonin, E.-M., & Scott, S. (2018). Research Review: Harnessing the power of individual participant data in a meta-analysis of the benefits and harms of the Incredible Years parenting program. *Journal of Child Psychology and Psychiatry*, *59*(2), 99–109. https://doi.org/10.1111/jcpp.12781

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review*, *26*(1), 86–104. https://doi.org/10.1016/j.cpr.2005.07.004

Mabe, P. A., Turner, M. K., & Josephson, A. M. (2001). Parent management training. *Child and adolescent psychiatric clinics of North America*, *10*(3), 451-464. http://dx.doi.org/10.1016/S1056-4993(18)30040-3

McCulloch, A., Wiggins, R. D., Joshi, H. E., & Sachdev, D. (2000). Internalising and externalising children's behaviour problems in Britain and the US:

Relationships to family resources1. *Children & Society*, *14*(5), 368–383. https://doi.org/10.1111/j.1099-0860.2000.tb00191.x

McGilloway, S., Mhaille, G. N., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., Comiskey., C & Donnelly, M. (2012). A parenting intervention for childhood behavioral problems: a randomized controlled trial in disadvantaged community-based settings. *Journal of consulting and clinical psychology*, *80*(1), 116. http://dx.doi.org/10.1037/a0026304

McMahon, R. J., & Forehand, R. L. (2019). Helping the Noncompliant Child. In J. L. Lebow, A. L. Chambers, & D. C. Breunlin (Eds.), *Encyclopedia of Couple and Family Therapy* (pp. 1359–1364). Springer International Publishing. https://doi.org/10.1007/978-3-319-49425-8_382

Menting, A., de Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *CLINICAL PSYCHOLOGY REVIEW*, *33*(8), 901–913. https://doi.org/10.1016/j.cpr.2013.07.006

Murrihy, R., Kidman, A., & Ollendick, T. (2010). *Clinical handbook of assessing and treating conduct problems in youth. New York:* Springer. http://dx.doi.org/10.1007/978-1-4419-6297-3

Niec, L., Barnett, M., Prewett, M., & Chatham, J. (2016). Group Parent-Child Interaction Therapy: A Randomized Control Trial for the Treatment of

Conduct Problems in Young Children. *Journal of Clinical and Consulting Psychology*, *84*(8), 682–698. https://doi.org/10.1037/a0040218

Overview | Antisocial behaviour and conduct disorders in children and young people: Recognition and management | Guidance | NICE. (2013, March 27). NICE. https://www.nice.org.uk/guidance/cg158

Perrin, E. C., Sheldrick, R. C., McMenamy, J. M., Henson, B. S., & Carter, A. S. (2014). Improving parenting skills for families of young children in pediatric settings: A randomized clinical trial. *JAMA pediatrics*, *168*(1), 16-24. http://dx.doi.org/10.1001/jamapediatrics.2013.2919

Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.

Phillips, S. T., & Mychailyszyn, M. P. (2022). Parent-Child Interaction Therapy for Preschool Aged Youth: A Meta-Analysis of Developmental Specificity. *Child & Youth Care Forum*. https://doi.org/10.1007/s10566-022-09694-w

Pidano, A. E., & Allen, A. R. (2015). The Incredible Years Series: A Review of the Independent Research Base. *Journal of Child and Family Studies*, *24*(7), 1898–1916. https://doi.org/10.1007/s10826-014-9991-7

Reitman, D., & McMahon, R. J. (2013). Constance “Connie” Hanf (1917–2002): The mentor and the model. *Cognitive and Behavioral Practice*, *20*(1), 106-116.

