Cochrane Library

Cochrane Database of Systematic Reviews

Personalised interventions for subgroups of children with conduct problems (Review)

Lane C, Hogg E, Karwatowska LA, French L, Ranieri VF, Jesnick LGD, Roberts C, Scott S, Senior R, Skinner GCM, Kennedy EMM

Lane C, Hogg E, Karwatowska LA, French L, Ranieri VF, Jesnick LG D, Roberts C, Scott S, Senior R, Skinner GCM, Kennedy EM M.
Personalised interventions for subgroups of children with conduct problems.
Cochrane Database of Systematic Reviews 2023, Issue 4. Art. No.: CD012746.
DOI: 10.1002/14651858.CD012746.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Personalised interventions for subgroups of children with conduct problems

Chloe Lane[1], Elizabeth Hogg[1], Lucy A Karwatowska[1], Lorna French[1], Veronica F Ranieri[2], Leah G D Jesnick[1], Christopher Roberts[3], Stephen Scott[4], Robert Senior[1], Guy CM Skinner[1,5], Eilis M M Kennedy[1,2]

[1]Tavistock Research Unit, Tavistock and Portman NHS Foundation Trust, London, UK. [2]Department of Clinical, Educational and Health Psychology, University College London, London, UK. [3]Division of Population Health, University of Manchester, Manchester, UK. [4]Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [5]Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

Contact: Eilis M M Kennedy, ekennedy@tavi-port.nhs.uk.

Editorial group: Cochrane Developmental, Psychosocial and Learning Problems Group.
Publication status and date: New, published in Issue 4, 2023.

Citation: Lane C, Hogg E, Karwatowska LA, French L, Ranieri VF, Jesnick LG D, Roberts C, Scott S, Senior R, Skinner GCM, Kennedy EM M. Personalised interventions for subgroups of children with conduct problems. *Cochrane Database of Systematic Reviews* 2023, Issue 4. Art. No.: CD012746. DOI: 10.1002/14651858.CD012746.pub2.

ABSTRACT

Background

Conduct problems are a range of disruptive behaviours in childhood that are associated with long-term adverse outcomes in adolescence and adulthood, including antisocial behaviour, substance misuse, and poor academic achievement. Children with conduct problems can vary according to age of onset, comorbidities, and environmental factors, and it has been suggested that certain groups of children may have different treatment outcomes. Therefore, it is important to assess the extent to which personalised interventions for different groups of children with conduct problems may affect outcomes. To our knowledge, this is the first review to systematically identify and appraise the effectiveness of personalised interventions, adapted, or developed, for prespecified subgroups of children with conduct problems.

Objectives

To assess whether personalised interventions, adapted or developed for subgroups of children with conduct problems are effective in improving outcomes.

Search methods

We used standard, extensive Cochrane search methods. The latest search was 1 February 2022.

Selection criteria

We included randomised controlled trials (RCTs), in any setting, in children (aged two to 12 years) with conduct problems and within a prespecified subgroup, comparing a personalised intervention with a non-personalised intervention, waitlist control, or treatment as usual. Personalised interventions included adaptations to standard practice, such as parent-training programmes; other recommended interventions for children with conduct problems; or interventions developed specifically to target subgroups of children with conduct problems. We excluded non-personalised and non-psychological interventions (e.g. pharmacological or dietary intervention). Prespecified subgroups of children with conduct problems, however defined, were eligible for inclusion.

Data collection and analysis

We used standard Cochrane methods. Our primary outcomes were 1. child conduct problems or disruptive behaviour and 2. adverse events. Our secondary outcomes were 3. personalised treatment outcomes relevant to each subgroup, 4. parenting skills and knowledge, 5. family functioning, engagement and decreased dropout, and 6. educational outcomes. We used GRADE to assess the certainty of the evidence.

Main results

We identified 13 RCTs (858 participants). Seven studies were conducted in the USA, five in Australia, and one in Germany. Eleven studies reported their source of funding, with five studies receiving grants from the National Institute of Mental Health. In total, 15 different funders supported the studies included in the review.

We separated subgroups of children with conduct problems into three broad categories: children with co-occurring conditions (e.g. emotional difficulties), parent characteristics (e.g. conflict between parents), or familial/environmental circumstances (e.g. rural families). All studies delivered a personalised intervention that was adapted or developed for a prespecified subgroup of children with conduct problems. We rated all trials at unclear or high risk of bias in most domains. Below, we report the results of improvement in child conduct problems and disruptive behaviour, personalised treatment outcomes, and parenting skills and knowledge for our main comparison: personalised versus non-personalised interventions.

Improvement in child conduct problems and disruptive behaviour

Compared with a non-personalised intervention, a personalised intervention may result in a slight improvement in child conduct problems or disruptive behaviour measured using the Eyberg Child Behavior Inventory (ECBI) Problem subscale in the short term (mean difference (MD) −3.04, 95% confidence interval (CI) −6.06 to −0.02; 6 studies, 278 participants; P = 0.05), but may have little to no effect on improving child conduct problems or disruptive behaviour measured by the ECBI Intensity subscale (MD −6.25, 95% CI −16.66 to 4.15; 6 studies, 278 participants; P = 0.24), or the Externalising subscale of the Child Behaviour Checklist (CBCL) (MD −2.19, 95% CI −6.97 to 2.59; 3 studies, 189 participants, P = 0.37) in the short term. We graded the certainty of evidence as very low for all three outcomes, meaning any estimate of effect is very uncertain.

Personalised treatment outcomes, relevant to each subgroup

Although six studies reported personalised treatment outcomes, relevant to each subgroup, we were unable to pool the data due to differences between the measures used in the studies and the heterogeneity this would produce in analysis. The results for this outcome were inconclusive.

Parenting skills and knowledge

Although seven studies reported parenting skills and knowledge, we were unable to pool the data due to differences between the measures used in the studies and the heterogeneity this would produce in analysis. The results for this outcome were inconclusive.

Adverse events

None of the trials reported monitoring adverse events.

Summary of results

In summary, there is limited evidence that personalised intervention improves child conduct problems, personalised treatment outcomes, relevant to each subgroup, or parenting skills and knowledge compared with a non-personalised intervention.

Authors' conclusions

There is limited evidence for the effectiveness of personalised interventions for subgroups of children with conduct problems. The certainty of evidence for all outcomes was very low, meaning that we have very little confidence in the estimated effects and the true effects may be different to our findings, which will limit the relevance of our findings to clinical decisions. To overcome the limitations of the evidence, large-scale RCTs are needed to determine whether personalised interventions, adapted or developed, for subgroups of children with conduct problems are effective in improving outcomes. Consensus on the most appropriate measures to use in these studies is needed in order to facilitate cross-study comparisons. Persistent conduct problems predict a range of adverse long-term outcomes, so future research should investigate the medium- and long-term effects of personalised treatments. Studies are needed in low- and middle-income countries as well as studies recruiting children aged between nine and 12 years, as they were under-represented in the studies.

PLAIN LANGUAGE SUMMARY

Personalised interventions for children with conduct problems

Key messages

There is currently little evidence to support personalising or tailoring interventions for children with conduct problems. What little evidence exists is of low quality. Therefore, it is unclear whether personalising interventions can improve outcomes for children with conduct problems. Further high-quality research is needed.

What are conduct problems?

Conduct problems are a range of disruptive behaviours in childhood that can have a negative impact on an individual's life. Conduct problems may lead to difficulties later in adolescence and adulthood, including antisocial behaviour, substance misuse, difficulties with education, and mental health problems.

It has been proposed that there are different subgroups of children with conduct problems. These subgroups include variations in the age conduct problems start, emotional difficulties, attention deficit hyperactivity disorder, parental characteristics, the level of aggression within antisocial behaviour, and the influence of genetic and environmental factors in relation to callous unemotional traits. These subgroups of children may respond differently to treatment, and it is therefore important to establish whether targeting interventions to these subgroups of children may result in better outcomes.

Targeted or 'personalised' interventions are those which tailor different aspects of treatment to the needs of particular subgroups of parents and children. For example, a subgroup of children with conduct problems experiencing interparental conflict (disagreements between parents) could potentially benefit from a parenting programme tailored to include additional sessions focusing on interparental conflict and offering particular techniques to address this issue. A non-personalised intervention would only provide the parenting programme, and not contain the additional sessions focusing on interparental conflict.

What did we want to find out?

This review assessed whether personalised interventions that have been adapted or developed for a specific subgroup of children with conduct problems are effective in improving conduct problems.

What did we do?

We undertook an extensive search of 13 databases. We also searched reference lists of included studies and contacted subject experts.

We only selected studies known as 'randomised controlled trials (RCTs)'. In this type of study, participants are allocated to groups randomly. One group receives the intervention and the other receives a different treatment or no treatment at all. RCTs aim to reduce the risk of introducing bias in clinical studies.

Subgroups of children with conduct problems were separated into three broad categories: children with co-occurring conditions (e.g. emotional difficulties), parent characteristics (e.g. conflict between parents) or familial/environmental circumstances (e.g. rural families).

All studies compared personalised interventions to non-personalised interventions.

We combined results from several studies that used the same measures to assess improvements in child conduct problems. Where this was not possible, we report studies individually.

What did we find?

We identified 13 RCTs with 858 participants to include in the review. Seven studies were conducted in the USA, five in Australia, and one in Germany. Eleven studies reported their source of funding, with five studies receiving grants from the National Institute of Mental Health. In total, 15 different funders supported the studies.

There was very little evidence from these studies that the personalised interventions were more effective than the non-personalised interventions in improving child conduct problems or disruptive behaviour.

What are the limitations of the evidence?

We have little confidence in the evidence because many of the studies had design limitations. There was also variation in the length of treatment and in how it was delivered. This means that we need to be cautious in interpreting the results that we found, and they may not be reliable.

How up to date is this evidence?

We searched for studies up to February 2022.

SUMMARY OF FINDINGS

Summary of findings 1. Personalised intervention versus non-personalised intervention for subgroups of children with conduct problems

Personalised intervention versus non-personalised intervention for subgroups of children with conduct problems[a]

Patient or population: subgroups of children with conduct problems
Setting: mixed (university clinic, online, community)
Intervention: personalised intervention
Comparison: comparison intervention (non-personalised parent education/parent training)

Outcomes		Illustrative comparative risks* (95% CI)		№ of participants (studies)	Certainty of the evidence (GRADE)	Direction of effects
		Assumed risk	Corresponding risk			
		Non-personalised	Personalised			
Improvement in child conduct problems or disruptive behaviour	Assessed with: Problem subscale of the Eyberg Child Behaviour Inventory (36 items scored yes or no) Measured at: short-term follow-up	The mean score in the control group ranged from 8.88 to 13.65	The mean score in the intervention group was, on average, **3.04 lower** (6.06 lower to 0.02 lower)	278 (6 RCTs)	⊕⊖⊖⊖ **Very low**[b]	A higher score indicates greater behavioural problems.
	Assessed with: Intensity subscale of the Eyberg Child Behaviour Inventory (36 items rated on 7-point Likert scale (1 = never occurs to 7 = always occurs) Measured at: short-term follow-up	The mean score in the control group ranged from 91.63 to 129.87	The mean score in the intervention group was, on average,**6.25 lower** (16.66 lower to 4.15 higher)	278 (6 RCTs)	⊕⊖⊖⊖ **Very low**[b]	A higher score indicates greater behavioural problems.
	Assessed with: Externalising subscale (reflects conflict with others and violation of social norms) of the Child Behaviour Checklist (36 items rated on 3-point Lik-	The mean score in the control group ranged from 19.47 to 48.82	The mean score in the intervention group was, on average, **2.19 lower** (6.97 lower to 2.59 higher)	189 (3 RCTs)	⊕⊖⊖⊖ **Very low**[b]	A higher score indicates greater behavioural problems.

	ert scale (0 = 'Absent', 1 = 'Occurs sometimes', 2 = 'Occurs often') Measured at: short-term follow-up				
Any adverse events	See comment	See comment	See comment	See comment	None of the included studies reported data related to adverse events.
Personalised treatment outcomes, relevant to each subgroup	See comment	See comment	See comment	See comment	Although 6 studies reported personalised treatment outcomes, relevant to each subgroup, we were unable to pool the data due to differences between the measures used and the statistical heterogeneity this would produce in analysis. The results for this outcome are inconclusive.
Parenting skills and knowledge	See comment	See comment	See comment	See comment	Although 7 studies reported personalised treatment outcomes, relevant to each subgroup, we were unable to pool the data due to differences between the measures used and the statistical heterogeneity this would produce in analysis. The results for this outcome are inconclusive.

*__The risk in the intervention group__ (and its 95% CI) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).
CI: confidence interval; **RCT:** randomised controlled trial.

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.

[a]There is no commonly agreed measure for the assessment of child conduct problems. This table reports on all measures and subscales which were suitable for pooling for each outcome.
[b]Downgraded three levels due to risk of bias (there was a high risk of bias because it was not possible to blind participants, personnel, and outcome assessors), indirectness (the sample consisted of only children aged three to nine years and so we cannot generalise to children aged between nine and 12 years), imprecision (the number of participants was small, fewer than 400), and inconsistency (there was heterogeneity in the population and outcomes measures used in the studies).

BACKGROUND

Description of the condition

Conduct problems are a range of antisocial and disruptive behaviours that can be diagnosed as conduct disorder (CD) or oppositional defiant disorder (ODD), with ODD symptoms sometimes acting as a precursor to the onset of the more severe CD symptoms (Frick 2012; Moffitt 2008). CD is characterised by a repetitive and persistent pattern of behaviour in which the basic rights of others or major age-appropriate norms or rules are violated, whereas children with ODD demonstrate defiant behaviour, irritability, and vindictiveness (APA 2013). Epidemiological studies have identified that between 5% and 10% of children and adolescents have significant problems with conduct and disruptive behaviour (Moffit 2009), making it the most common behavioural and mental health problem in children and young people globally (Collishaw 2004), and the most common reason for referral of young children to child and adolescent mental health services in the UK (NICE 2017).

Conduct problems are an important, long-term condition of childhood (Murphy 2013). They predict the development of antisocial behaviour and substance misuse in adulthood, and poor educational outcomes and increased physical health burden throughout life (Odgers 2007). They are the most common precursor of adult mental health problems across the spectrum (Copeland 2009; Kim-Cohen 2003). The 2010 Global Burden of Disease Study identified CD as a significant contributor to global years lived with disability (YLD), ranking it the 30th leading cause of non-fatal burden worldwide (Erskine 2014). CD is ranked the fourth leading cause of global YLDs for children aged five to nine years, and second for boys in this age group. As well as the impact on the individual child and family, there is an increased cost to the public purse; Rivenbark 2018 found that childhood conduct problems signalled high future costs, up to the age of 38 years, across criminal justice, health, and social welfare services. The early treatment and prevention of conduct problems is therefore of tremendous importance.

In recent years, there has been increasing awareness of substantial heterogeneity within conduct problems, so that it is now recognised that there are several 'subgroups' of children with conduct problems (Fairchild 2019; Frick 2016; Klahr 2014; McKay 2020). Subgroups of children in this review refers to a group of children who share similar characteristics; it should not be confused with subgroup analyses. Subgroups of conduct problems may include variations in the of age of onset (Silberg 2015), level of aggression within antisocial behaviour (Loeber 1985), comorbidity with attention deficit hyperactivity disorder (ADHD; Waschbusch 2002), and influence of genetic and environmental factors in relation to level of callous-unemotional (CU) traits or 'limited prosocial emotion' (LPE; Viding 2005). These heterogeneous subgroups can exhibit differences in aetiology, developmental trajectory, and likely prognosis (Frick 2016; Klahr 2014), with some studies reporting differential treatment outcomes (Hawes 2014; Reyno 2006).

Particular family characteristics have also been identified, both as risk factors for the development of conduct problems and as moderators of treatment effectiveness; for example, maternal mental health (Hutchings 2012), and contact with child protection services (Drugli 2010). Maternal ADHD symptoms have been associated with child ADHD and ODD symptoms (Zisser 2012), and may limit the improvement shown by children with ADHD in response to treatment (Chronis-Tuscano 2011; Sonuga-Barke 2002). Therefore, it is critical to identify subgroups defined by familial factors in addition to the recognition of heterogeneity on an individual level.

An example of the greater recognition of the importance of such subgroup heterogeneity is the decision by the APA 2013 to incorporate a new specifier in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) to describe children and young people with conduct problems who present with LPE. LPE is characterised by the presence of two or more of the following criteria over at least 12 months, and in multiple relationships and settings: 1. lack of remorse or guilt; 2. callous – lack of empathy; 3. shallow or deficient affect; and 4. lack of concern about performance (APA 2013; Jambroes 2016). Two reviews regarding chronic irritability and anger in ODD have also recommended the inclusion of a specific irritability subtype for ODD in the 11th revision of the *International Classification of Diseases and Related Health Problems* (ICD-11; ICD 2023), in response to the greater level of severity and impairment experienced by some children (Evans 2017; Lochman 2014). Subsequently, a "with chronic irritability and anger" qualifier has been included in ICD-11 to characterise presentations of ODD with prevailing, persistent irritable mood or anger and it has also been recognised that this presentation significantly increases risk for subsequent depression and anxiety (Reed 2019).

Additionally, outcomes of programmes may vary for different subgroups of parents and children (e.g. depressed mothers or children with callous unemotional traits) (Koerting 2013). Personalised approaches have therefore been proposed to enable intervention components to be tailored to the particular needs of such subgroups (Kennedy 2017). Such differences between subgroups have prompted debate as to whether a 'one-size-fits-all' model of intervention, which fails to take account of this heterogeneity, may be limited in its effectiveness. It is hoped, for example, that the addition of the LPE specifier to the DSM-5 will encourage a more precise diagnosis and the acknowledgement of 'an emerging subgroup within conduct problems', thereby promoting more targeted treatment research (APA 2013). It is important to improve measurement, particularly in relation to neurocognitive and social risk factor indices in order to develop more targeted approaches to intervention (Viding 2020). However, evidence from individual participant data meta-analyses of Incredible Years parenting programmes suggest that subgroup differences may not be as significant as previously thought (Gardner 2019; Leijten 2018; Leijten 2019). Therefore, it is important to consider the extent to which personalised interventions may affect outcomes for subgroups of children with CD.