Sanders, M. R. (2008). Triple P-Positive Parenting Program as a public health approach to strengthening parenting. *Journal of Family Psychology*, *22*, 506–517. https://doi.org/10.1037/0893-3200.22.3.506

Sanders, M. R. (2012). Development, Evaluation, and Multinational Dissemination of the Triple P-Positive Parenting Program. *Annual Review of Clinical Psychology*, *8*(1), 345–379. https://doi.org/10.1146/annurev-clinpsy-032511-143104

Sanders, M. R., Baker, S., & Turner, K. M. T. (2012). A randomized controlled trial evaluating the efficacy of Triple P Online with parents of children with early-onset conduct problems. *Behaviour Research and Therapy*, *50*(11), 675–684. https://doi.org/10.1016/j.brat.2012.07.004

Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, *34*(4), 337–357. https://doi.org/10.1016/j.cpr.2014.04.003

Seabra-Santos, M. J., Gaspar, M. F., Azevedo, A. F., Homem, T. C., Guerra, J., Martins, V., Leitão, S., Pimentel, M., Almeida, M., & Moura-Ramos, M. (2016). Incredible Years parent training: What changes, for whom, how, for how long? *Journal of Applied Developmental Psychology*, *44*, 93–104. https://doi.org/10.1016/j.appdev.2016.04.004

Shen, J., Zhao, S., Horn, T., Benkart, R., Busch, T., Vrabec, A., & Taylor, H. G. (2023). Family matters: A systematic review and meta-analysis on the efficacy of family-oriented interventions for children with acquired brain injuries. *Clinical Psychology Review*, *99*, 102218. https://doi.org/10.1016/j.cpr.2022.102218

Skotarczak, L., & Lee, G. K. (2015). Effects of parent management training programs on disruptive behavior for children with a developmental disability: A meta-analysis. *Research in Developmental Disabilities*, *38*, 272–287. https://doi.org/10.1016/j.ridd.2014.12.004

Sonuga-Barke, E., Barton, J., Daley, D., Hutchings, J., Maishman, T., Raftery, J., Stanton, L., Laver-Bradbury, C., Chorozoglou, M., Coghill, D., Little, L., Ruddock, M., Radford, M., Yao, G., Lee, L., Gould, L., Shipway, L., Markomichali, P., McGuirk, J., Thompson, M. (2018). A comparison of the clinical effectiveness and cost of specialised individually delivered parent training for preschool attention-deficit/hyperactivity disorder and a generic, group-based programme: A multi-centre, randomised controlled trial of the New Forest Parenting Programme versus Incredible Years. *European Child & Adolescent Psychiatry* , *27*(6), 797–809. https://doi.org/10.1007/s00787-017-1054-3

Squires, G., Farrell, P., Woods, K., Lewis, S., Rooney, S., & O'Connor, M. (2007). Educational psychologists' contribution to the Every Child Matters agenda: The parents' view. *Educational Psychology in Practice*, *23*(4), 343-361.

Ringoot, A. P., Tiemeier, H., Jaddoe, V. W., So, P., Hofman, A., Verhulst, F. C., & Jansen, P. W. (2015). Parental depression and child well-being: young children's self-reports helped addressing biases in parent reports. *Journal of clinical epidemiology*, *68*(8), 928-938.

Tarver, J., Daley, D., Lockwood, J., & Sayal, K. (2014). Are self-directed parenting interventions sufficient for externalising behaviour problems in childhood? A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, *23*(12), 1123–1137. https://doi.org/10.1007/s00787-014-0556-5

Tarver, J., Palmer, M., Webb, S., Scott, S., Slonims, V., Simonoff, E., & Charman, T. (2019). Child and parent outcomes following parent interventions for child emotional and behavioral problems in autism spectrum disorders: A systematic review and meta-analysis. *Autism*, *23*(7), 1630–1644. https://doi.org/10.1177/1362361319830042

Tellegen, C. L., & Sanders, M. R. (2013). Stepping Stones Triple P-Positive Parenting Program for children with disability: A systematic review and meta-analysis. *Research in Developmental Disabilities*, *34*(5), 1556–1571. https://doi.org/10.1016/j.ridd.2013.01.022

Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral Outcomes of Parent-Child Interaction Therapy and Triple P—Positive Parenting Program: A Review and Meta-Analysis. *Journal of Abnormal Child Psychology*, *35*(3), 475–495. https://doi.org/10.1007/s10802-007-9104-9

Tucker, S., & Gross, D. (1997). Behavioral Parent Training: An Intervention Strategy for Guiding Parents of Young Children. *The Journal of Perinatal Education*, *6*(2), 35–43. https://www.proquest.com/docview/203577108/abstract/89DD314E27F49E4PQ/1

Tully, L., & Hunt, C. (2016). Brief Parenting Interventions for Children at Risk of Externalizing Behavior Problems: A Systematic Review. *Journal of Child and Family Studies, 25*(3), 705–719. https://doi.org/10.1007/s10826-015-0284-6

Valero-Aguayo, L., Rodríguez-Bocanegra, M., Ferro-García, R., & Ascanio-Velasco, L. (2021). Meta-analysis of the Efficacy and Effectiveness of Parent Child Interaction Therapy (PCIT) for Child Behaviour Problems. *Psicothema*, *33.4*, 544–555.https://doi.org/10.7334/psicothema2021.70i.org/10.7334/psicothema2021.70

van der Veen-Mulders, L., Hoekstra, P. J., Nauta, M. H., & van den Hoofdakker, B. J. (2018). Preschool children's response to behavioural parent training and

parental predictors of outcome in routine clinical care. *Clinical Psychology & Psychotherapy*, *25*(1), 1–9. https://doi.org/10.1002/cpp.2117

Ward, M. A., Theule, J., & Cheung, K. (2016). Parent–Child Interaction Therapy for Child Disruptive Behaviour Disorders: A Meta-analysis. *Child & Youth Care Forum*, *45*(5), 675–690. https://doi.org/10.1007/s10566-016-9350-5

Webster-Stratton, C. (2002). The incredible years: Parents and children videotape series: A parenting course (ADVANCE). *Seattle, WA: Incredible Years*. https://doi.org/10.1007/s10826-015-0284-6

Appendix A: The Incredible Years, Triple P and Parent Child Interaction Therapy

The Incredible Years parent training programme is a group-based programme where parents of young children watch video footage modelling parent child interaction in different situations (Webster-Stratton, 2001). Two trained facilitators lead the group to discuss the models and to use the techniques in role-play. In the basic IY preschool programme, topics include play, misbehaviour, praise and rewards (Menting et al., 2013; Webster-Stratton, 2001).

Triple P is an intervention designed to support behavioural, emotional and developmental problems in children aged zero to 16 (de Graaf et al., 2008). Triple P is multi-tiered, and level 4 is a group parent programme that meets NICE recommendations (2013). Parents learn a range of child management techniques and how to apply them to their own child (Sanders,

2012). There are various versions of Triple P including an online version that parents work through independently, which has a growing body of supportive literature (Baumel et al., 2018; Sanders et al., 2012).

Parent Child Interaction Therapy is an intervention where a therapist coaches a parent working with their child, often from behind a one-way mirror (Brinkmeyer & Eyberg, 2003; Hembree-Kigin & McNeil, 2013). Parents are coached through child-directed interaction, where parents learn play therapy skills, and parent-direct interaction, where parents learn skills relating to compliance and disruptive behaviour (Hembree-Kigin & McNeil, 2013). There is not a set length of treatment as parents only move forward when they are considered to have mastered a skill, mastery is assessed through a standardised coding system (Brinkmeyer & Eyberg, 2003).

Appendix B: Studies excluded at the full text review stage

Study reference	Exclusion Reason
1. Azevedo, A. F., Seabra-Santos, M. J., Gaspar, M. F., & Homem, T. C. (2015). Do Portuguese Preschoolers With High Hyperactive Behaviors Make More Progress Than Those With Low Hyperactivity After Parental Intervention?. *Journal of Early Intervention*, *37*(2), 119-137.	Same data as included study
2. Azevedo, Andreia Fernandes, et al. "The incredible years basic parent training for Portuguese Preschoolers with AD/HD behaviors: Does it make a difference?." *Child & Youth Care Forum*. Vol. 42. No. 5. Springer US, 2013.	Same data as included study
3. Baker, S., Sanders, M. R., Turner, K. M., & Morawska, A. (2017). A randomized controlled trial	Diagnosis exclusion criteria