Description of the intervention

Current recommended interventions

The gold standard, evidence-based intervention for the treatment of conduct problems in children is behavioural parent training (Scott 2009). However, evaluations of even the best parent-training programmes estimate that a quarter to a third of families and their children do not benefit (Scott 2009). Parent training also requires substantial commitment and organisation from parents and can be undermined as a treatment due to dropout or failure to engage.

Although recent interventions have sought to trial new methods of delivery that could address particular issues with provision and attendance, such as internet-delivered parent training (Högström 2015; Sourander 2016), there are still inherent difficulties with implementing behavioural parent training. Families with children diagnosed with ODD, CD, or ADHD who are appropriate for behavioural parent training, commonly do not enrol, enrol but never attend treatment, drop out prematurely, or do not fully engage in within-session or between-session skills implementation (Chacko 2012; Fernandez 2011; Peters 2005). One review of 262 studies of behavioural parent training found a combined dropout rate of 51%, with 25% not enroling despite being appropriate for the programme, and 26% beginning but not completing the training (Chacko 2016). Limitations in the reach of parent-training programmes are therefore a significant problem (NICE 2017; Pilling 2013).

In addition to such limitations in reach and effectiveness, differential outcomes of parent training have been associated with subgroups of children with conduct problems. High CU traits (or LPE) can predict poor outcomes across parent-training interventions (Hawes 2014), and there is evidence that poor economic circumstances, marital discord, parental mental health problems, and parental hostility are associated with poorer outcomes (Reyno 2006). Paternal substance abuse and child comorbid anxiety or depression have been identified also as factors predicting poorer outcomes (Beauchaine 2005). However, the evidence in this area is not clear-cut and one comprehensive meta-analysis found that a range of family characteristics, which are usually associated with a poorer outcome from parent training, did not moderate a less favourable response (Gardner 2016). One recent network meta-analysis investigating the most effective parenting programme content for disruptive child behaviour concluded that future research should focus on testing individual family differences (Leijten 2022). This will improve understanding of the effectiveness of interventions for each individual family by exploring the complex interactions between family characteristics and programme components.

Although parent training is the primary recommended intervention for children with conduct or behavioural problems, other treatments, such as cognitive problem-solving programmes, are recommended (NICE 2017), and cognitive behavioural treatments have been investigated in the treatment of aggressive behaviour in children (Smeets 2015; Sukhodolsky 2004). Meta-analyses of cognitive behavioural therapy (CBT) for aggression in children and adolescents have demonstrated medium effect sizes (Smeets 2015; Sukhodolsky 2004), and have suggested that further research is necessary to determine whether subgroups of individuals with predominantly reactive or proactive aggression may respond differentially to CBT intervention (Smeets 2015). Therefore, additional investigation is vital to clarify whether subgroup classification is associated with differential outcome across available interventions and, if so, whether understanding the underlying reasons for this could potentially lead to the development of more effective treatments.

Personalisation

Whilst parent-training programmes are recognised as an effective treatment (Dretzke 2009), personalisation seeks to address possible limitations in the effectiveness of such programmes. Differences across developmental pathways and clinical presentations have been identified within the wider diagnostic classification of conduct problems (Frick 2016; Klahr 2014). Recognition of these differences could therefore aid the tailoring or personalisation of interventions to address the specific needs of particular subgroups (Frick 2016). Such personalisation aligns with the recent strategy of the Medical Research Council (MRC) to 'embrace a stratified medicine approach' (MRC 2017). Stratified medicine is described as "identifying groups of people with shared characteristics within or across specific disorders … looking beyond standard diagnostic categories to find new treatments and better ways of using existing treatments." Personalised interventions may, therefore, include novel treatments or may involve additional or adjunctive interventions alongside existing standard evidence-based interventions. While the National Institute for Mental Health (NIMH) in the USA has called for mental health researchers to "expand and deepen the focus to personalise intervention research" (Fisher 2015), the science of personalisation in relation to child mental health is a novel field in the early stage of development (Ng 2016; Scott 2016). The nomenclature of 'personalisation,' 'precision,' 'targeted,' 'tailored,' 'stratified,' and 'adapted' interventions are often used interchangeably in the literature, and often encompass the same approach to intervention. In this review, we use the term 'personalised interventions' to capture these interchangeable terms.

How the intervention might work

Personalised interventions are tailored for particular subgroups based on shared subgroup characteristics. It is possible that personalised treatments may include elements of parent-training programmes, or supplement existing interventions with additional techniques to address subgroup heterogeneity. For example, a subgroup of children with conduct problems experiencing parental hostility could potentially benefit from a parenting programme tailored to include additional sessions focusing on hostility and offering particular techniques to address this issue. Alternatively, personalised interventions may be entirely novel treatments without any reference to parent training, or adaptations of existing non-parent-training-based interventions for conduct problems.

Subgroups of children with conduct problems that may benefit from personalisation

An example of a subgroup difference that could be addressed by a personalised approach to intervention is that of children who have high versus low CU traits (Frick 2014). Children with low CU traits are more likely to be sensitive to traditional disciplinary strategies employed in parenting programmes (Scott 2015), whereas children with high CU traits appear genetically vulnerable to antisocial behaviour (Viding 2005) and relatively insensitive to punishment, threat, and others' distress (Pardini 2012). These vulnerabilities may cause insensitivity to certain critical components of traditional behavioural approaches (Hawes 2005), and children with high CU traits may benefit from programmes focusing on the positive dimension of parenting (Muratori 2016). Programmes that have been successful in the treatment of CU traits may contain elements that are more beneficial for this particular subgroup; for example, supporting an increase in parental warmth (e.g. Fast Track Intervention; Pasalich 2016).

Similarly, maternal ADHD symptoms have been associated with poorer parent-training outcomes for children with ADHD (Chronis-

Tuscano 2011; Sonuga-Barke 2002). Lack of reduction in negative parenting behaviours has been identified as a possible explanation for the relationship between maternal ADHD symptoms and poorer post-treatment child behavioural outcomes (Chronis-Tuscano 2011). The ability of parents to exhibit positive parenting behaviours is of vital importance for behavioural change and has been shown to act as a protective factor against the development of conduct problems in children with ADHD (Chronis 2007). For these families, management of maternal ADHD symptoms to aid implementation of positive parenting strategies could be beneficial.

Targeting other aspects of parental mental health may also be beneficial; for example, treating maternal depression appears to improve outcomes for children with conduct problems (Hutchings 2012). Further, following evidence suggesting that subgroups of children presenting with emotional dysregulation may respond differentially to parent-training programmes, Scott 2012 proposed that it may be worthwhile to screen children prior to allocation of parenting interventions to ensure individual differences are accounted for.

Personalisation can also focus on familial and environmental characteristics such as parents on low income, single parent households, foster parents, and families from rural areas. For example, rural families may struggle to access services. Therefore, services may need to be adapted for this particular subgroup of families (e.g. online delivery).

Personalised interventions, therefore, may have the potential to improve outcomes by targeting the specific needs of predefined subgroups. For example, children with co-occurring conditions, parents with particular characteristics, or wider familial/environmental factors. This review investigates each of these.

Why it is important to do this review

While existing reviews have identified considerable heterogeneity within conduct problems and have investigated differential response to treatment (Gardner 2016; Hawes 2014; Klahr 2014; Shelleby 2014; Wilkinson 2016), to date, there has been no attempt to identify and synthesise the evidence on personalised interventions for subgroups of children with conduct problems. Previous Cochrane Reviews focusing on the treatment of conduct problems have evaluated standard, group-based, parenting programmes for improving emotional and behavioural adjustment in young children (Barlow 2016); improving early-onset conduct problems in children aged three to 12 years (Furlong 2012); and improving conduct problems in older children and adolescents (Woolfenden 2001). This review, therefore, aims to address a gap in the treatment literature by systematically identifying and appraising the evidence for personalised treatments for subgroups of children with conduct problems.

OBJECTIVES

To assess whether personalised interventions, adapted or developed, for subgroups of children with conduct problems are effective in improving outcomes.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs).

Types of participants

We included children aged between two and 12 years, in any setting, with conduct problems. Prespecified subgroups of children with conduct problems, however defined, were eligible for inclusion. This included children with diagnoses of CD and ODD but was not restricted to those with formally diagnosed conditions.

Due to differences in associated risk factors and developmental trajectory between child-onset and adolescent-onset conduct problems (Silberg 2015), we excluded studies that included a proportion of children older than 12 years of age. We also excluded studies that focused on a geographical area of disadvantage rather than prespecifying a subgroup of socioeconomically disadvantaged families. This excluded studies exploring generalisability rather than heterogeneity.

Types of interventions

We included any personalised intervention that was adapted or developed for a subgroup of children with conduct problems. This could have included adaptations to standard practice, such as parent-training programmes (e.g. Helping the Non-Compliant Child) or other recommended interventions for children with conduct problems (e.g. Triple P), or interventions developed specifically to target subgroups of children with conduct problems (e.g. Integrated Family Intervention for Child Conduct Problems). We excluded non-personalised and non-psychological interventions (e.g. pharmacological or dietary intervention).

Relevant comparators included non-personalised interventions, waitlist control, treatment as usual (TAU), or no intervention.

Types of outcome measures

Outcomes were used as criteria for inclusion in this review, rather than a list of outcomes of interest. This was to ensure that the included studies met the review's objective, which was to evaluate the (often disparate) interventions designed to reduce subgroups of children with conduct problems, as specified within the DSM-5.

Primary outcomes

1. Improvement in child conduct problems or disruptive behaviour, measured by, for example, the Eyberg Child Behavior Inventory (ECBI) (Eyberg 1978).
2. Any adverse events (such as emotional or psychological trauma of any type, perhaps if a parent was to experience an increase in anxiety or depression throughout the course of a parent-focused treatment; or an increase in negative parenting practices, such as shouting or criticism). These could be measured through validated tools such as the Family Life Interview (Llewellyn 2010), or through practitioner reports.

Secondary outcomes

1. Personalised treatment outcomes relevant to each subgroup (e.g. reduction in ADHD symptoms, measured by, for example,

the Conners Abbreviated Parent/Teacher Rating Scale (CAP/TRS; Conners 1994); reduction in CU traits, measured by, for example, the Inventory of Callous-Unemotional Traits (ICU; Frick 2004); or maternal depression, measured by, for example, the Beck Depression Inventory-II (BDI-II; Beck 1996)).

2. Parenting skills and knowledge, measured by direct observation or self-report (e.g. Parenting Scale (PS); Arnold 1993).
3. Family functioning, measured by, for example, the Family Assessment Device (FAD; Epstein 1983).
4. Engagement and decreased dropout, measured by number of sessions attended.
5. Educational outcomes, measured by, for example, the items capturing child academic performance from the MacArthur Health Behavior Questionnaire (MacArthur HBQ; Boyce 2002; Essex 2002), or developmental assessments such as the Bayley Scales of Infant and Toddler Development – Third Edition (Bayley-III; Bayley 2006), or the Mullen Scales of Early Learning (MSEL; Mullen 1995).

Primary and secondary outcomes could have been measured by child, parent, or carer reports, though questionnaires, interviews, or observational assessments.

We collected outcomes for the following time points: short term (one month postintervention or less), medium term (greater than one month to less than 12 months postintervention), and long term (12 months postintervention or greater).

Search methods for identification of studies

We conducted searches for all available years in each database, modifying the search terms as necessary. Searches were conducted for this review in May 2017, March 2019, June 2020, and February 2022. Searches were not limited by language or publication date. We sought translations of any studies of potential relevance.

Electronic searches

We identified relevant trials by searching the electronic databases and trials registers listed below, using the search strategies in Appendix 1.

1. Cochrane Central Register of Controlled Trials (CENTRAL), which includes the Cochrane Developmental, Psychosocial and Learning Problems Group Specialised Register, accessed via Cochrane Register of Studies Online (CRSO; crso.cochrane.org; searched 1 February 2022).
2. MEDLINE Ovid (1946 to 1 February 2022).
3. MEDLINE EPub Ahead of Print Ovid (searched 1 February 2022).
4. MEDLINE In-Process and Other Non-Indexed Citations Ovid (searched 1 February 2022).
5. Embase Ovid (1974 to 1 February 2022).
6. APA PsycINFO Ovid (1967 to 1 February 2022).
7. CINAHLPlus EBSCOhost (Cumulative Index to Nursing and Allied Health Literature; 1937 to 3 February 2022).
8. ERIC EBSCOhost (Education Resources Information Center; 1966 to 3 February 2022).
9. Conference Proceedings Citation Index – Science Web of Science (CPCI-S; 1990 to 3 February 2022).
10. Conference Proceedings Citation Index – Social Science & Humanities Web of Science (CPCI-SS&H; 1990 to 3 February 2022).
11. *Cochrane Database of Systematic Reviews* (CDSR; 2022, Issue 2), part of the Cochrane Library (searched 3 February 2022).
12. Database of Abstracts of Reviews of Effects (DARE; 2015, Issue 2. Final issue), part of the Cochrane Library (searched 17 May 2017 – no issue after this date).
13. Epistemonikos (www.epistemonikos.org; searched 3 February 2022).
14. ClinicalTrials.gov (clinicaltrials.gov; searched 3 February 2022).
15. World Health Organization International Clinical Trials Registry Platform (ICTRP) (apps.who.int/trialsearch; 3 February 2022).

Searching other resources

We examined the reference lists of included studies and relevant review articles up to February 2022, in order to identify further studies. We contacted authors of the identified RCTs to request further information, and contacted experts and researchers working in the field in order to search for unpublished and ongoing studies. On 5 October 2022 we ran searches in Retraction Watch, MEDLINE, and Embase to identify retractions of the included studies. We found no retractions.

Data collection and analysis

We were unable to use some of the preplanned methods presented in the protocol (Kennedy 2017). In successive sections here, we describe only the methods we were able to use.

Selection of studies

For the first search in May 2017, four review authors (LF, EH, LK, and LJ) independently selected studies in a two-phase process. First, three review authors (LF, LJ, and LK) removed duplicate records and obviously irrelevant records based on a preliminary screen of titles. Then, they screened the remaining titles and abstracts for eligibility and retrieved full-text reports of potentially relevant studies. Second, two review authors (LF and EH) assessed the full-text reports for inclusion against the selection criteria (Criteria for considering studies for this review).

For the three top-up searches in March 2019, June 2020, and February 2022, the same two-phase process was implemented independently: four review authors (EH, VR, CL, and GS) screened the titles and abstracts for eligibility; then four review authors (EH, VR, CL, and GS) retrieved the full-text reports of potentially relevant studies and assessed them for inclusion against the selection criteria.

Within the original and updated searches, one review author (EK) checked the full-text reports of studies which appeared to meet inclusion criteria but were excluded and the studies that two review authors (LF and EH) deemed eligible. Three review authors (LF, EH, and EK) resolved any disagreements by discussion.

Any studies that appeared to meet inclusion criteria but were excluded, are reported in the Characteristics of excluded studies table.

The flow of information throughout the phases of this review is documented and presented in a flow chart, as described by the PRISMA Statement (Moher 2009).

Data extraction and management

For each eligible study, three review authors (EH, GS, and LK) independently extracted information on a number of key characteristics (listed below) using electronic data collection form templates provided by the review group, which we modified and piloted. We addressed any differences through discussion. We contacted authors of studies directly when target information was unreported or unclear, in order to clarify or complete extracted data. We entered and organised citations and data in Review Manager 5 (Review Manager 2020) and RevMan Web (RevMan Web 2022), and entered the collected data onto predesigned Characteristics of included studies tables. Within the Discussion section, we considered the implications of any bias on outcomes or meta-analyses in this review.

1. General information about the study: title; authors; year of publication; eligibility
2. Methods: study design; unit allocation; and duration of the study
3. Participants: setting; recruitment method; withdrawal from study; relevant diagnostic details; age; sex; race/ethnicity; further sociodemographic detail; subgroup allocation
4. Intervention: considerations and components related to the intervention, including theoretical basis, duration, session frequency, individual or group-based delivery, staff qualifications, outcome measures, and scales; economic information; compliance and integrity of delivery
5. Outcomes: primary and secondary outcomes and time points considered in the review

Assessment of risk of bias in included studies

Four review authors (EH, LF, CL, and GS) independently assessed the risk of bias in each included study using the risk of bias criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). For each included study, review authors assessed the risk of bias using the RoB 1 tool across the following domains, and assigned ratings of low, high, or unclear risk of bias: sequence generation, allocation concealment, blinding of participants or personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, and other sources of bias (lack of adherence to treatment manual and group differences). A third review author (EK) resolved any disagreements by discussion until we reached consensus. The final judgements are reported in a risk of bias table for each included study with a brief rationale for each decision. Overall risk of bias assessments were formed through discussion between four review authors (EH, LF, CL, and GS) following the guidance in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Measures of treatment effect

Continuous data

Where means and standard deviations (SD) were available or were not but could be computed from other test statistics (standard errors (SE) and 95% confidence intervals (CI)), we calculated the mean difference (MD) and 95% CI for continuous data measured on the same scale. We requested additional information from the study authors when reports had insufficient data.

Dichotomous data

We calculated risk ratios (RR) with 95% CIs for dichotomous data.

Unit of analysis issues

Cluster-randomised trials

We identified no eligible cluster-randomised trials for inclusion in the review.