Study reference	Exclusion Reason
evaluating a low-intensity interactive online parenting intervention, Triple P Online Brief, with parents of children with early onset conduct problems. *Behaviour Research and Therapy*, *91*, 78-90.	
4. Chan, S., Leung, C., & Sanders, M. (2016). A randomised controlled trial comparing the effects of directive and non-directive parenting programmes as a universal prevention programme. *Journal of Children's Services*, *11*(1), 38-53.	Pharmacological exclusion criteria
5. Chung, S., Leung, C., & Sanders, M. (2015). The Triple P–Positive Parenting Programme: The effectiveness of group Triple P and brief parent discussion group in school settings in Hong Kong. *Journal of Children's Services*, *10*(4), 339-352.	Pharmacological exclusion criteria
6. Cova, F., Rincón, P., Bustos, C., Streiner, D., King, M., Saldivia, S., ... & Novoa, C. (2020). Randomized cluster trial of a parenting program in Chile: key mediators in the decrease in behavior problems in preschool children. *Clinical child psychology and psychiatry*, *25*(2), 320-332.	Programme exclusion criteria
7. Dittman, C. K., Farruggia, S. P., Keown, L. J., & Sanders, M. R. (2016). Dealing with disobedience: An evaluation of a brief parenting intervention for young children showing noncompliant behavior problems. *Child Psychiatry & Human Development*, *47*(1), 102-112.	Programme exclusion criteria
8. Gross, D., Belcher, H. M., Budhathoki, C., Ofonedu, M. E., Dutrow, D., Uveges, M. K., & Slade, E. (2019). Reducing preschool behavior problems in an urban mental health clinic: A pragmatic, non-inferiority trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(6), 572-581.	Pharmacological exclusion criteria
9. Heinrichs, N., Kliem, S., & Hahlweg, K. (2014). Four-year follow-up of a randomized controlled trial of Triple P group for parent and child outcomes. *Prevention Science*, *15*(2), 233-245.	Diagnosis exclusion criteria
10. Highlander, A., Quetsch, L., Girard, E., & McNeil, C. B. (2021). Preliminary Outcomes of an	Diagnosis exclusion criteria

Study reference	Exclusion Reason
Incentive-based Parent-training Intervention. *Journal of Child and Family Studies*, *30*(11), 2845-2859.	
11. Kohlhoff, J., Cibralic, S., Wallace, N., Morgan, S., McMahon, C., Hawkins, E., Eapen, V., Briggs, N., Huber, A., & McNeil, C. (2020). A randomized controlled trial comparing parent child interaction therapy—Toddler, circle of security- parenting (TM) and waitlist controls in the treatment of disruptive behaviors for children aged 14-24 months: Study protocol. *BMC psychology*, *8*(1). https://doi.org/10.1186/s40359-020-00457-7	Age of children
12. Jent, J. F., Rothenberg, W. A., Weinstein, A., Stokes, J., Barnett, M., Srivatsa, N., ... & Garcia, D. (2021). Comparing traditional and Ebook-augmented Parent-Child Interaction Therapy (PCIT): a randomized control trial of pocket PCIT. *Behavior therapy*, *52*(6), 1311-1324.	Diagnosis exclusion criteria
13. Kohlhoff, J., Cibralic, S., Wallace, N., Morgan, S., McMahon, C., Hawkins, E., ... & McNeil, C. (2020). A randomized controlled trial comparing parent child interaction therapy-toddler, circle of security–parenting™ and waitlist controls in the treatment of disruptive behaviors for children aged 14–24 months: study protocol. *BMC psychology*, *8*(1), 1-14.	Diagnosis exclusion criteria
14. McGilloway, S., Mhaille, G. N., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., ... & Donnelly, M. (2012). A parenting intervention for childhood behavioral problems: a randomized controlled trial in disadvantaged community-based settings. *Journal of consulting and clinical psychology*, *80*(1), 116.	Same data as included study
15. 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.	Control group (PCIT) is the intervention relevant to this review, and therefore the study has no relevant control or TAU.