Multiple treatment groups

Where a study included multiple arms, we first assessed whether all arms met the inclusion criteria. Where two or more arms involved personalised interventions of interest and shared an eligible comparator (non-personalised intervention, TAU, or waitlist), we split the shared group into two or more groups with small sample sizes to avoid double counting the participants. These comparisons were separated into different forest plots, such that for three-arms trials with both a comparison intervention and no treatment control group, we made the following separate comparisons: personalised intervention compared to the comparison intervention and personalised intervention compared to the no treatment control. When combining groups in multiple-arm trials, we summed the sample sizes and events across groups for dichotomous data; and we combined sample sizes, means and SD for continuous data, in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021).

Dealing with missing data

We assessed missing data and dropouts or attrition for each included study and contacted the study authors when there were missing or unclear data. We attempted to retrieve any missing data from the study authors in addition to numbers, characteristics and reasons for dropout.

Assessment of heterogeneity

We assessed clinical and methodological heterogeneity by examining differences in the characteristics of participants, methodology, the type and intensity of interventions, and types of outcome measures used. We assessed statistical heterogeneity using a Chi^2 test (with a $P < 0.10$), with heterogeneity being indicated by a Chi^2 statistic greater than the degrees of freedom and a small P value, and by visual inspection of forest plots where heterogeneity was indicated by limited overlap of studies on the forest plot or by outliers. We also used the I^2 statistic to detect inconsistencies across studies and determine the approximate proportion of variation that was due to heterogeneity rather than sampling error (chance). We interpreted I^2 values as follows: 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; 75% to 100% shows considerable heterogeneity. We accounted for the strength of evidence for heterogeneity (e.g. the P value from the Chi^2 test) while interpreting I^2 values. As part of a random-effects meta-analysis, we reported Tau^2 as an estimate of between-study variability. Where there was evidence of heterogeneity, we discussed the possible reasons for it within the Discussion (see Overall completeness and applicability of evidence).

Assessment of reporting biases

We were unable to assess reporting bias as we did not combine data from a sufficient number of studies in the meta-analyses (a minimum of 10 studies was required). However, we did search trial

registrations and protocols to compare the outcomes listed with those in the published reports, to assess for selective reporting.

Data synthesis

Where appropriate, we pooled the results from the included studies using RevMan 5 (Review Manager 2020) and RevMan Web (RevMan Web 2022). Due to the large number of different measures and different subscales of measures included in this review, and the heterogeneity between these measures and subscales, we chose not to pool multiple measures for a particular outcome using standardised mean differences. This decision was made because the results of pooling analyses in this way would produce inconclusive findings which we would be unable to draw any meaningful conclusions from. We therefore only pooled data for an outcome where we had multiple studies reporting the same measure or the same subscale of a measure. Providing there were no serious reporting or publications biases, we performed separate meta-analyses and syntheses of the data for each measure within a particular outcome. We assessed reporting bias through discussion between study authors (EH, LF, CL, and GS), using unclear, low, or high thresholds. Studies that reported a high threshold of reporting bias would have been excluded from meta-analyses. However, no studies were rated as such, and all available studies were included in the analyses. As we expected, identified studies to be estimating different but related intervention effects, we used a random-effects model with an inverse-variance method to calculate MD effect sizes with 95% CIs and displayed them with forest plots. We considered separate meta-analyses for short-term (one month postintervention or less), medium-term (greater than one month to less than 12 months), and long-term (12 months postintervention or greater) outcomes. Where studies were clinically diverse, provided insufficient data, was the only included study reporting a measure, or all three, we provided a narrative description of the results.

Subgroup analysis and investigation of heterogeneity

Due to the small number of trials that could be included in each meta-analysis, we did not conduct subgroup analyses.

Sensitivity analysis

Planned sensitivity analyses were not possible due to too few studies.

Summary of findings and assessment of the certainty of the evidence

Pairs of review authors (EH, LF, CL or GS) independently assessed the certainty of the body of evidence for all outcomes (improvement in child conduct problems or disruptive behaviour, any adverse events personalised treatment outcomes, parenting skills and knowledge, family functioning, engagement and decreased dropout, educational outcomes) using the GRADE approach (Schünemann 2021). We rated the overall certainty of evidence as high, moderate, low, or very low based on the following GRADE criteria: risk of bias (judgements on the overall risk of bias were agreed after discussion between four review authors (EH, LF, CL, and GS)); indirectness of evidence; unexplained heterogeneity or inconsistent results; imprecision of results; and high risk of publication bias. We resolved any disagreements through discussion between all authors.

In line with our protocol, we reported improvement in child conduct problems or disruptive behaviour, personalised treatment outcomes, relevant to each subgroup, and parenting skills and knowledge outcomes in the summary of findings table. We chose to present the results at short-term follow-up in the summary of findings table because a diagnosis of a CD is strongly associated with poor educational performance, social isolation, and, in adolescence, substance misuse, and increased contact with the criminal justice system. This association continues into adult life with poorer educational and occupational outcomes, involvement with the criminal justice system, and a high level of mental health problems (NICE 2017). Taking this into account, a focus on preventing or reducing the escalation of existing conduct problems or disruptive behaviour as early as possible is key (NICE 2017). As there is no clinically recommended measure for conduct problems, we reported the pooled analyses for all measures of our outcomes in the summary of findings table to provide the most representative data to inform clinical decision-making. We reported the GRADE ratings in a summary of findings table, which we created using GRADEpro GDT (GRADEpro GDT), for our main comparison of personalised intervention versus non-personalised intervention. In this table, we also present the effects of the interventions as MDs, with accompanying 95% CIs for each outcome assessed.

RESULTS

Description of studies

See Characteristics of included studies table.

Results of the search

The electronic searches up to February 2022 found 39,852 records. We also examined the reference lists of included studies and relevant review articles and identified a further six systematic reviews for reference checking (Bakker 2017; Dretzke 2009; Epstein 2015; Fossum 2009; Kaminski 2008; Lundahl 2006). We identified no unique studies at this stage. After the removal of duplicates, review authors screened the titles and abstracts of 27,795 records and excluded 27,183. We were unable to access five full texts, which are awaiting classification (see Characteristics of studies awaiting classification table). We assessed 607 full-text records for eligibility, excluded 592 reports (Excluded studies), and included 15 reports (containing 13 studies) in the review (Included studies). We found no ongoing studies. See Figure 1.

Figure 1. PRISMA flow diagram.

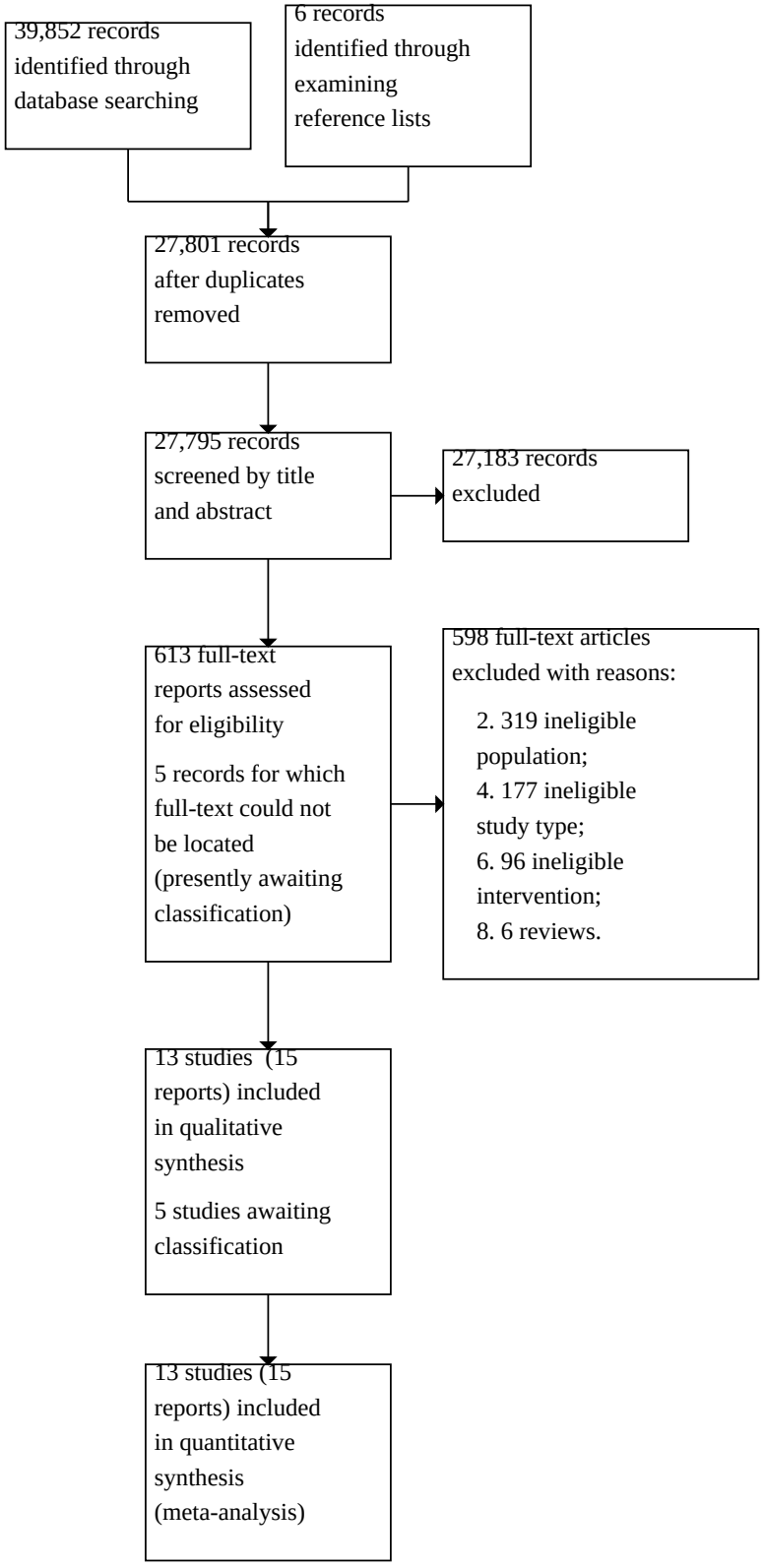
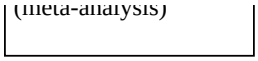


Figure 1. (Continued)



Included studies

This review included 13 studies (from 15 reports) that were deemed eligible after full-text screening against the inclusion criteria (Bor 2002; Dadds 1992; Dadds 2019; Görtz-Dorten 2019; Greene 2004; Jones 2014; Jouriles 2001; Jouriles 2009; Markie-Dadds 2006; McCabe 2009; Mersky 2016; Nicholson 1999; Parent 2022). The studies were published over a 27-year period, with the earliest paper published in 1992 (Dadds 1992) and the most recent in 2021 (Parent 2022).

We only received missing data from Dadds 2019. In this case we requested the Diagnostic Interview Schedule For Children, Adolescents, And Parents (DISCAP; Holland 1997) scores postintervention and at follow-up. For all other studies with missing data, we conducted analyses using the participant and summary available data only, and we noted any missing data in the risk of bias table.

Design

All 13 included studies were RCTs. Eight studies used two arms, whilst the remaining five studies included three groups (Bor 2002; Markie-Dadds 2006; McCabe 2009; Mersky 2016; Nicholson 1999).

All studies with two arms compared the personalised intervention to a non-personalised intervention or existing services (Dadds 1992; Dadds 2019; Greene 2004; Görtz-Dorten 2019; Jones 2014; Jouriles 2001; Jouriles 2009; Parent 2022). One study compared the personalised intervention to both a non-personalised intervention and a TAU group (McCabe 2009). The remaining studies compared the personalised intervention to both a non-personalised intervention and a waitlist control group (Bor 2002; Markie-Dadds 2006; Mersky 2016; Nicholson 1999).

Setting

Seven studies were conducted in the USA (Greene 2004; Jones 2014; Jouriles 2001; Jouriles 2009; McCabe 2009; Mersky 2016; Parent 2022), five in Australia (Bor 2002; Dadds 1992; Dadds 2019; Markie-Dadds 2006; Nicholson 1999), and one in Germany (Görtz-Dorten 2019).

Five studies recruited participants from community settings using a variety of methods, including television, radio, and newspaper advertisements, posters and flyers, each targeted at the specific subgroup of children presenting with conduct problems (Bor 2002; Dadds 1992; Jones 2014; Markie-Dadds 2006; Parent 2022). Three studies recruited participants from outpatient community health settings (Görtz-Dorten 2019; Greene 2004; McCabe 2009). Two studies recruited parents from both the community (e.g. media advertisements) and referrals from organisations, including schools, therapy groups, and child welfare agencies, or outpatient clinics (Dadds 2019; Nicholson 1999). Two studies recruited mothers with dependent children in domestic violence shelters (Jouriles 2001; Jouriles 2009). One study recruited participants from referrals from agencies, including schools, child welfare agencies, and community clinicians (Mersky 2016).

Five studies were conducted in community clinics (Görtz-Dorten 2019; McCabe 2009; Mersky 2016; Nicholson 1999; Parent 2022). Two studies were conducted in university clinics (Dadds 1992; Jones 2014). Two studies were conducted in the family's home following departure from a shelter (Jouriles 2001; Jouriles 2009). One study delivered the intervention online using interactive and educational videomodules and videoconferencing (Dadds 2019) and one study completed treatment in community centres (Bor 2002). Markie-Dadds 2006 used a self-directed treatment and thus there was no external setting for the delivery of the intervention. One study did not report the location of the intervention (Greene 2004).

Sample size

Overall, the 13 included studies initially randomised 858 participants, with a mean number of 66 participants per study. Sample size ranged from 22 in both Dadds 1992 and Jones 2014 to 133 in Dadds 2019.

Participants

Subgroups of children with conduct problems in the 13 included studies were separated into three broad categories: co-occurring conditions (three studies); parent characteristics (two studies); and familial/environmental factors (eight studies). The three studies targeting subgroups of children with conduct problems and a co-occurring conditions included: children with comorbid attentional/hyperactive difficulties (Bor 2002); boys with ODD or CD and peer-related aggression (Görtz-Dorten 2019); and affectively dysregulated children (Greene 2004). The two studies targeting subgroups associated with parent characteristics included children exposed to intimate partner conflict (Jouriles 2001; Jouriles 2009). The remaining eight studies targeting subgroups related to familial or environmental factors, or both, included: children with isolated, single parents (Dadds 1992); children from rural families (Dadds 2019); children from low-income families (Jones 2014; Parent 2022); children living in rural and remote areas (Markie-Dadds 2006); young Mexican-American children (McCabe 2009); children from foster families (Mersky 2016); and stepfamilies (Nicholson 1999).

Participants ranged in age from two (Markie-Dadds 2006) to 12 (Görtz-Dorten 2019) years. Two studies only reported the mean age of participants (Dadds 1992: mean 4.57 years; Mersky 2016: mean 4.60 years). Most studies recruited both boys and girls with one study recruiting boys only (Görtz-Dorten 2019). Most studies provided data on ethnicity, with four reporting a predominantly Caucasian sample (Bor 2002; Greene 2004; Markie-Dadds 2006; Parent 2022), two reporting ethnically diverse samples (Jouriles 2001; Jouriles 2009), two reporting that most participants were of an ethnic minority (Jones 2014; Mersky 2016), and one study recruiting participants of a specific ethnicity (i.e. Mexican American; McCabe 2009).

Studies varied in the primary carers recruited. Two studies recruited only mothers (Jouriles 2001; Jouriles 2009). Three studies included parents and did not restrict recruitment to mothers or fathers, although most carers were mothers (Dadds 1992; Dadds 2019; Jones 2014). Four studies recruited parents but did not report details of their gender (Bor 2002; Görtz-Dorten 2019; McCabe 2009; Parent 2022). Two studies recruited both mothers and fathers and collected separately reported outcomes (Greene 2004; Markie-Dadds 2006). Mersky 2016 recruited foster parents, and most foster parents identified as the child's primary carer were female. Nicholson 1999 included parents and stepparents (81% stepfather families) and collected outcome data from both. Görtz-Dorten 2019 also included teacher-reported outcomes.

Interventions

Types of intervention

All 13 studies delivered a personalised intervention that was adapted or developed for a prespecified subgroup of children with conduct problems.

Three studies delivered a personalised parent-training intervention (Dadds 1992; Dadds 2019; Markie-Dadds 2006). Dadds 1992 delivered child management training (CMT) with adjunctive ally support training to parents. The CMT was developed by Sanders 1982, to help parents learn techniques and encourage problem-solving. Parents nominated an ally responsible for supporting them in three ways: 1. being available to offer support when required; 2. regularly communicating with and listening to the parent in casual discussions; and 3. participating in problem-solving sessions. Dadds 2019 delivered a manualised therapist-assisted online intervention based on the Integrated Family Intervention for Child Conduct Problems (Dadds 2006), designed to reduce child conduct problems and improve parent–child relationships. The intervention was implemented in two ways. First, parents were taught parenting strategies based on social learning theory in six interactive and educational video modules; they viewed role-plays, received handouts, activity ideas, and psychoeducation; and completed interactive exercises. Second, parents attended video sessions with a clinician to review module content and implementation of strategies. Markie-Dadds 2006 delivered a personalised adaptation of Triple P, a positive parenting programme, which aims to encourage positive parent–child relationships and help parents acquire skills for effective child behavioural management (Tully 1999). Markie-Dadds 2006 evaluated an Enhanced Self-Directed Triple P intervention in which parents worked through exercises in a 10-unit workbook and learned 17 core child-management strategies. Families also participated in weekly telephone consultations with a practitioner, with the intention of encouraging parents' problem-solving skills.

Two studies delivered a personalised cognitive behavioural intervention (Görtz-Dorten 2019; Greene 2004). Görtz-Dorten 2019 delivered an individual, social competence training that combined knowledge from child-, parent-, teacher-, and group-focused interventions. Although primarily delivered to children, additional support was provided to parents based upon the individual needs of the child. The intervention targeted problem-maintaining factors in specific life situations for each child, with the aim of changing social-cognitive functioning (such as information processing, impulsivity, problem-solving, and social skills). Greene 2004 delivered a collaborative problem-solving intervention that taught problem-solving skills to help parents collaboratively resolve disagreements and reduce conflict with their children. Parents learned how to understand cognitive factors that may contribute to aggressive behaviour and were taught strategies to handle unmet expectations. The intervention was primarily delivered to parents with children included at the therapist's discretion.