Study reference	Exclusion Reason
16. Perrin, E. C., Sheldrick, R. C., McMenamy, J. M., Henson, B. S., & Carter, A. S. (2014). Improving parenting skills for families of young children in pediatric settings: A randomized clinical trial. *JAMA pediatrics*, *168*(1), 16-24.	Pharmacological exclusion criteria
17. Sampaio, F., Sarkadi, A., Salari, R., Zethraeus, N., & Feldman, I. (2015). Cost and effects of a universal parenting programme delivered to parents of preschoolers. *The European Journal of Public Health*, *25*(6), 1035-1042.	Diagnosis exclusion criteria
18. Stattin, H., Enebrink, P., Özdemir, M., & Giannotta, F. (2015). A national evaluation of parenting programs in Sweden: The short-term effects using an RCT effectiveness design. *Journal of consulting and clinical psychology*, *83*(6), 1069.	Age exclusion criteria

Appendix C: Mapping the field

Characteristics of the six studies

Study	Aim	Location	Sample size	Intervention	Child age	Outcome	Follow up
Bjorseth & Wichstrom (2020)	To assess the effectiveness of PCIT in Norway for young children's behaviour problems.	Norway	81	PCIT	2-7	Overall children who had received intervention showed the biggest decrease in behaviour problems compared with TAU.	At follow up, parent report showed fewer behaviour challenges compared to TAU.
Franke et al. (2020)	Evaluate online - self help version of Triple P programme for parents of pre-schoolers with ADHD symptoms.	New Zealand	53	Triple-P online	3-4 years	Intervention group reported fewer ADHD symptoms immediately after intervention.	At the six month follow up some results were maintained but not changes in child behaviour.
Healey & Healey (2019)	To compare the effectiveness of ENGAGE and Triple P in reducing behaviour problems in preschool children.	New Zealand	60	Triple-P	3-4	ENGAGE and Triple P were both effective in reducing parent rated problem behaviours.	Information not included for comparing waitlist and follow up data.

Study	Aim	Location	Sample size	Intervention	Child age	Outcome	Follow up
Leckey et al. (2019)	To compare IY parent training versus NFPP parent training and child training programmes for ADHD behaviours.	Ireland	45	Incredible Years	3-7 years	Children in the PT + CT group did not fare better than those in the PT group only, both groups improved ADHD behaviours and social skills compared with TAU.	
Seabra-Santos et al. (2016)	Examine the effects of IY with Portuguese families of pre-schoolers.	Portugal	124	IY	3-6	Reduction in behaviour problems in the intervention group with a wide range of families.	Information not included for comparing waitlist and follow up data.
Sonuga-Barke et al., (2018)	Comparing the clinical effectiveness and cost of the New Forest Parenting Programme versus Incredible Years for ADHD behaviours.	UK	306	Incredible Years	Between 2 years and 9 months and 4 years and six months	No significant findings. NFPP cost considerably less.	Results not significant.

Appendix D – Criteria for Weight of evidence ratings with rationale.

Coding Protocol for Weight of Evidence A
Protocol: Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and
quasi experimental research in special education. Exceptional children, 71(2), 149-
164.

Study: Bjorseth & Wichstrom, 2016

Essential Quality Indicators

Quality Indicators for Describing Participants

1. Was sufficient information provided to determine/confirm whether the participants
demonstrated the disability(ies) or difficulties presented?

☒Yes ☐No/ Unknown

Rationale: Inclusion required referral to medical service for behaviour problems or ECBI score in 90th percentile of norms.

2. Were appropriate procedures used to increase the probability that teachers or
interventions were comparable across conditions?

☒Yes ☐No/ Unknown

Rationale: Therapists were clinical practitioners who either offered TAU or PCIT correspondingly.

3. Was sufficient information given characterising the interventions or teachers provided? Did it indicate whether they were comparable across conditions?

☒Yes ☐No/ Unknown

Rationale: years of experience given and specialism.

Quality Indicators for Implementation of the Intervention and Description of Comparison Conditions

1. Was the intervention clearly described and specified?

☒Yes ☐No/ Unknown

Rationale: Outlined intervention in full.

2. Was the fidelity of implementation described and assessed?

☒Yes ☐No/ Unknown

Rationale: Integrity checklists of co-therapists monitored each other in 45% of sessions.

3. Was the nature of services provided in comparison conditions described?

☒Yes ☐No/ Unknown

Rationale: Full description of alternate treatment.

Quality Indicators for Outcome Measures

1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?

☒Yes ☐No/ Unknown

Rationale: ECBI, Child Behaviour Checklist, Dyadic parent-child interaction coding system by clinicians.

2. Were outcomes for capturing the intervention's effect measured at the appropriate times?

☒Yes ☐No/ Unknown

Rationale: Measured pre, immediate post and 6, 12 and 18 months.