Two studies delivered adapted Parent–Child Interaction Therapy (PCIT) (McCabe 2009; Mersky 2016). These interventions were based on social learning models. McCabe 2009 maintained the core principles of PCIT, to establish secure parent–child relationships, in Guiando a Niños Activos (McCabe 2005). Cultural adaptations included: reference to cultural concepts throughout; rapport building due to increased session time; and adding representations of Mexican American families. Mersky 2016 delivered an extended PCIT to a group of four to eight carer–child dyads. The aim was to modify attitudes towards parenting and teach behavioural management skills through direct coaching, modelling, role play, didactic instruction, and homework and telephone consultations in two stages. First, carers practiced and acquired authoritative parenting skills and positive parent–child communication during child-directed interactions, and second, parents developed effective positive discipline and behaviour management skills during carer-directed interactions.

The remaining six studies delivered personalised behavioural family interventions (BFI) (Bor 2002; Jones 2014; Jouriles 2001; Jouriles 2009; Nicholson 1999; Parent 2022). Bor 2002 delivered an enhanced behavioural family intervention (EBFI), which consisted of parent training plus two adjunctive treatments: partner support training and coping skills training. The parent training taught parents how to increase positive interactions with their children and reduce inconsistent parenting strategies through modelling, role play, and homework assignments, with children attending six out of 10 of the sessions. Jones 2014 delivered a smartphone-enhanced, behavioural parent-training intervention, technology-enhanced helping the non-compliant child (McMahon 2003), to facilitate both improved child behaviour and family functioning. Carer–child dyads completed two stages: 1. learning to increase positive attention and remove questions and criticism during child-directed play; and 2. parents were taught a clear instruction sequence and to use a non-physical discipline procedure in the context of a parent's game. The smartphone components included: example videos for each helping the non-compliant child's skill, daily questionnaires of skill practice and progress, mid-week video-calls, weekly videotaped practice at home, and text message reminders. Jouriles 2001 and Jouriles 2009 delivered child management parent training as part of a multicomponent, project support, family intervention. The intervention was delivered primarily to mothers, but children received mentoring, and in Jouriles 2009, children were brought into sessions to evaluate the mother–child interaction. The intervention had two components; first, to provide support for mothers as they left the shelter; and second to teach child management strategies to positively enhance the mother–child relationship, increase desirable child behaviour, and decrease child undesirable behaviour. Skills were taught through instruction, role play, and homework. Nicholson 1999 delivered a BFI, supplemented by parenting and conflict resolution skills. The intervention included five main components: 1. stepfamily education; 2. positive

parenting skills training based on the behaviour family intervention model (Sanders 1993); 3. co-operative parenting skills training; 4. problem-solving and communication skills training; and 5. family activities training. Parent 2022 delivered 'Helping the Noncompliant Child' (HNC) to families (McMahon 2003). HNC is a therapist-delivered, criteria based (i.e. therapists conduct weekly observation and coding of skill use to determine progression through skills and programme completion) behavioural parent-training intervention. HNC included weekly face-to-face therapy sessions (60 minutes), as well as a brief midweek telephone check-in. HNC consists of two phases: differential attention (e.g. increasing positive attention, ignoring inappropriate behaviour), and compliance training (e.g. utilising time-outs). When parents progress to Phase II (i.e. compliance training), they continue to practice Phase I skills to maintain skill proficiency. Parent 2022 supplemented HNC with technological enhancement for their experimental intervention arm, which entailed detailing the full HNC protocol in addition to a digital companion, Tantrum Tamers. Tantrum Tamers is a Health Insurance Portability and Accountability Act compliant, interactive system that allowed therapists to monitor carer activity on the mobile application, as well as tailor the focus and pace of treatment based on parent practice and progress. The Tantrum Tamers application included: 1. daily surveys of skills practice, 2. weekly videorecorded home practice, 3. daily text reminders for skill practice and appointments; 4. video calls with the family midweek to problem-solve obstacles; and 5. skills video series to model new skills and share with other carers. Additionally, a homework checklist was added to remind carers of daily and weekly assignments.

Comparators

Three studies used both a non-personalised and waitlist control comparison. Bor 2002 compared their personalised EBFI to a non-personalised Standard Behavioural Family Intervention (SBFI). This non-personalised intervention consisted of the Standard Triple P – Positive Parenting Program. Bor 2002 also randomly allocated families to a waitlist control, in which families received no treatment and had no contact with the research team. Markie-Dadds 2006 compared their enhanced self-directed arm with a non-personalised self-directed arm and a waitlist control group. Families in the non-enhanced intervention received a 10-unit self-directed programme comprising *Every Parent and Every Parent's Workbook*. Families allocated to the waitlist intervention arm received no treatment and had no contact with the research team for 12 weeks. These families completed the postintervention measures and then received the programme of their choice, namely Enhanced Self-Directed Triple P or Self-Directed Triple P. These families took no further part in the study. Nicholson 1999 provided a comparison between their personalised therapist-directed programme, their self-directed programme, and a waitlist control group. The self-directed group received an identical BFI to the personalised arm, without the guidance of a therapist. Nicholson 1999 did not provide any further information on their comparison group.

Three studies used only TAU as their comparator. Mersky 2016 had three arms in their RCT. The first two consisted of brief PCIT and extended PCIT compared to a TAU control. Foster parents in the control group continued to receive their usual services, including case management and standard parent training. Foster children also continued to receive standard care options designated by their case plan, including medication and other mental health services such as play therapy. After completing their final assessment at 14 weeks postbaseline, waitlist controls were eligible to attend PCIT workshops. Families in Jouriles 2001 existing services intervention arm were contacted monthly either in person or by telephone. They encouraged families in the comparison group to use existing community or shelter services. That is, there were no restrictions on families' receipt of services from other sources, and indeed, they encouraged them to make use of the resources available to them. Jouriles 2009 existing services intervention permitted project staff to contact families in the comparison group. These monthly contacts were structured so that these families could receive instrumental and emotional support services similar to those provided to Project Support families. In addition, there were no restrictions placed on comparison families' receipt of services from other sources; indeed, they encouraged them to make use of community resources.

One study used a non-personalised intervention and TAU as their comparator. McCabe 2009 compared their Guiando a Ninos Activos (GANA) personalised arm with a non-personalised PCIT arm. In this non-personalised arm parents were taught skills to establish a nurturing and secure relationship with their child while increasing their child's prosocial behaviour and decreasing negative behaviour. Therapists actively coached parents and terminated when parents demonstrated mastery of the skills and their child's behaviour was within half a SD of the normative mean on the ECBI Intensity scale. McCabe 2009 assigned TAU families to therapists without training in PCIT at the same clinic. The three TAU therapists described their orientations as 'person-centred cognitive behavioural,' 'trauma-focused cognitive behavioural,' and 'family systems' and were allowed complete freedom in the approaches they used.

Six studies used only non-personalised interventions as their comparators. Dadds 1992 compared their outcome of CMT with an ally personalised arm with a non-personalised CMT arm. Parents in the non-personalised intervention received contact with the therapist during group discussion sessions. In the first session parents were involved in a general discussion about issues and special problems for the single parent. Parents then received CMT in mixed groups. As well as receiving technique training, mothers were encouraged to problem-solve for and with each other. In these sessions, mothers were encouraged to act as "therapists" and discuss alternative solutions to remaining problems. Dadds 2019 compared between their AccessEI, a six- to 10-week online therapist-assisted parenting programme, and a face-to-face (FTF) treatment, whereby they received the same programme presented FTF during a one-week treatment. Görtz-Dorten 2019 compared their CBT-based Treatment Program For Children With Aggressive Behavior (THAV) social competence training programme with their educational group play (PLAY) intervention arm. PLAY consisted of three to five children in each group and techniques to activate resources and provided the opportunity to practice prosocial interactions in groups. Greene 2004 compared their Collaborative Problem Solving (CPS) personalised intervention with parenting training (PT). Families assigned to the PT intervention received a 10-week behaviour management programme, as prescribed by the treatment manual, with specified weekly session content. Jones 2014 compared their Technology-Enhanced Helping The Noncompliant Child (TE-HNC) programme to a non-personalised HNC programme. All families received the standard, two-phase HNC programme, as

described above. Parent 2022 also compared their personalised TE-HNC intervention arm with a non-personalised HNC. The HNC intervention followed the same structure as that in Jones 2014, with HNC including weekly FTF therapy sessions (60 minutes), as well as a brief midweek telephone check-in.

Delivery of intervention

In most studies, therapists (graduate psychology students, clinical psychology students, and clinical psychologists) who received training in the content and techniques of the intervention prior to treatment and regular supervision, delivered the intervention. Clinical psychologists and postgraduate psychology trainees were trained and supervised to deliver the intervention with planned session contents in Bor 2002, Dadds 2019, and Parent 2022. Greene 2004 selected experienced, doctoral-level clinical psychologists to deliver the treatment and the psychologists received weekly supervision from the primary investigator. Master's level graduate students were trained in clinical practice for the intervention and participated in weekly supervision in Jones 2014. Clinical psychology graduate students and one clinical psychologist were the therapists delivering the intervention in Jouriles 2001; the therapists received extensive training, were provided with standardised treatment manuals, and received close supervision. Jouriles 2009 selected master's level graduate clinicians and one clinical psychologist as therapists to deliver the intervention. Therapists received comprehensive training and were required to demonstrate their ability to facilitate the intervention in the trial by successfully completing mastery tests, co-facilitation of the intervention, and a training case. The therapists received supervision whilst delivering the intervention. Bilingual students on professional psychology doctoral programmes delivered the intervention in McCabe 2009; they completed 40 hours of training by the principal investigator and received weekly individual supervision. PCIT clinicians delivered the intervention in Mersky 2016. Skilled graduate level psychologists who received comprehensive training in delivering the BFI and the treatment manual facilitated the intervention in Nicholson 1999. Child therapists, or therapists in training, provided the intervention in Görtz-Dorten 2019 and received weekly supervision from a senior child therapist. Psychologists practising in child psychopathology delivered the intervention in Dadds 1992. Markie-Dadds 2006 used a self-directed intervention with weekly telephone contact with a clinical psychologist.

Monitoring treatment fidelity

Eleven studies conducted fidelity checks and monitored adherence to the intervention (Bor 2002; Dadds 2019; Görtz-Dorten 2019; Greene 2004; Jones 2014; Jouriles 2001; Jouriles 2009; Markie-Dadds 2006; McCabe 2009; Nicholson 1999; Parent 2022).

Duration of the intervention

Treatment duration varied between studies. Dadds 1992 conducted six sessions over eight weeks. Although the duration of the intervention was not clearly reported in Dadds 2019, it was reported that parents received six to 10 weekly sessions with a clinician. Markie-Dadds 2006 did not report the duration of the intervention, presumably because the treatment was self-directed. Treatment consisted of weekly or fortnightly, 90- to 120-minute sessions over 10 weeks in Nicholson 1999. Of the studies that delivered interventions to parent–child dyads, Bor 2002 delivered a mean of 14 hours of intervention in 12 sessions of 60 to 90 minutes over a 17-week period; Greene 2004 delivered treatment over seven to 16 weeks, with a mean duration of 11 weeks; Jones 2014 conducted 12-weekly sessions; McCabe 2009 reported that the mean number of sessions was 18; Mersky 2016 provided three days of group training and 14 weeks of in-home services. Görtz-Dorten 2019, Jouriles 2001, and Jouriles 2009 conducted longer interventions than all other studies delivered to parent–child dyads, with treatment lasting 24 weeks in Görtz-Dorten 2019, and up to eight months in Jouriles 2001 and Jouriles 2009. Parent 2022 did not specify the duration of their intervention.

Duration of follow-up

Follow-up duration can be grouped into short term (one month postintervention or less), medium term (greater than one month to less than 12 months postintervention), and long term (12 months postintervention or greater). All studies collected short-term outcome data. Some studies also completed medium-term follow-up assessments (Dadds 1992; Dadds 2019; Greene 2004; Jouriles 2001; Markie-Dadds 2006; Mersky 2016; Parent 2022). Of these studies, two had follow-up assessments at six months (Dadds 1992; Markie-Dadds 2006), one at three months (Dadds 2019), one at four months (Greene 2004), one at four and eight months (Jouriles 2001), one at six weeks (Mersky 2016), and one at three and six months (Parent 2022). Two studies completed a long-term follow-up assessment at 12 months (Bor 2002; Jouriles 2009); Jouriles 2009 collected one-year follow-up data for all intervention groups, whereas Bor 2002 collected long-term outcome data for the personalised intervention group and the comparison intervention group, but not the waitlist control group. We extracted additional data for long-term follow-up (24 months) for Jouriles 2001 from the secondary report of the study (McDonald 2006). We extracted long-term follow-up data for McCabe 2009 from the secondary report of the study (McCabe 2012); such long-term follow-up data were collected over 17 months, with a mean time point of collection of 15 months postintervention.

Outcomes and measures

Primary outcomes

Improvement in child conduct problems or disruptive behaviour

All studies measured child conduct problems or disruptive behaviour. Six studies used the Child Behaviour Checklist (CBCL; Achenbach 2001) (Görtz-Dorten 2019; Jouriles 2001; Jouriles 2009; McCabe 2009; Mersky 2016; Nicholson 1999). Of these six, five studies used the Externalizing subscale (Görtz-Dorten 2019; Jouriles 2001; Jouriles 2009; McCabe 2009; Mersky 2016), whilst one used the Total score (Nicholson 1999). In addition to the CBCL, Görtz-Dorten 2019 used three further measures; the parent- and teacher-rated DISYPS-II Symptom Checklist for Disruptive Behavior Disorders (SCL-DBD; Lochman 2004) to assess all ODD and CD symptoms according to ICD-10 and DSM-IV criteria, the Externalizing subscale of the Teacher Report Form (TRF; Achenbach 1991), and the aggressive behaviour subscale of the Social Problem-Solving Test (SPST; clinician reported). Jouriles 2009 used an oppositional child behaviour measure.

Seven studies used the ECBI (Bor 2002; Jones 2014; Jouriles 2009; Markie-Dadds 2006; McCabe 2009; Mersky 2016; Parent 2022). Five studies used both the Problem and Intensity subscales (Jones 2014; Markie-Dadds 2006; McCabe 2009; Mersky 2016; Parent 2022), one used the Problem scale only (Jouriles 2009), and one used the

Intensity, Problem, Oppositional Defiant, and Conduct subscales (Bor 2002).

In addition to the CBCL and the ECBI, one study, McCabe 2009, used the Early Childhood Inventory (ECI; Sprafkin 1997) to measure ODD and CD, while Nicholson 1999 used interviewer ratings of clinical levels of symptoms for ODD and CD. Four studies used the Parent Daily Report (PDR; Chamberlain 1987) (Bor 2002; Dadds 1992; Markie-Dadds 2006; Nicholson 1999). Bor 2002 used the mean Daily and Target scales, and Markie-Dadds 2006 used the Problem and Parget subscales. In addition to the PDR, Dadds 1992 used the CD subscale of the Revised Behaviour Problem Checklist (RBPC; Quay 1983), to measure child conduct problems.

Of the remaining studies, Dadds 2019 used three measures of improvement in child conduct problems/disruptive behaviour, clinical symptom severity (and diagnosis) of ODD/CD from the DISCAP, the Oppositional Behavior subscale of the Conners' Parent Rating Scale – Revised (CPRS-R; Conners 1998) and the Strengths and Difficulties Questionnaire (SDQ; Goodman 1997) Total difficulties score. Greene 2004 used an ODD Rating Scale (ODDRS; an unpublished scale created by Ross W Greene) and the Clinical Global Impression Scale (CGI) (NIMH 1985).

Any adverse events

None of the included studies explicitly measured adverse events.

Secondary outcomes

Personalised treatment outcomes relevant to each subgroup

Six studies measured personalised treatment outcomes relevant to the subgroup of interest (Bor 2002; Dadds 1992; Görtz-Dorten 2019; Jouriles 2001; Jouriles 2009; Nicholson 1999). Bor 2002 measured inattention using the ECBI factor structure (Burns 2000). Dadds 1992 included four measures of personalised treatment outcomes: the Beck Depression Inventory (BDI) (Beck 1979); Inventory of Socially Supportive Behaviours (ISSB; Barrera 1981); Perceived Social Support from Family (PSS-Fa); and Perceived Social Support from Friends (PSS-Fr; Procidiano 1983). Görtz-Dorten 2019 included three measures of personalised treatment outcomes relevant to the subgroup: the parent- and teacher-reported peer-related aggression subscale of the Questionnaire for Aggressive Behavior of Children (FAVK; Görtz-Dorten 2010), the parent- and teacher-rated SCL-DBD Prosocial Behaviour subscale (Lochman 2004), and the socially competent behaviour subscale of the SPST. Both Jouriles 2001 and Jouriles 2009 measured personalised treatment outcomes using the Symptom Checklist-90 – Revised (SCL-90-R; Derogatis 1976), and Jouriles 2009 also used the Impact Events Scale (IES; Horowitz 1979). Nicholson 1999 used the Parent Problem Checklist (PPC) (Dadds 1991).

Parenting skills and knowledge

Seven studies measured parenting skills and knowledge (Bor 2002; Dadds 1992; Greene 2004; Jouriles 2001; Jouriles 2009; Markie-Dadds 2006; McCabe 2009). Of these, two studies used the Parenting Scale (PS; Arnold 1993), with one using the Total score (Bor 2002), and one using three subscales, Laxness, Over-Reactivity, and Verbosity (Markie-Dadds 2006). Both Bor 2002 and Markie-Dadds 2006 also measured parenting skills and knowledge using the Parenting Sense of Competency scale (PSOC; Gibaud-Wallston 1978), with Bor 2002 using the Total score and Markie-Dadds 2006 using two subscales, Satisfaction and Efficacy. Bor 2002 also measured negative parenting behaviour through observation. Dadds 1992 used the correct implementation scale of the Family Observation Schedule (FOS; Sanders 1982). Greene 2004 measured parenting skills and knowledge using three subscales of the Parent–Child Relationship Inventory (PCRI; Gerard 1994): Limit setting; and Communication and Autonomy. Jouriles 2001 used a direct observation of mothers' child management skills to measure parenting skills and knowledge whereas Jouriles 2009 used the Consistency subscale of the Parenting Dimensions Inventory (PDI; Power 1993). McCabe 2009 measured parenting skills and knowledge using the Parenting Practices Scale (PPS; Strayhorn 1988).

Family functioning

Two studies, Bor 2002 and Markie-Dadds 2006, measured family functioning using the PPC, with Bor 2002 using the Problem subscale, and Markie-Dadds 2006 using both the Problem and Intensity subscales.

Engagement and decreased dropout

Three studies measured engagement and dropout (Jones 2014; McCabe 2009; Parent 2022). Jones 2014 reported session attendance and mid-week call availability, McCabe 2009 reported attendance and dropouts, and Parent 2022 reported attendance.

Educational outcomes

None of the included studies measured educational outcomes.

Funding sources

Eleven studies reported their sources of funding. Five studies received grants from the NIMH (Jones 2014; Jouriles 2001; Jouriles 2009; McCabe 2009; Parent 2022). Three of these also received additional financial support: Jouriles 2001 from the Hogg Foundation for Mental Health and Texas Higher Education Coordinating Boards; Jouriles 2009 from the US Department of Justice; and Parent 2022 from The National Institute of Child Health and Human Development, The National Science Foundation, The National Institute of Minority Health and Health Disparities, and The National Institute on Drug Abuse). Studies received grants from the: Queensland Health and National Health and Medical Research Council for Bor 2002; the National Health and Medical Research Council for Dadds 2019; the School of Child and Adolescent Behavior Therapy at the University Hospital Cologne for Görtz-Dorten 2019; the Theodore and Vada Stanley Foundation for Greene 2004; the National Institute of Health and National Institute of Child Health and Human Development for Mersky 2016; the National Health and Medical Research Council, and Public Health Research and Development Council in the form of a Postdoctoral Fellowship for Nicholson 1999.

Contact with study authors

In order to obtain missing information or data, or both, we contacted the authors of six studies (Dadds 2019; Greene 2004; Jouriles 2001; Jouriles 2009; Markie-Dadds 2006; Nicholson 1999). We received a response from two study authors (Dadds 2019; Nicholson 1999); however, we only received the data requested from Dadds 2019.

Excluded studies

We excluded 592 studies after reading the full-text reports. Reasons for exclusion of included: participants being outside the prespecified age range and not being from a predefined subgroup category (319 reports); non-randomised studies (177 reports); and interventions that were not personalised psychological treatments (96 reports). From these, we selected six studies to list in the Characteristics of excluded studies tables because they did not report data on any of our primary or secondary outcome outcomes (Basile 1993; Costantino 1984; Eckenrode 2010; Eddy 2003; Graham 2015; Oden 1982).

Studies awaiting classification

When assessing full-text reports for eligibility we were unable to access five papers, therefore the studies are awaiting classification. We were able to access only the abstracts of four of these studies (Bloom 1980; Jones 1991; Robbins 1988; Walker 1984). We were unable to access an abstract for Greene 1999, and therefore, are only able to report limited details about this study. We presented the available information in the Characteristics of studies awaiting classification table.

Design

Bloom 1980 randomly assigned participants to four groups and compared the personalised intervention to three control groups (one received a comparison treatment, and two groups did not receive treatment). Three studies were two-arm trials with participants randomly assigned to the personalised intervention or control group (Jones 1991; Robbins 1988; Walker 1984). Jones 1991 and Robbins 1988 did not report the details of the control groups. Walker 1984 compared the personalised intervention to a waitlist control group. Greene 1999 compared operant parent management training and CBT; however, there was no further information reported, and the design is not known.

Participants

The subgroups of children with conduct problems in all five studies were based on co-occurring conditions. The studies targeting subgroups of children with conduct problems and a co-occurring condition included: children with comorbid language impairment (Bloom 1980); children with comorbid bipolar disorder (Greene 1999); impulsive boys (Jones 1991); children hospitalised for severe conduct problems (Robbins 1988); and boys (Walker 1984). Although Greene 1999 and Walker 1984 reported using family treatment models, details of parent involvement were not explicitly reported in the five studies' titles or abstracts.

Bloom 1980 recruited an equal number of boys and girls aged seven to 11 years and 11 months. Jones 1991 recruited boys only, aged nine to 12 years, and Walker 1984 recruited boys only from second to sixth grade (aged seven to 12 years). Robbins 1988 recruited children aged five to 12 years, but did not report the gender of children. Greene 1999 did not include the age and gender of children in the title of the report. None of the five studies reported ethnicity in the titles or abstracts.

Intervention

In Bloom 1980, the personalised intervention was a structured language therapy programme facilitated by speech and language teachers. Jones 1991 delivered a 10-week, multicomponent, cognitive-behavioural group training programme and Walker 1984 delivered a counselling model treatment programme regarding troubled families. Robbins 1988 delivered movement treatment; in this intervention participants attended 45-minute group sessions twice a week and implemented the following three specific movement strategies: change in context, recombination, and substitution, and leader intervention. Robbins 1988 reported that follow-ups were completed two weeks following treatment. The other four studies did not report follow-up duration. Greene 1999 compared operant parent management training and CBT; due to the limited information reported, it is not known which group was the personalised intervention.

Outcome measures

Greene 1999 did not report the outcome measures used in the study title.

Primary outcomes

Improvement in child conduct problems or disruptive behaviour

Bloom 1980 reported using the Jesness Inventory. Jones 1991 did not report the outcome measure used but measured "behavioural adjustment." Robbins 1988 used the Children's Behavior Inventory, calculating both the general level of pathology and the Anger/Hostility Standard Score, and the Aggression Incident Frequency Score, which measured the frequency of aggressive incidents. Walker 1984 used the following measures: Parents' Daily Report, Negative Behavior Score; the Aggression subscale of the Achenbach Child Behavior Checklist; and teacher-reported Daily Behavior Checklist.

Any adverse events

No studies reported measuring adverse events.

Secondary outcomes

Personalised treatment outcomes relevant to each subgroup

Bloom 1980 reported using the Ward Heasley Evaluation of Expressive language, The Wide Range Achievement Test and Coopersmith Self-Esteem Inventory. Jones 1991 measured problem-solving skills; the exact outcome measure was not reported. Walker 1984 used the Conflict subscale of the Family Environment Scale.

Other secondary outcomes

No studies reported data on the other secondary outcomes: parenting skills and knowledge; family functioning; engagement and decreased dropout; and educational outcomes.

Risk of bias in included studies

See the risk of bias graph for judgements about each risk of bias domain presented as percentages across all 13 included studies (Figure 2), and the risk of bias summary for review authors' judgements about each risk of bias item for each included study (Figure 3).

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

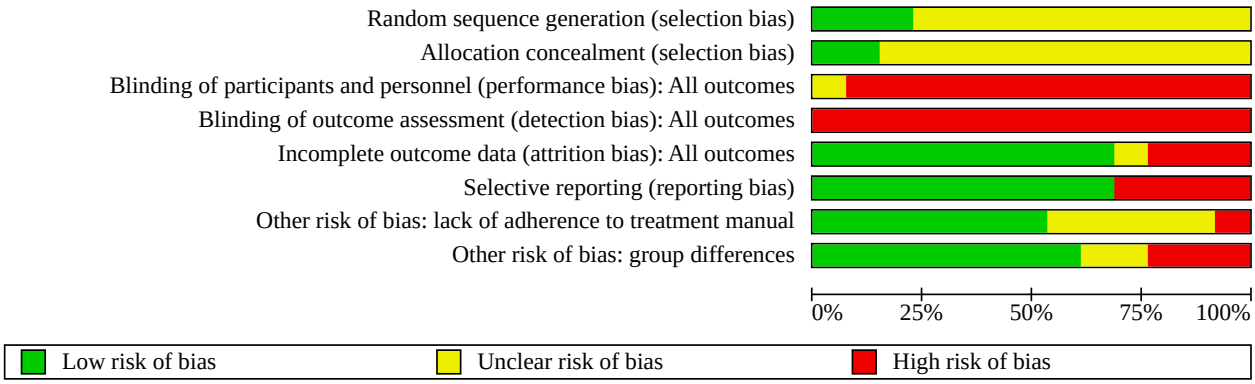
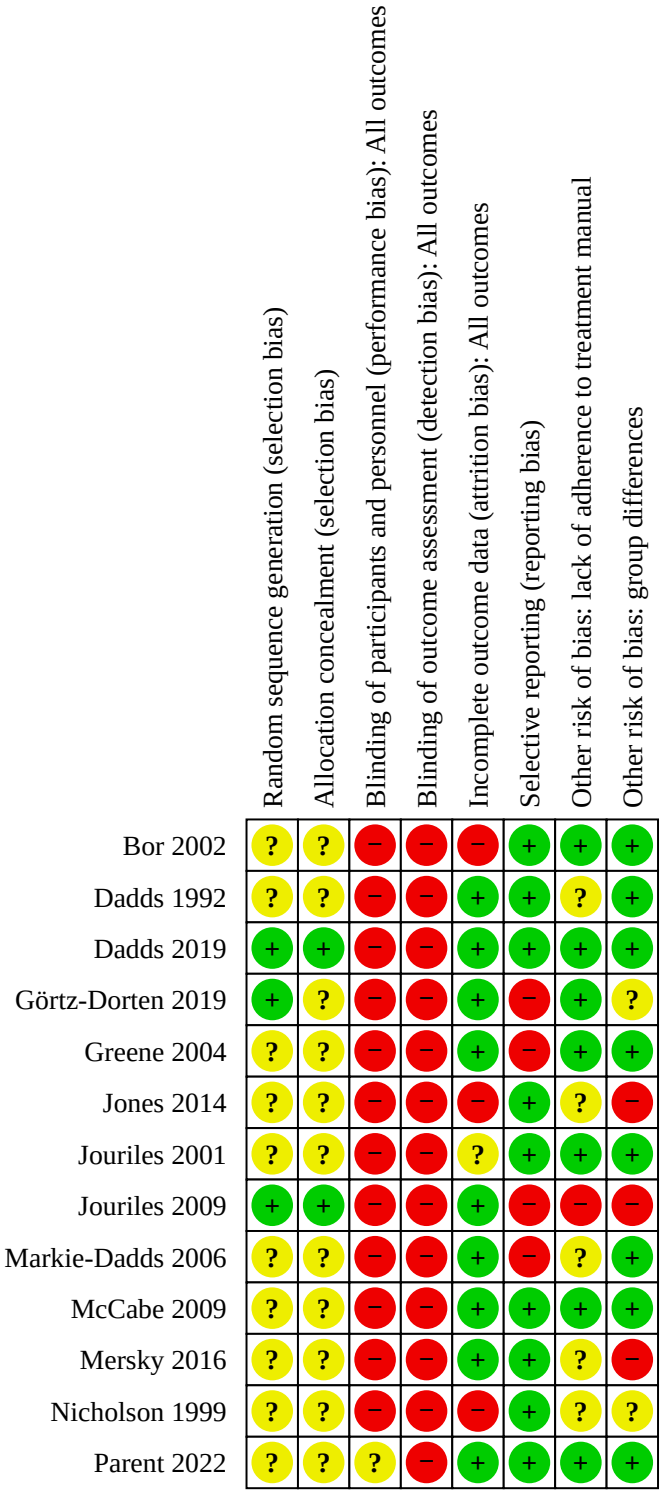


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Allocation

Random sequence generation

Three studies were at low risk of selection bias for sequence generation. Dadds 2019 used a block randomised, sequentially numbered, opaque sealed envelopes method, and a coin-toss method to allocate the order of blocks in the sequence. Görtz-Dorten 2019 also used a block randomisation procedure with random selections from permutations, and Jouriles 2009 randomised participants by allocating a randomisation code to each participating shelter with a random numbers tab. The remaining 10 studies did not report methods of randomisation, and were at unclear risk of bias.

Allocation concealment

Two studies were at low risk of selection bias for allocation (Dadds 2019 and Jouriles 2009 (a project co-ordinator who did not conduct assessments randomised all shelters at the same time)). The remaining 11 studies were at unclear risk of bias because the studies did not report whether they had used any processes to conceal the random allocation.

Blinding

Performance bias

Blinding of the participants and personnel was not possible in 12 of the 13 studies due to the nature of the treatment. The 12 studies where blinding of participants was not possible were at high risk of performance bias. Parent 2022 was at unclear risk of bias as they provided no details of their blinding process.

Detection bias

All studies were at high risk of detection bias because the assessors of the primary outcome measures within each study were carers, teachers, or clinicians who were not blinded to the group allocation.

Incomplete outcome data

Nine studies were at low risk of attrition bias because they adequately addressed incomplete outcome data. Four studies performed intention-to-treat (ITT) analyses (Dadds 2019; Jouriles 2009; McCabe 2009; Mersky 2016). Similarly, Görtz-Dorten 2019 imputed missing values using the expectation maximisation procedure and Greene 2004 reported a low attrition rate of 13% and used last-observation-carried-forward methodology to account for missing data at follow-up. There were no dropouts in Dadds 1992 (all 22 participants initially randomised were analysed at follow-up) or Parent 2022 (all 101 participants initially randomised were analysed at follow-up) and Markie-Dadds 2006 reported low attrition rates (2%).

Three studies were at high risk of attrition bias (Bor 2002; Jones 2014; Nicholson 1999). Bor 2002 only included participants who completed each intervention arm in the analysis, and reported differential attrition based on mothers' reports on the ECBI. The other two studies reported high numbers of dropouts and did not perform ITT analyses (Jones 2014: four participants, two in each group; Nicholson 1999: 30% of families).

One study provided insufficient information about whether missing data at each of the five assessments were accounted for in the analyses and, consequently, was at unclear risk of attrition bias (Jouriles 2001).

Selective reporting

One study had a published online protocol available (Görtz-Dorten 2019).

Nine studies were at low risk of reporting bias as the studies seemed to report all prespecified outcomes (outcome measures and all subscales) (Bor 2002; Dadds 1992; Dadds 2019; Jones 2014; Jouriles 2001; McCabe 2009; Mersky 2016; Nicholson 1999; Parent 2022).

Four studies were at high risk of reporting bias: there were inconsistencies in the planned outcome measures in the protocol and outcome measures reported in Görtz-Dorten 2019; Greene 2004 did not report on four of the seven subscales of the PCRI; Jouriles 2009 did not report the outcome measures data for all time points at which they were assessed; Markie-Dadds 2006 collected outcome measures from both mothers and fathers but only reported analyses of the outcomes completed by mothers.

Other potential sources of bias

For other potential sources of bias, we assessed a lack of adherence to treatment manual and group differences.

Lack of adherence to treatment manual

One study was at high risk of bias due to lack of adherence to their treatment manuals because, despite stating that their undergraduate and postbaccalaureate students who delivered the intervention received extensive training and supervision from either a master's level clinician or a clinical psychologist, an assessment of their proficiency was only detailed as the students having 'completed the training' and being deemed 'adequately trained'. No formal assessment or double coding of fidelity to the treatment manual was reported (Jouriles 2009). Seven studies were at low risk of bias due to lack of adherence to their treatment manuals because all studies reported reviewing their training manuals; establishing reliability with coding systems; assessed the training of therapists or individuals conducting the intervention by providing session role-play, observations, supervision, and feedback; and reported high levels of treatment fidelity which was blind double coded (Bor 2002; Dadds 2019; Görtz-Dorten 2019; Greene 2004; Jouriles 2001; McCabe 2009; Parent 2022). Five studies were at unclear risk of bias due to lack of adherence to their treatment manuals because they provided no information on their treatment manuals (Dadds 1992; Jones 2014; Markie-Dadds 2006; Mersky 2016; Nicholson 1999).

Group differences

Three studies were at high risk of bias due to differences between groups because they stated that they collected demographic information but did not subsequently report any demographic information or comment upon the comparability of their study groups based on the information they stated they gathered (Jones 2014; Jouriles 2009; Mersky 2016). Eight studies were at low risk due to differences between groups because they all reported clear tables detailing the characteristics of both groups for comparability and reported equivalent sociodemographic characteristics for their groups (Bor 2002; Dadds 1992; Dadds 2019; Greene 2004; Jouriles 2001; Markie-Dadds 2006; McCabe 2009; Parent 2022). Two studies were at unclear risk of bias due to group differences because they

did not report information on their groups (Görtz-Dorten 2019; Nicholson 1999).

Effects of interventions

See: **Summary of findings 1** Personalised intervention versus non-personalised intervention for subgroups of children with conduct problems

In the results below, P values are reported exactly, where possible, apart from values less than 0.001, which are reported as P < 0.001. We have reported where exact P values were not provided in the study. For some comparisons, it was not possible to conduct a meta-analysis due to heterogeneity in the participants and outcome measures; therefore, we presented a narrative synthesis.

Comparison 1. Personalised intervention versus non-personalised intervention

All 13 studies (858 participants) compared a personalised intervention to a non-personalised intervention. See Summary of findings 1.