Quality Indicators for Data Analysis

1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

☒Yes ☐No/ Unknown

Rationale: Appropriate and explained techniques used.

2. Did the research report include not only inferential statistics but also effect size calculations?

☒Yes ☐No/ Unknown

Rationale: Effect size reported.

Desirable Quality Indicators

1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

☒Yes☐ No/ Unknown

Rationale: Attrition was listed and below 30%.

2. Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scores blind to study conditions and equally (un)familiar to examinees across study conditions?

☐Yes ☒No/ Unknown

Rationale: Not reported.

3. Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

☒Yes ☐No/ Unknown

Rationale: Up to 18 months.

4. Was evidence of the criterion-related validity and construct validity of the measures provided?

☒Yes ☐No/ Unknown

Rationale: References criterion-related validity and construct validity of measures.

5. Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

☒Yes ☐No/ Unknown

Rationale: Co-therapists examined quality.

6. Was any documentation of the nature of instruction or series provided in comparison conditions?

☒Yes ☐No/ Unknown

Rationale: checklists and video

7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

☒Yes ☐No/ Unknown

8. Were results presented in a clear, coherent fashion?

☒Yes ☐No/ Unknown

Study: Franke et al., 2016

Essential Quality Indicators

Quality Indicators for Describing Participants

1. Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?

☒Yes ☐No/ Unknown

Rationale: measures for ADHD behaviours including disruptive behaviour.

2. Were appropriate procedures used to increase the probability that teachers or interventions were comparable across conditions?

☒Yes ☐No/ Unknown

Rationale: Telephone consultations with Triple P trained facilitators

3. Was sufficient information given characterising the interventions or teachers provided? Did it indicate whether they were comparable across conditions?

☐Yes ☒No/ Unknown

Rationale: No evidence given

Quality Indicators for Implementation of the Intervention and Description of Comparison Conditions

1. Was the intervention clearly described and specified?

☒Yes ☐No/ Unknown

Rationale: Outlined intervention in full.

2. Was the fidelity of implementation described and assessed?

☒Yes ☐No/ Unknown

Rationale: 25% pf consultations checked on inter-observer agreement.

3. Was the nature of services provided in comparison conditions described?

☒Yes ☐No/ Unknown

Rationale: Delayed treatment waitlist.

Quality Indicators for Outcome Measures

1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?

☒Yes ☐No/ Unknown

Rationale: Teacher and parent report scales used.

2. Were outcomes for capturing the intervention's effect measured at the appropriate times?

☒Yes ☐No/ Unknown

Rationale: Measured pre, immediate post and 6 months.

Quality Indicators for Data Analysis

1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

☒Yes ☐No/ Unknown

Rationale: Appropriate and explained techniques used.

2. Did the research report include not only inferential statistics but also effect size calculations?

☒Yes ☐No/ Unknown

Rationale: Effect size reported.

Desirable Quality Indicators

1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

☐Yes ☒No/ Unknown

Rationale: High teacher attrition

2. Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scores blind to study conditions and equally (un)familiar to examinees across study conditions?

☒Yes ☐No/ Unknown

Rationale: internal consistency and test-retest reported.

3. Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

☒Yes ☐No/ Unknown

Rationale: 6 months.

4. Was evidence of the criterion-related validity and construct validity of the measures provided?

☐Yes ☒No/ Unknown

Rationale: References criterion-related validity and construct validity of measures.

5. Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

☒Yes ☐No/ Unknown

Rationale: Observers examined quality.

6. Was any documentation of the nature of instruction or series provided in comparison conditions?

☐Yes ☒No/ Unknown

7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

☐Yes ☒No/ Unknown

8. Were results presented in a clear, coherent fashion?

☒Yes ☐No/ Unknown

Table D1

WoE A Scores for Each Study

Study	Essential Quality Indicators	Desirable Quality indicators	WoE B
Bjorseth & Wichstrom, (2016)	9	8	3
Franke et al., (2020)	9	4	2
Healey & Healey, (2019)	8	2	1
Leckey et al., (2019)	10	2	2
Seabra-Santos et al., (2016)	8	5	1
Sonuga-Barke et al., (2018)	10	24	2