Primary outcomes

Improvement in child conduct problems or disruptive behaviour

Pooled results

Compared with a non-personalised intervention, a personalisation intervention may result in a slight improvement in child conduct problems or disruptive behaviour in the short term when assessed using the ECBI (higher scores indicate greater child behaviour difficulties) Problem subscale but not with the ECBI Intensity subscale (Problem: MD −3.04, 95% CI −6.06 to −0.02; 6 studies, 278 participants; P = 0.05; I^2 = 51%; Analysis 1.1; very low-certainty evidence; Intensity: MD −6.25, 95% CI −16.66 to 4.15; 6 studies, 278 participants; P = 0.24; I^2 = 48%; Analysis 1.2; very low-certainty evidence). However, the evidence was very uncertain. There was moderate heterogeneity for both ECBI subscales (Problem: Chi^2 = 10.23, degrees of freedom (df) = 5 (P = 0.07); I^2 = 51%; Tau^2 = 6.91; Intensity: Chi^2 = 9.65, df = 5 (P = 0.09); I^2 = 48%; Tau^2 = 77.21).

Compared with a non-personalised intervention, a personalised intervention may improve child conduct problems or disruptive behaviour in the medium term when assessed with both the ECBI Problem and Intensity subscales, but the evidence is very uncertain (Problem: MD −5.08, 95% CI −6.03 to −4.14; 3 studies, 186 participants; P < 0.001; I^2 = 0%; Analysis 1.3; very low-certainty evidence; Intensity: MD −8.98, 95% CI −12.95 to −5.01; 3 studies, 186 participants; P < 0.001; I^2 = 0%; Analysis 1.4; very low-certainty evidence). There was no heterogeneity for either analysis.

Three studies (189 participants) assessed improvement in child conduct problems or disruptive behaviour using the Externalizing subscale of the CBCL (higher scores indicate greater behaviour problems) at postintervention. There was no evidence of a difference between the groups in CBCL Externalizing scores in the short term (MD −2.19, 95% CI −6.97 to 2.59; P = 0.37; Analysis 1.5; very low-certainty evidence). Heterogeneity was substantial (Chi^2 = 4.42, df = 2 (P = 0.11); I^2 = 55%; Tau^2 = 9.79).

Single study results

Bor 2002 reported no evidence of a difference between groups on all measures of child behaviour in the short term (PDR mean Daily: t = 0.78; PDR mean Target: t = 0.55; ECBI Intensity: t = −0.02; ECBI Problem: t = 1.01; no analyses reported for the ECBI Oppositional Defiant and Conduct Disorder subscales). There was also no difference in the long term (PDR mean Daily: f < 1; PDR mean Target: f = 4.13; ECBI Intensity: f < 1; ECBI Problem: f < 1; no analyses reported for the ECBI Oppositional Defiant and Conduct Disorder subscales). Using means and SDs, we calculated an MD of 1.42 (95% CI −0.29 to 3.13; 36 participants; Analysis 1.6; very low-certainty evidence) in the short term, and an MD of 1.00 (95% CI −1.43 to 3.43; 32 participants; Analysis 1.7) in the long term for the PDR mean Daily subscale, both of which suggest no group differences. For the PDR mean Target subscale, we calculated an MD of 1.11 (95% CI −0.75 to 2.97; 36 participants; Analysis 1.8; very low-certainty evidence) in the short term, suggesting no group differences, and an MD of 2.23 (95% CI 0.01 to 4.45; 32 participants; Analysis 1.9) in the long term, in favour of the non-personalised intervention. For the ECBI Problem and Intensity subscales, we calculated an MD of −2.54 (95% CI −7.65 to 2.57; 32 participants; Analysis 1.10) and an MD of 3.35 (95% CI −20.68 to 27.38; 32 participants; Analysis 1.11) in the long term, both of which indicate no group differences. For the ECBI Oppositional Defiant subscale, we calculated an MD of 0.99 (95% CI −7.48 to 9.46; 36 participants; Analysis 1.12; very low-certainty evidence) in the short term and an MD of 1.91 (95% CI −6.40 to 10.22; 32 participants; Analysis 1.13) in the long term, suggesting no group differences. Finally, for ECBI Conduct Disorder subscale, we calculated an MD of 1.64 (95% CI −4.37 to 7.65; 36 participants; Analysis 1.14; very low-certainty evidence) in the short term, and an MD of 0.91 (95% CI −4.70 to 6.52; 32 participants; Analysis 1.15) in the long term, again suggesting no group differences.

Dadds 2019 reported no evidence of interaction effects between time and intervention on any of the measures used for improvement in child conduct problems in the short and medium term: clinical symptom severity of ODD/CD using the DISCAP, the Oppositional Behavior subscale of the CPRS-R, and the Strengths and Difficulties Questionnaire Total Difficulties score (SDQ-T). Dadds 2019 reported no evidence of differences in the proportion of children showing clinical improvement or deterioration on child diagnostic status on the DISCAP at short- and medium-term follow-up. Using the data provided and additional data received from Dadds 2019 when requested, we calculated an MD of −0.33 (95% CI −0.87 to 0.22; 113 participants; Analysis 1.16; very low-certainty evidence) at postintervention, suggesting no group differences, and an MD of −0.56 (95% CI −1.10 to −0.03; 105 participants; Analysis 1.17) at medium-term follow-up, in favour of the personalised parent training intervention arm, for clinical symptom severity of ODD/CD. We also calculated an RR of 1.41 (95% CI 0.61 to 3.26; 86 participants; P = 0.42; Analysis 1.18; very low-certainty evidence) for children rated as 'non-clinical or subclinical/no diagnosis' on the DISCAP at postintervention, and an RR of 2.37 (95% CI 0.74 to 7.65; 82 participants; P = 0.15; Analysis 1.19) at medium-term follow-up, with higher scores indicating greater symptom severity. Dadds 2019 measured oppositional behaviour using a subscale of the CPRS-R (range of scores were not reported, but higher scores seemed to indicate greater oppositional behaviour). We calculated an MD of −0.65 (95% CI −2.13 to 0.83; 91 participants; Analysis 1.20; very low-certainty evidence) at postintervention, and −0.34 (95% CI −2.01

to 1.33; 91 participants; Analysis 1.21) at medium-term follow-up, both of which suggest no group differences. Dadds 2019 measured behavioural difficulties using the SDQ-T (scores range from 0 to 40, with higher scores indicating greater emotional/behavioural problems). We calculated an MD of 1.14 (95% CI −1.07 to 3.35; 91 participants; Analysis 1.22; very low-certainty evidence) at postintervention and an MD of 1.52 (95% CI −1.06 to 4.10; 91 participants; Analysis 1.23) at medium-term follow-up, both of which suggest no group differences.

Dadds 1992 reported short- and medium-term data for the PDR Total score (range of scores not reported). The results suggested no evidence for a main effect for the treatment group, and no evidence of an interaction between the treatment group and phase for the PDR Total score, either at postintervention or follow-up. Using the available data, we calculated an MD of −2.40 (95% CI −5.82 to 1.02; 21 participants; Analysis 1.24; very low-certainty evidence) at postintervention, and an MD of 0.20 (95% CI −1.81 to 2.21; 21 participants; Analysis 1.25) at medium-term follow-up, both of which suggest no group differences.

Görtz-Dorten 2019 analysed all results using an analysis of covariance (ANCOVA) and reported evidence for a between-group difference on the CBCL Externalizing subscale at postintervention in favour of the personalised intervention group (P = 0.030; d = −0.54). They used additional measures of child conduct problems and reported evidence for a between-group difference on the following measures at postintervention in favour of the personalised intervention group: parent-rated DISYPS-II SCL-DBD ODD subscale (P = 0.015; d = −0.61) and the clinician-reported SPST Aggressive behaviour subscale (P = 0.025; d = −0.50). They reported no evidence of a difference on the following measures: parent-rated SCL-DBD CD subscale (P = 0.207; d = −0.27), teacher-rated SCL-DBD ODD (P = 0.134; d = −0.34), and CD (P = 0.109; d = −0.36) subscales, and the TRF Externalised subscale (P = 0.282; d = −0.23). Görtz-Dorten 2019 measured ODD and CD using the SCL-DBD. The SCL-DBD measures symptoms according to ICD-10 and DSM-IV criteria, and a higher score indicates a greater number of symptoms. Using this measure, with postintervention data, we calculated an MD of −4.05 (95% −6.25 to −1.85; 91 participants; Analysis 1.26; very low-certainty evidence) for the parent-rated SCL-DBD ODD subscale, in favour of the personalised intervention, and an MD of −1.44 (95% CI −3.94 to 1.06; 91 participants; Analysis 1.27; very low-certainty evidence) for the teacher-rated SCL-DBD ODD subscale, indicating no group differences. We calculated an MD of −0.80 (95% CI −1.90 to 0.30; 91 participants; Analysis 1.28; very low-certainty evidence) for the parent-rated SCL-DBD, CD subscale, which suggests no group differences. We calculated an MD of −1.28 (95% CI −2.22 to −0.34; 91 participants; Analysis 1.29; very low-certainty evidence) for the teacher-rated SCL-DBD CD subscale, in favour of the personalised intervention group. Görtz-Dorten 2019 measured externalising behaviours using the TRF. The TRF measures child behavioural problems, and although the range of scores were not reported, higher scores indicate greater behavioural problems. Using this measure, for the TRF Externalised subscale, we calculated an MD of −2.38 (95% CI −6.57 to 1.81; 91 participants; Analysis 1.30; very low-certainty evidence), suggesting no group differences. Görtz-Dorten 2019 measured child conduct problems or disruptive behaviour using the Aggressive behaviour subscale of the SPST. Although the scale and direction of scores was not reported in the paper, it appears that a higher score indicates greater aggressive behaviour during the role play. We calculated the MD and found evidence that, compared with a non-personalised intervention, a personalised intervention may have little to no effect on child conduct or disruptive behaviour, but the evidence is very uncertain (MD −5.95, 95% CI −13.55 to 1.65; 91 participants; see Analysis 1.31; very low-certainty evidence).

Greene 2004 reported no evidence of a difference between the personalised intervention and comparison intervention groups on the ODD Rating Scale (ODDRS; range of scores not reported; very low-certainty evidence). However, they did find evidence that compared to children in the comparison intervention, children in the personalised intervention improved on the CGI at post-treatment (t (46) = −4.26, P < 0.01; exact P value not reported; very low-certainty evidence) and at four-month follow-up (t (41) = −3.04, P < 0.01; exact P value not reported).

Jouriles 2001 did not report the number of participants in each group in the analyses, and therefore, we were unable to calculate MDs and CIs. In a secondary report of Jouriles 2001, we were able to extract and analyse long-term follow-up data. Jouriles 2001 administered the CBCL Externalizing subscale at postintervention, two medium-term follow-ups(four and eight months), and at long-term (24 months) follow-up. The study authors found no evidence of a difference between the two groups at postintervention or medium-term follow-up (statistical analyses not reported).

Using means and SDs at long-term follow-up, we calculated an MD of −5.50 (95% CI −14.88 to 3.88; 30 participants; see Analysis 1.32), in favour of the personalised intervention group, indicating a reduction in the number of symptoms on the CBCL Externalizing subscale.

Jouriles 2009 did not report SDs, and therefore, we were unable to calculate MDs and CIs. Jouriles 2009 collected data using the CBCL Externalizing subscale and ECBI Intensity subscale at short- and long-term follow-up. At postintervention, externalising problems had decreased more rapidly in the personalised intervention than in the control group (b = 0.23, t (64) = 2.78, P < 0.01; exact P value not reported). The study authors did not report the between-group differences on the CBCL Externalizing subscale in the long term or the analyses of ECBI Problem subscale.

McCabe 2009 suggested that there was no evidence for changes in prescores to postscores on the CBCL Externalizing subscale, and that the ECBI Problem and Intensity subscales did not differ between the personalised intervention and non-personalised intervention groups (statistical analyses not reported). McCabe 2009 also reported evidence that no group differences between the personalised intervention and non-personalised intervention on the CBCL Externalizing subscale and ECBI Problem and Intensity subscales at long-term follow-up existed (collected at a mean time point of 15 months postintervention).

Markie-Dadds 2006 completed short- and medium-term follow-up assessments. The study authors reported evidence of a difference between the personalised intervention group and the non-personalised intervention group on PDR mean Target subscale scores in the short term; this was not maintained at medium-term follow-up. Using available means and SDs, we calculated an MD of −3.86 (95% CI −5.94 to −1.78; 28 participants; Analysis 1.33; very low-certainty evidence) at postintervention, and an MD of −2.01 (95% CI −3.50 to −0.52; 27 participants; Analysis 1.34) at medium-

term follow-up, both which favour the personalised intervention. The study authors also reported no evidence of a difference for PDR mean Problem scores at short- and medium-term follow-up.

Using means and SDs, we calculated an MD of −2.48 (95% CI −5.45 to 0.49; 28 participants; Analysis 1.35; very low-certainty evidence) at postintervention, and an MD of −0.81 (95% CI −3.46 to 1.84; 27 participants; Analysis 1.36) at medium-term follow-up, both of which suggest no group differences. Markie-Dadds 2006 collected outcome data from both mothers and fathers but reported that all father-reported outcomes showed no evidence of differences, and did not provide the analyses or descriptive statistics. Therefore, we were unable to calculate the MDs and SDs for father-reported outcomes in this trial.

Mersky 2016 reported that, for the CBCL Externalizing subscale, there was evidence that the trajectories of the two groups differed when time was modelled as a quadratic function (P = 0.045, effect size (ES) = 0.03). Thus, there was a greater improvement in children in the personalised intervention group compared to the non-personalised intervention group in the short term; however, between the short-term and medium-term follow-ups, the non-personalised intervention group increased in their mean externalising scores while the personalised group scores continued to decrease. For the CBCL Externalizing subscale, we calculated an MD of −5.60 (95% CI −12.16 to 0.96; 52 participants; Analysis 1.37) in the medium-term, in favour of the personalised intervention.

Nicholson 1999 found no evidence of a difference between the personalised intervention and non-personalised intervention groups on any measures of improvement in child conduct problems or personalised treatment outcomes. Due to the lack of group differences, they combined the data from the two interventions and compared them to the control group. They reported no further descriptive statistics or analyses and, as such, we were unable to calculate MDs and CIs for this study.

Any adverse events

None of the trials reported monitoring adverse events.

Secondary outcomes

Personalised treatment outcomes relevant to each subgroup

Bor 2002 reported greater inattentive behaviour in children in the intervention plus adjunctive treatment group, compared to the comparison intervention group, in the short term (no further analyses reported). The study authors also found no evidence of a difference between the intervention group and the control group in the long term (no further analyses reported). Using postintervention means and SDs, we calculated an MD of 3.76 (95% CI −0.05 to 7.57; 36 participants; Analysis 1.38) for the ECBI Inattention subscale, and using means and SDs at one-year follow-up, we calculated an MD of −0.31 (95% CI −5.12 to 4.50; 32 participants; Analysis 1.39), both of which indicate no group differences in the short and long term. There is very low-certainty evidence for short- and long-term personalised treatment outcomes.

Dadds 1992 used four measures to assess personalised treatment outcomes: BDI (range not reported, lower scores indicate fewer depression symptoms); ISSB (range not reported, higher scores indicate greater social support); PSS-Fa (range not reported, higher scores indicate greater social support); and PSS-Fr (range not reported, higher scores indicate greater social support). They found no evidence of a difference between the two groups on any of the outcomes at short- and medium-term follow-up. For the BDI, we calculated an MD of −3.50 (95% CI −9.38 to 2.38; 21 participants; Analysis 1.40) at short-term and an MD of −4.80 (95% CI −10.92 to 1.32; 21 participants; Analysis 1.41) at medium-term follow-up. For the ISSB, we calculated an MD of 4.30 (95% CI −10.01 to 18.61; 21 participants; Analysis 1.42) at short-term and an MD of 4.40 (95% CI −10.01 to 18.81; 21 participants; Analysis 1.43) at medium-term follow-up. For the PSS-Fa, we calculated an MD of 4.60 (95% CI −0.09 to 9.29; 21 participants; Analysis 1.44) at short-term and an MD of 1.70 (95% CI −2.00 to 5.40; 21 participants; Analysis 1.45) at medium-term follow-up. For the PSS-Fr, we calculated an MD of −1.80 (95% CI −7.27 to 3.67; 21 participants; Analysis 1.46) at short-term and an MD of −1.20 (95% CI −7.24 to 4.84; 21 participants; Analysis 1.47) at medium-term follow-up. Thus, we found no evidence for group differences for all the personalised treatment outcomes in Dadds 1992 at short-term (postintervention) and medium-term follow-up.

Görtz-Dorten 2019 reported evidence of between-group differences on the following measures at postintervention, all of which were in favour of the personalised intervention group: the parent-reported, peer-related, FAVK Aggression subscale (P = 0.043; d = −0.46); the parent-reported SCL-DBD Prosocial Behaviour subscale (P = 0.048; d = 0.42), the teacher-reported SCL-DBD Prosocial Behaviour subscale (P = 0.048; d = 0.42), and the SPST Socially Competent Behaviour subscale (P = 0.021; d = 0.66). They found no evidence of a difference for the teacher-reported peer-related FAVK Aggression subscale (P = 0.209; d = −0.24). For the peer-related FAVK Total Aggression subscale, which assesses several maintaining factors of peer-related aggression (range of scores not reported; however, it appears that a higher score indicates greater peer aggression), we used postintervention data and calculated an MD of −5.25 (95% CI −10.04 to −0.46; 91 participants; Analysis 1.48) for the parent-reported FAVK peer-related aggression subscale, in favour of the personalised intervention group, and an MD of −3.50 (95% CI −9.82 to 2.82; 91 participants; Analysis 1.49) for the teacher-reported FAVK peer-related Aggression subscale, suggesting no short-term group differences. For the SCL-DBD Prosocial Behaviour subscale, which measures parent- and teacher-reported prosocial behaviour (range of scores not reported, but a higher score seems to indicate greater prosocial behaviour), we calculated an MD of 2.16 (95% CI −0.19 to 4.51; 91 participants; Analysis 1.50) for the parent-rated prosocial behaviour, suggesting no group differences, and an MD of 3.00 (95% CI 1.07 to 4.93; 91 participants; Analysis 1.51) for teacher-rated prosocial behaviour, in favour of the personalised intervention group at postintervention. For the SPST Socially Competent Behaviour subscale, which measures role-played socially competent behaviours (range of scores not reported, but it appears that a higher score indicates greater socially prosocial behaviours), we calculated an MD of 8.00 (95% CI 3.14 to 12.86; 91 participants; Analysis 1.52), in favour of the personalised intervention group. We graded the certainty of the evidence as very uncertain for short-term personalised treatment outcomes.