Table D2

WoE B Criteria

Criteria	Low (1)	Medium (2)	High (3)	Rationale
1. Randomisation	Limited information about randomisation process	Appropriate randomisation information	Highly detailed randomisation information	As all studies included are RCTs, it is crucially important that they include randomisation and that the process is clear.
2. Follow up	Only includes post-test outcomes	6 months outcomes	Further follow up outcomes	Follow up shows how the effectiveness of the intervention is maintained.
3. Control group	Wait list control only	Active comparison intervention control only	Active control and waitlist control.	Using an active control group helps to control for placebo effect. Studies that had a wait list control, the intervention therapy and another type of therapy scored the highest.

Table D3

WoE B Scores for Each Study

Study	Randomisation	Follow up	Control Group	WoE B
Bjorseth & Wichstrom, (2016)	2	3	2	2.33
Franke et al., (2020)	1	2	2	1.67
Healey & Healey, (2019)	3	3	3	3
Leckey et al., (2019)	3	2	3	2.67
	1	1	2	1.33
Seabra-Santos et al., (2016)	1	1	2	1.33
Sonuga-Barke et al., (2018)	3	2	3	2.31

Table D4

WoE C Criteria

Criteria	Low (1)	Medium (2)	High (3)	Rationale
1. Source of report data relevant to this study	Only parent report	Parent and observer/pre-school teacher report	Parent, observer and pre-school teacher report	Solely parent report data may contain the most bias, whereas independent observer data may be the least biased. Pre-school teacher data may provide a less biased review of parent and child progress.
2. Country of study	Non English speaking country	English speaking non UK country	UK	Studies from the UK are the most generalisable to UK context. The pre-school provision in English speaking countries is more relevant to the UK context as changes to the BPT programmes may have been made when used in other languages.
3. Format of intervention	A non-standard version	Protocol amended version	Standard version	Non-standard versions of the intervention have less evidence base

4. Child age	Less than half of the children are aged 3-5 or it's not possible to say	The majority of the children are aged 3-5	All children are aged 3-5	3-5 is the typical preschool age in the UK.
5. Focus of study	Does not focus on children presenting with challenging behaviour	Includes children with challenging behaviour as partial focus	Focuses on children presenting with challenging behaviour	This review focuses specifically on challenging behaviour.

Table D5

WoE C Scores for Each Study

Study	Report source	Country	Format	Child age	Study Focus	WoE C
Bjorseth & Wichstrom, (2016)	1	3	3	1	2	1.6
Franke et al., (2020)	1	2	2	3	2	1.8
Healey & Healey, (2019)	2	3	2	3	3	2.6
Leckey et al., (2019)	1	2	2	1	2	1.8
Niec et al., (2016)	1	2	3	2	3	2.2
Seabra-Santos et al., (2016)	1	3	2	2	1	1.6
Sonuga-Barke et al., (2018)	2	2	2	3	2	2.6

Appendix E: Outcome measures descriptions

Measures	Description	Cronbach's Alpha
Behaviour Assessment System for Children-Second Addition (BASC-2) (aggression)	Multidimensional assessment system that assesses adaptive and clinical parts of behaviour and emotional functioning.	.81
The Eyberg Child Behaviour Inventory (ECBI) (intensity and problem)	Assesses the frequency and severity of disruptive behaviours.	.86
The Child Behaviour Checklist (CBCL)	Measures a wide set of problems including externalising behaviour.	.84
The Conners Early Childhood Behaviour Scale (Conners EC-BEH) (defiance/aggression/oppositional)	Assesses young child behaviour.	.94
The Preschool and Kindergarten Behaviour Scale Second Addition (PKBS-2 external)	Assesses problem behaviours and social skills.	.9
The Conners Parent Rating Scale, 3rd edition, short form (CPRS-SF)	Assesses behaviour, social, and academic behaviour related issues. Often uses in the diagnosis of ADHD.	.89
The Strengths and Difficulties Questionnaire (SDQ) (conduct problems)	Behavioural screening questionnaire relating to specific attributes.	.62
The Swanson, Nolan, and Pelham Rating Scale SNAP-IV (parent and teacher)	Measures core symptoms of ADHD.	Parent 0.89 Teacher 0.93