Jouriles 2001 reported that there was no evidence of change in parameters of the SCL-90-R (a measure of psychological distress, range of scores not reported) between groups at postintervention or at medium-term follow-up.

Jouriles 2009 reported no evidence of a difference between the intervention and the control group in scores on the IES (range of scores not reported) and SCL-90-R at postintervention and long-term follow-up.

Nicholson 1999 provided insufficient descriptive statistics or analyses and, as such, we were unable to calculate MDs and CIs for this study.

Parenting skills and knowledge

Bor 2002 used three measures to assess parenting skills and knowledge: observations of negative parent behaviour, PS, and PSOC. The study authors found no evidence of a difference in negative parent behaviour, parenting skills, and parenting confidence between the personalised intervention group and the non-personalised intervention group at short- or long-term follow-up (short term: negative parent behaviour: t = 1.76; PS: t = −0.18; PSOC: t = −0.28; long-term: negative parent behaviour: f = 2.94; PS: f < 1; PSOC: f < 1). Using the data provided, we calculated an MD of −1.08 (95% CI −2.71 to 0.55; 36 participants; Analysis 1.53) at postintervention, and an MD of 0.35 (95% CI −0.17 to 0.87; 32 participants; Analysis 1.54) at one-year follow-up for negative parent behaviour; an MD of −0.22 (95% CI −0.71 to 0.27; 36 participants; Analysis 1.55) at short-term follow-up, and an MD of −0.38 (95% CI −0.82 to 0.06; 32 participants; Analysis 1.56) at long-term follow-up for the PS Total score; an MD of 1.36 (95% CI −7.51 to 10.23; 36 participants; Analysis 1.57) at short-term follow-up, and an MD of 6.94 (95% CI −0.21 to 14.09; 32 participants; Analysis 1.58) at long-term follow-up for the PSOC Total score, all of which suggest no group differences. We graded the certainty of the evidence as very low for parenting skills and knowledge.

Dadds 1992 reported no evidence of an interaction between treatment group and phase at postintervention and follow-up for parenting skills and knowledge. Using available data, we calculated an MD of 0.90 (95% CI −12.03 to 13.83; 21 participants; Analysis 1.59) in the short term, and an MD of −1.80 (95% CI −18.29 to 14.69; 21 participants; Analysis 1.60) in the medium term, both of which indicate no group differences.

Greene 2004 reported evidence of improvements in the PCRI (range of scores not reported) Limit Setting subscale (Z score = −3.52, P < 0.01; exact P value not reported) and the PCRI Communication subscale (Z score = 2.27, P < 0.05; exact P value not reported), both of which were in favour of the personalised intervention group at post-treatment. For the PCRI Autonomy subscale, Greene 2004 reported evidence of a time × group interaction (Z score = −1.93, P < 0.06; exact P value not reported) at post-treatment.

Markie-Dadds 2006 found evidence of an effect for mother-reported scores on the PS Laxness subscale (F (3,32) = 3.49, P = 0.03), such that mothers in the personalised intervention reported lower scores than mothers in the non-personalised intervention at postintervention. At follow-up, there was no longer evidence of a difference between the groups. The study authors also found no evidence for differences between groups for the other two subscales of the PS (Over-reactivity and Verbosity) and the PSOC Satisfaction and Efficacy subscales at short- and medium-term follow-up. Using means and SDs reported for the Laxness subscale, we calculated an MD of −0.61 (95% CI −1.12 to −0.10; 28 participants; Analysis 1.61) at short-term follow-up, in favour of the personalised intervention, and an MD of −0.02 (95% CI −0.58 to 0.54; 27 participants; Analysis 1.62) at medium-term follow-up, which suggests no group differences. For the PS Over-reactivity subscale, we calculated an MD of −1.19 (95% CI −1.77 to −0.61; 28 participants; Analysis 1.63) at short-term follow-up, in favour of the personalised intervention, and an MD of −0.34 (95% CI −1.00 to 0.32; 27 participants; Analysis 1.64) at medium-term follow-up, suggesting no group differences. For the PS Verbosity subscale, we calculated an MD of −0.71 (95% CI −1.40 to −0.02; 28 participants; Analysis 1.65) at short-term follow-up, in favour of the personalised intervention, and an MD of −0.30 (95% CI −0.99 to 0.39; 27 participants; Analysis 1.66) at medium-term follow-up, suggesting no group differences. For the PSOC Satisfaction subscale, we calculated an MD of 8.98 (95% CI 3.36 to 14.60; 28 participants; Analysis 1.67) at short-term follow-up, in favour of the personalised intervention, and an MD of 4.60 (95% CI −0.92 to 10.12; 27 participants; Analysis 1.68) at medium-term follow-up, suggesting no group differences. For the PSOC Efficacy subscale, we calculated an MD of 9.48 (95% CI 4.93 to 14.03; 28 participants; Analysis 1.69) at short-term follow-up, and an MD of 5.59 (95% CI 1.64 to 9.54; 27 participants; Analysis 1.70) at medium-term follow-up, both of which were in favour of the personalised intervention. Markie-Dadds 2006 did not report any of the range of score for the above measures of parenting skills and knowledge. We graded the certainty of the evidence as very low for parenting skills and knowledge in the short and medium term.

McCabe 2009 found no evidence of a difference in pretreatment-to-post-treatment mother-reported data on the PPS (range of scores on the scale not reported), between the personalised and non-personalised intervention groups (statistical analyses not reported). Using change-from-baseline-to-postintervention data, we calculated an MD of 7.50 (95% CI −3.72 to 18.72; 40 participants; Analysis 1.71), suggesting no group differences. We graded the certainty of the evidence as very low for short- and medium-term parenting skills and knowledge.

Jouriles 2001 and Jouriles 2009 provided insufficient data for analysis.

Family functioning

Pooled results

Two studies assessed family functioning using the PPC Problem subscale (Bor 2002; Markie-Dadds 2006). The PPC is a 16-item measure designed to assess parental conflict over child-rearing issues. A higher score indicates greater parental conflict.

We pooled the data in a meta-analysis and found no evidence of a difference between the personalised intervention group and the non-personalised intervention group in PPC Problem subscale scores in the short term, but the evidence is very uncertain (MD 0.30, 95% CI −1.23 to 1.83; 64 participants; P = 0.70; I^2 = 0%; Analysis 1.72; very low-certainty evidence).

Single study results

Bor 2002 reported no evidence of a difference in scores on the PPC Problem subscale between groups in the long term (follow-up less than one year). We calculated an MD of −1.33 (95% CI −2.72 to 0.06; 32 participants; Analysis 1.73) in the long term, suggesting no group differences. There is very low-certainty evidence for family functioning.

Markie-Dadds 2006 reported no evidence of a difference between the personalised intervention and control group on mothers' ratings on the PPC Problem and Intensity subscales at both short- and medium-term follow-up. Using the postintervention means and SDs for each group, we calculated an MD of 8.93 (95% CI −8.04 to 25.90; 28 participants; Analysis 1.74) for the PPC Intensity subscale, which suggests no group differences. Using means and SDs at medium-term follow-up, we calculated an MD of −0.75 (95% CI −3.60 to 2.10; 27 participants; Analysis 1.75) for the PPC Problem subscale, and an MD of 15.39 (95% CI −1.63 to 32.41; 27 participants; Analysis 1.76) for the PPC Intensity subscale, both of which also suggest no group differences. There is very low-certainty evidence for family functioning in the short term.

Jouriles 2001 reported a higher mean level of child management skills (observational data) at postintervention in mothers in the personalised intervention compared to the non-personalised intervention (t (110) = −1.91, P = 0.03). The analysis of medium-term data was not reported.

Jouriles 2009 reported no evidence of a difference between the two groups in scores on the Inconsistency subscale of the Parenting Dimensions Inventory (PDI) in the short and long term.

Engagement and decreased dropout

Pooled results

Three trials (156 participants) measured engagement and decreased dropout in relation to session attendance (Jones 2014; McCabe 2009; Parent 2022). Using means and SDs, we calculated an MD of 0.97% (95% CI −0.32 to 2.27; 156 participants; P = 0.14; Analysis 1.77) for the number of sessions attended between the personalised intervention group and the non-personalised intervention group, suggesting no group differences in attendance.

Single study results

Effect sizes reported in Jones 2014 demonstrated that families in the personalised intervention group were more likely to attend weekly sessions (d = 0.88) and participate in mid-week calls (d = 0.63) than families in the non-personalised intervention group. Using means and SDs, we calculated an MD of 35.00% (CI 19.87 to 50.13; 15 participants; Analysis 1.78) for mid-week call availability, which is in favour of the personalised intervention group.

McCabe 2009 reported no evidence of a difference in dropout rates between the personalised intervention and the non-personalised intervention groups. We calculated an RR of 1.36 (95% CI 0.59 to 3.10; 40 participants; Analysis 1.79) for dropout rates, suggesting no group differences. There is very low-certainty evidence for engagement and decreased dropout.

Educational outcomes

None of the included studies reported outcome data related to educational outcomes.

Comparison 2. Personalised intervention versus waitlist control

Four studies (279 participants) compared a personalised intervention group to a waitlist control group (Bor 2002; Markie-Dadds 2006; Mersky 2016; Nicholson 1999).

Primary outcomes

Improvement in child conduct problems or disruptive behaviour

Pooled results

Three studies (119 participants) assessed improvement in child conduct problems using the ECBI Problem and Intensity subscales (higher scores indicate greater child behaviour difficulties) (Bor 2002; Markie-Dadds 2006; Mersky 2016). We found evidence of a difference between the personalised intervention group and the waitlist control group on both subscales in the short term in favour of the personalised intervention (ECBI Problem: MD −11.50, 95% CI −16.73 to −6.27; 119 participants; P < 0.001; I^2 = 69%; Analysis 2.1; very low-certainty evidence; ECBI Intensity: MD −28.44, 95% CI −50.10 to −6.78; 119 participants; P = 0.01; I^2 = 76%; Analysis 2.2; very low-certainty evidence).

Single study results

Bor 2002 reported evidence of an effect for scores on the PDR mean Daily and Target subscales (both P < 0.001). Children in the personalised intervention group showed lower levels of disruptive child behaviour on the PDR than children in the control group at postintervention. Using postintervention means and SDs, we calculated an MD of −5.20 (95% CI −7.78 to −2.62; 42 participants; Analysis 2.3) for the mean Daily subscale, and an MD of −4.42 (95% CI −6.94 to −1.90; 42 participants; Analysis 2.4) for the mean Target subscale. Both results favoured the personalised intervention. The study authors also found evidence of a difference between the groups for the ECBI factor score of Operational Defiance, but found no evidence of a difference for the Conduct Disorder factor score. We calculated an MD of −9.25 (95% CI −16.74 to −1.76; 42 participants; see Analysis 2.5) Operational Defiance, which is in favour of the personalised intervention, and an MD of −3.96 (95% CI −10.17 to 2.25; 42 participants; Analysis 2.6) for Conduct Disorder, which indicates no group differences.

Markie-Dadds 2006 reported that children in the personalised intervention group showed lower levels of disruptive behaviour on the PDR mean Problem (P < 0.001) and Target (P < 0.01) subscales, than children in the control group at postintervention. Using postintervention means and SDs, we calculated an MD of −4.53 (95% CI −7.26 to −1.80; 25 participants; Analysis 2.7) for the PDR Problem subscale, and an MD of −4.43 (95% CI −6.02 to −2.84; 25 participants; Analysis 2.8) for the PDR Target subscale, both of which favour the personalised intervention.

Mersky 2016 reported short-term (postintervention) and medium-term data for the ECBI Problem and Intensity subscales. For both subscales, the study authors observed an intervention × time interaction effect (intensity: P = 0.12, composite ES = 0.11; problem: P < 0.001, composite ES = 0.18), such that scores decreased at a faster rate in the personalised intervention group than in the control group. Using medium-term data, we calculated an MD of −8.90 (95% CI −13.85 to −3.95; 52 participants; Analysis 2.9) for the ECBI Problem subscale, which is in favour of the personalised intervention, and an MD of −15.80 (95% CI −37.01 to 5.41; 52 participants; Analysis 2.10) for the ECBI Intensity subscale, which suggests no group differences. Additionally, the study authors found that CBCL Externalizing scores decreased for both groups, with an omnibus intervention × time interaction effect (P < 0.001, composite ES = 0.17). We calculated an MD of −2.20 (95% CI −8.52 to 4.12; 52 participants; Analysis 2.11) for the CBCL Externalizing

behaviours subscale at short-term follow-up, and an MD of −5.60 (CI −12.16 to 0.96; 52 participants; see Analysis 2.12) at medium-term follow-up, both of which indicate no group differences. There is very low-certainty evidence for improvements in child conduct problems in the short and medium term.

Nicholson 1999 combined data from the personalised intervention group and the non-personalised group, to compare a single 'active intervention' to the control group. Consequently, they did not report means, SDs, and analyses for the personalised intervention versus control. As a result, we were unable to analyse the data from that trial.

Any adverse events

None of the trials reported monitoring adverse events.

Secondary outcomes

Personalised treatment outcomes relevant to each subgroup

Bor 2002 found no between-group differences in personalised treatment outcome assessed with the ECBI Inattention subscale at postintervention. Using available data, we calculated an MD of −0.20 (95% CI −3.73 to 3.33; 42 participants; Analysis 2.13), which also suggests no group differences.

Parenting skills and knowledge

Bor 2002 reported evidence of an effect for time on the PS and PSOC scales, such that, at postintervention, mothers in the personalised intervention group reported lower levels of dysfunctional parenting practices (P < 0.001) and higher levels of parenting satisfaction and competence (P < 0.01) than mothers in the control group (exact P values not reported). The study authors found no effect for observed negative parenting behaviour at postintervention. We calculated an MD of −0.03 (95% CI −0.96 to 0.90; 42 participants; Analysis 2.14), which also suggests no group differences. Additionally, using postintervention means and SDs, we calculated an MD of −0.78 (95% CI −1.23 to −0.33; 42 participants; Analysis 2.15) for the PS Total and 12.07 (95% CI 3.97 to 20.17; 42 participants; Analysis 2.16) for the PSOC Total, both of which favour the personalised intervention.

Markie-Dadds 2006 found that mothers in the personalised intervention reported lower scores on the PS Laxness subscale than mothers in the control group, at short-term follow-up (P < 0.01; exact P value not reported). We calculated an MD of −0.75 (95% CI −1.23 to −0.27; 25 participants; Analysis 2.17), in favour of the personalised intervention. The study authors found no evidence of a difference between the two groups for the PS Over-reactivity or PS Verbosity subscales at postintervention. However, we calculated an MD of −1.05 (95% CI −1.69 to −0.41; 25 participants; Analysis 2.18) for Over-reactivity, and −0.97 (95% CI −1.61 to −0.33; 25 participants; Analysis 2.19) for Verbosity, both of which favour the personalised intervention. Markie-Dadds 2006 also found that mothers in the personalised intervention reported higher levels of parenting satisfaction using the PSOC Satisfaction subscale (P < 0.01) and efficacy using the PSOC Efficacy subscale (P < 0.001) than mothers in the control group in the short term. We calculated an MD of 10.70 (95% CI 5.44 to 15.96; 25 participants; Analysis 2.20) for satisfaction, and 10.04 (95% CI 5.40 to 14.68; 25 participants; Analysis 2.21) for efficacy, both of which favour the personalised intervention group also. The certainty of evidence for parenting skills and knowledge in the short term was very low.

Family functioning

Pooled results

Two studies assessed family functioning using the PPC Problem subscale. We pooled the data in a meta-analysis and found that, compared to a waitlist, a personalised intervention may improve family functioning in the short term (MD −1.82, 95% CI −3.53 to −0.12; 67 participants; P = 0.04; I^2 = 0%; Analysis 2.22; very low-certainty evidence); however, the evidence is very uncertain.

Single study results

Markie-Dadds 2006 found no evidence of a difference between the two groups on the PPC Intensity (t = 7.18) subscale at postintervention. Using postintervention means and SDs, we calculated an MD of −1.29 (95% CI −21.57 to 18.99; 25 participants; Analysis 2.23) for the PPC Intensity subscale, suggesting no group differences. There was very low-certainty evidence for family functioning in the short-term.

Engagement and decreased dropout

None of the studies assessed engagement or dropout.

Educational outcomes

None of the included studies reported outcome data related to educational outcomes.

DISCUSSION

Summary of main results

A comprehensive search of the literature identified 13 RCTs (858 participants) evaluating the effectiveness of a personalised intervention for improving disruptive behaviour in subgroups of children with conduct problems aged between two and 12 years. Studies were disparate in focus, with heterogeneity in the subgroups, interventions, and outcome measures used. We conducted meta-analyses when appropriate; however, given the clinical and methodological heterogeneity in the included trials, we provided a narrative description of most results. There was considerable variation in the subgroups of children with conduct problems, with three studies targeting interventions to children with a co-occurring condition (e.g. emotional dysregulation), two studies targeting interventions to children with a prespecified parent characteristic (e.g. children exposed to intimate partner conflict), and eight studies targeting interventions to children with familial or environmental factors (e.g. children from rural and remote areas). No studies investigated interventions tailored to children with callous unemotional traits or limited prosocial emotion. Eight were two arm trials, whilst the remaining five studies included three groups. All studies with two arms compared a personalised intervention to a non-personalised intervention or TAU. Of the studies with three groups, four included both a non-personalised intervention group and a waitlist control group. All included studies collected short-term outcome data (one-month postintervention or less); seven studies also collected medium-term follow-up data (greater than one month to less than 12 months postintervention), and two studies collected long-term follow-up data (12 months postintervention or greater).

Improvement in child conduct problems or disruptive behaviour

Overall, there was limited, very low-certainty evidence to suggest personalised interventions improved child conduct problems in the short and medium term, compared to non-personalised intervention outcomes. The studies used a range of different measures to assess improvements in child conduct problems or disruptive behaviour and the most commonly used measure was the ECBI. There were inconsistencies between studies with some studies reporting improvements on some measures and others reporting no evidence of differences in improvements. In addition, there was very low-certainty evidence that personalised interventions improved child conduct problems compared to a waitlist control group in both the short and medium term.

Any adverse events

None of the studies explicitly monitored and documented adverse events that occurred during the intervention period. It is unclear whether the lack of reporting of adverse events reflects an absence of adverse events or a failure to report them.

Personalised treatment outcomes relevant to each subgroup

Six studies provided data on personalised treatment outcomes, relevant to the subgroup of children with conduct problems. Of these, four studies found very low-certainty evidence for no evidence of a difference in improvements in personalised treatment outcomes in the short and medium term, compared to both non-personalised interventions and a waitlist control group. One further study found very low-certainty of evidence of greater improvements in personalised treatment outcomes on two of the measures but there were no group differences on three of the measures. Nicholson 1999 provided insufficient descriptive statistics or analyses and, as such, we were unable to calculate MDs and CIs for this study.

Parenting skills and knowledge

Seven studies assessed parenting skills and knowledge. Within single studies, there was limited, very low-certainty evidence to suggest that personalised interventions improved parenting skills and knowledge compared to non-personalised interventions in the short and medium term, with only one study, Markie-Dadds 2006, reporting improvements in the short term. When we compared personalised interventions to a waitlist control group, very low-certainty evidence suggested short-term improvements in only some measures of parenting skills and knowledge (Markie-Dadds 2006: PS Laxness subscale, and PSCS). Jouriles 2001 and Jouriles 2009 provided insufficient data for analysis.

Family functioning

Only two studies reported outcome data for family functioning that could be pooled (Bor 2002; Markie-Dadds 2006). There was very low-certainty evidence suggesting no evidence of a difference in changes to family functioning between the personalised interventions and either comparator (non-personalised interventions and waitlist control group) in the short, medium, and long term within pooled analyses.

Engagement and decreased dropout

Only three studies addressed this outcome, and there were inconsistencies between the results for these studies (Jones 2014; McCabe 2009; Parent 2022). Very low-certainty evidence suggested no evidence for group differences in session attendance in the personalised intervention compared to the non-personalised intervention.

Educational outcomes

None of the included studies reported outcome data related to educational outcomes.

Overall completeness and applicability of evidence

This review includes 13 RCTs. Data were incomplete in several studies for the computation of effect sizes, and it was not possible to obtain these data despite our efforts to contact study authors. Such missing data and small sample sizes decrease the completeness and certainty of the evidence. All studies were conducted in high-income countries, and, therefore, results cannot be generalised more widely than this. In most studies, the majority of the participants were mothers and within those that delivered the personalised intervention to both mothers and fathers, father-reported data were lacking. Consequently, it is unclear whether the results of this review are applicable to fathers.

This review was interested in all subgroups of CDs in children aged two to 12 years, however defined. We have discussed the concept of subgroups throughout the review within the Background, 'Plain language summary', Description of studies, and Discussion sections. As the information in this review was so limited, of poor quality, and heterogeneous, we documented the evidence in the clearest way possible – we sought to provide a review of the overarching concept of personalisation, and an overall assessment of whether personalised intervention versus non-personalised intervention was effective. In addition, we described in the Background section a particular subgroup of children with conduct problems, children with limited prosocial emotion or CU traits, for which there is an emerging evidence base relating to the development of potential new treatments which may be evaluated in future RCTs.

As noted above, included studies targeted a range of subgroups, taking into account a variety of individual and systemic characteristics of children with conduct problems. This heterogeneity in the populations might impact the interpretation of the results. Although studies covered a large number of subgroups, we could have under-represented or missed relevant subgroups, thus reducing the completeness and applicability of the evidence. For example, research has suggested an association between maternal mental health and child conduct problems (Reyno 2006). The relationship between conduct problems and levels of CU traits or LPE has also been investigated (Viding 2005), and the DSM-5 includes a specifier for children and young people with conduct problems and LPE. Thus, we had expected to find at least one RCT delivering personalised treatment to a predefined subgroup of children with conduct problems and CU traits. However, we identified no such trial, perhaps because studies that have investigated effectiveness with children with conduct problems and CU traits have not predefined these children as a subgroup. This may be due to potential ethical concerns related to reluctance to cause distress to parents by defining their children as 'CU' for the purposes of research.

Personalised interventions varied considerably in many aspects, including with regard to the recipient of the intervention, the

theoretical basis and content of the intervention, and the duration of the intervention. Five included studies incorporated the existing gold standard treatment recommendation of parent training (Scott 2009), personalising the parenting programme to the specific subgroup.

All 13 studies measured child conduct problems. However, the completeness and applicability of this evidence may be compromised by the lack of independent outcome measures, as all included studies relied upon self-report measures completed by parents. Furthermore, no studies reported CU traits or LPE. We found limited evidence for the secondary outcomes of parenting skills and knowledge, personalised treatment outcomes, family functioning, and engagement and dropouts. None of the studies reported on adverse events and there was no evidence available on the effect of interventions on educational outcomes. We conducted meta-analyses of the effects of interventions, where possible, and reported individual study level findings otherwise. For all outcomes, the heterogeneity in the measures used might influence the interpretation of the results. Short-term data were available for all 13 included studies and five measured outcomes over the medium term (greater than one month to less than 12 months postintervention) and two over the long term (12 months postintervention or greater). Therefore, there is limited evidence for the longer-term effectiveness of personalised interventions for conduct problems in this review.

Quality of the evidence

Several methodological limitations reduced the certainty of the evidence in all included trials, thereby limiting the conclusions that can be drawn. We rated the certainty of the evidence using GRADE and considered the evidence for all included outcomes to be of very low certainty. We downgraded one level for indirectness, as the samples only consisted of children aged three to nine years. All trials had used small sample sizes, so there was low confidence for precision in our assessments. There was heterogeneity in the population and outcome measures used in the studies, which may account for some of the inconsistencies in the results and the observed statistical heterogeneity. For many outcomes, we downgraded the certainty of the evidence due to inconsistencies in the results reported across studies.

We rated all trials at high risk of bias overall. Most studies did not report sufficient information about the randomisation procedure and allocation concealment, so they were at unclear risk of selection bias. Given the nature of the treatment, it was not possible to blind participants and personnel providing the intervention, so we judged the risk of performance bias to be high. We also considered the risk of detection bias to be high for all trials because the primary outcome measures were completed by carers or teachers who were not blinded to the group allocation. We rated nine trials at low risk of attrition bias and reporting bias because they either adequately addressed incomplete outcome data by conducting ITT analyses or had low or equal (or both) attrition rates across groups and reported all prespecified outcomes. However, it is important to note that some studies reported a high number of dropouts and did not state whether they used ITT analyses. Excluding participants from the analyses may have biased estimates of treatment effects in these studies.

Potential biases in the review process

To minimise bias, we conducted the review in line with the published protocol (Kennedy 2017), and reported any deviations within this review (see Differences between protocol and review). We conducted a comprehensive and extensive search of the literature, which resulted in the inclusion of eligible studies with different populations, interventions, comparisons, and outcomes. Despite our persistent efforts, we were unable to obtain the full texts of five reports; however, based on title and abstract screening, such reports were unlikely to meet the eligibility criteria. At least two review authors independently screened abstracts and full texts in the selection of studies, and we resolved any disagreements by discussion. Two review authors independently and thoroughly extracted data, assessed the risk of bias of each included trial using the Cochrane RoB 1 tool (Higgins 2011), and judged the certainty of evidence for each outcome using GRADE. When necessary, we resolved any differences through discussion and with a third review author. We attempted to obtain any missing, inconsistent, or incomplete data by contacting the study authors, but we were unable to obtain these data from any of the authors contacted. Where raw data were unavailable, we could not include these trials in any pooled analyses or calculate the MD and CIs, and thus reported the results narratively. The missing data in our analyses likely resulted in biased effect estimates in single studies. We included postintervention means and SDs in our meta-analyses if they were reported in the published trial. With the exception of one trial in which we used change-from-baseline data, we used postintervention means and SDs, and follow-up means and SDs for medium-term and long-term outcomes, to calculate effect sizes for single studies. Heterogeneity in baseline scores in studies may have resulted in conservative estimates of effect sizes calculated in this way and may explain why, for some trials, there is a difference between the P values reported and those that we calculated. We reduced the impact of reporting bias by undertaking comprehensive searches of multiple sources and identifying whether all outcomes had been reported by the included studies. We have no conflicts of interest to report.

Agreements and disagreements with other studies or reviews

This review is the first of its type to systematically identify and appraise the evidence for personalised treatments for subgroups of children with conduct problems.

Previous reviews have not investigated subgroups or personalisation and have instead focused on the treatment of child conduct problems, evaluating standard group-based parenting programmes for improving emotional and behavioural adjustment in young children (Barlow 2016), improving early-onset conduct problems in children aged three to 12 years (Furlong 2012), and improving conduct problems in older children and adolescents (Woolfenden 2001). All three reviews have provided evidence that standard group-based parenting programmes can improve emotional and behavioural adjustment, reduce conduct problems, and reduce the time spent by juvenile delinquents with conduct problems in institutions. The findings from the three existing Cochrane Reviews focusing on the treatment of child conduct problems call for more research in the area, a conclusion our review reinforces.

AUTHORS' CONCLUSIONS

Implications for practice

The results of this review are limited and the evidence of very low-certainty. Overall, the current evidence for the effectiveness of personalised interventions for subgroups of children with conduct problems does not allow us to draw definite conclusions for the following child and parent outcomes: child conduct problems, personalised outcomes, parenting skills and knowledge, family functioning, and engagement. In the absence of data, it was not possible to determine the influence of personalised interventions on educational outcomes and adverse events. It is also important to note that this review only focused on randomised controlled trials that prespecified that the intervention was personalised to specific subgroups of children with conduct problems. In addition, the review was restricted to children aged two to 12 years.

Personalised interventions targeting children with callous-unemotional traits or limited prosocial emotion were notable in their absence, despite a body of research suggesting that this is an important subgroup of children with conduct problems who may benefit from a personalised approach. Our search strategy did not specifically include these terms and it may be advisable to do so in future searches.

In summary, there is currently insufficient evidence to reach any firm conclusions regarding the effectiveness of personalised interventions, adapted or developed, for subgroups of children with conduct problems.

Implications for research

The current evidence for the effectiveness of personalised interventions for subgroups of children with conduct problems comprises studies at high risk of bias, compromising their ability to detect true differences. As such, the evidence base would benefit from large-scale randomised controlled trials that address several key issues, as discussed below, before any firm conclusions can be made regarding the effectiveness of personalised interventions, adapted or developed, for subgroups of children with conduct problems.

There was considerable variability between trials in the measures used to assess conduct problems and therefore, future research would benefit from consensus regarding the most appropriate measures, in order to facilitate cross-study comparisons. Future research should consider selecting outcome measures that are completed by blinded assessors to provide more objective and robust measures of treatment effectiveness. Furthermore, it would be beneficial for future randomised controlled trials to assess the effectiveness of personalised treatments in relation to outcomes about which we know relatively little, such as parenting skills and knowledge, family functioning, and educational outcomes. None of the included studies reported monitoring adverse events, so it will be important for future trials to assess these. Implementation of a standardised method of recording and reporting adverse events should enable more consistent reporting. Additional important outcomes to include in future trials are dropouts and engagement, as this is relevant to the effectiveness of any intervention, and economic evaluations.

It is important for future trials to provide sample size calculations to demonstrate that the trial is sufficiently powered to detect differences between groups. Furthermore, fathers should be encouraged to participate in trials to improve generalisability. In order to reduce the risk of attrition bias, authors of future studies should consider methods to impute missing outcome data in their analyses, and report the characteristics of, and reasons for, the missing data.

As persistent conduct problems predict a range of adverse long-term outcomes, future research should investigate the medium- and long-term effectiveness of personalised treatments for children with conduct problems. In particular, it is important for future trials to measure outcomes over the medium and long term in order to assess changes in behaviour in children with conduct problems and, if applicable, changes in parents' behaviours, skills, and knowledge. Evidence concerning the longer-term effectiveness of treatments will improve understanding of the maintenance of effects over time and identify whether further input is required at a later date for the child or parent (or both) who received the intervention.

All studies were conducted in high-income countries, and, therefore, it will be important to investigate studies conducted in middle- and low-income countries. Furthermore, as this review only identified studies that included children from age three to nine years, there is a lack of evidence for children between ages nine and 12 years. It is important to determine whether the personalised interventions would be effective for children in this older age group. It will also be useful for future research to consider the cost-effectiveness of personalised interventions and to include economic analyses.

In summary, large-scale and well-designed, high-quality trials are required to address randomisation procedures, attrition, analysis of missing data, sample size, and treatment fidelity. The lack of medium- and long-term follow-up data emphasise the need for further research to assess the long-term effectiveness of interventions.

ACKNOWLEDGEMENTS

We would like to thank the Managing Editor, Dr Joanne Duffield, and Deputy Managing Editor, Dr Sarah Davies, of Cochrane Developmental, Psychosocial and Learning Problems (CDPLP) for their helpful suggestions, guidance, and advice. We would also like to thank Margaret Anderson (Information Specialist, CDLP) who helped in the construction, development, and completion of our search strategy.

The CRG Editorial Team are grateful to the following peer reviewers for their time and comments: Mohsen Ghaffari Darab, Iran; Sturla Fossum, RKBU, UiT The Arctic University in Norway, Norway; and Prof John E Lochman, PhD, ABPP, The University of Alabama, USA.

REFERENCES

References to studies included in this review

Bor 2002 *{published data only}*

Bor W, Sanders MR, Markie-Dadds C. The effects of the Triple P-Positive Parenting Program on preschool children with co-occurring disruptive behavior and attentional/hyperactive difficulties. *Journal of Abnormal Child Psychology* 2002;**30**(6):571-87. [DOI: 10.1023/A:1020807613155] [PMID: 12481972]

Dadds 1992 *{published data only}*

Dadds MR, McHugh TA. Social support and treatment outcome in behavioral family therapy for child conduct problems. *Journal of Consulting and Clinical Psychology* 1992;**60**(2):252-9. [DOI: 10.1037//0022-006x.60.2.252] [PMID: 1592955]

Dadds 2019 *{published data only}*

* Dadds MR, Thai C, Mendoza Diaz A, Broderick J, Moul C, Tully LA, et al. Therapist-assisted online treatment for child conduct problems in rural and urban families: two randomized controlled trials. *Journal of Consulting and Clinical Psychology* 2019;**87**(8):706-19. [DOI: 10.1037/ccp0000419] [PMID: 31204839]

Dadds MR. Therapist-Assisted Online Treatment for Child Conduct Problems in Rural and Urban Families: Two Randomized Controlled Trials (Dadds 2019) – additional information required [personal communication]. Email to: E Hogg 27 August 2020.

Hogg E. Therapist-Assisted Online Treatment for Child Conduct Problems in Rural and Urban Families: Two Randomized Controlled Trials (Dadds 2019) – additional information required [personal communication]. Email to: MR Dadds 25 August 2020.

Görtz-Dorten 2019 *{published data only}*

Görtz-Dorten A, Benesch C, Berk-Pawlitzek E, Faber M, Hautmann C, Hellmich M, et al. Efficacy of individualized social competence training for children with oppositional defiant disorders/conduct disorders: a randomized controlled trial with an active control group. *European Child & Adolescent Psychiatry* 2019;**28**(2):165-75. [DOI: 10.1007/s00787-018-1144-x] [PMID: 29594368]

Greene 2004 *{published data only}*

* Greene RW, Ablon JS, Goring JC, Raezer-Blakely L, Markey J, Monuteaux MC, et al. Effectiveness of collaborative problem solving in affectively dysregulated children with oppositional-defiant disorder: initial findings. *Journal of Consulting and Clinical Psychology* 2004;**72**(6):1157-64. [DOI: 10.1037/0022-006X.72.6.1157] [PMID: 15612861]

Hogg E. Effectiveness of collaborative problem solving in affectively dysregulated children with oppositional-defiant disorder: initial findings (Greene 2004) – additional information required [personal communication]. Email to: RW Greene 23 January 2018.

Jones 2014 *{published data only}*

Jones DJ, Forehand R, Cuellar J, Parent J, Honeycutt A, Khavjou O, et al. Technology-enhanced program for child disruptive behavior disorders: development and pilot randomized control trial. *Journal of Clinical Child & Adolescent Psychology* 2014;**43**(1):88-101. [DOI: 10.1080/15374416.2013.822308] [PMCID: PMC3871925] [PMID: 23924046]

Jouriles 2001 *{published data only}*

Hogg E. Reducing conduct problems among children of battered women (Jouriles 2001) – additional information required [personal communication]. Email to: EN Jouriles 23 January 2018.

* Jouriles EN, McDonald R, Spiller L, Norwood WD, Swank PR, Stephens N, et al. Reducing conduct problems among children of battered women. *Journal of Consulting and Clinical Psychology* 2001;**69**(5):774-85. [DOI: 10.1037//0022-006X.69.5.774]

McDonald R, Jouriles EN, Skopp NA. Reducing conduct problems among children brought to women's shelters: intervention effects 24 months following termination of services. *Journal of Family Psychology* 2006;**20**(1):127-36. [DOI: 10.1037/0893-3200.20.1.127]

Jouriles 2009 *{published data only}*

Hogg E. Reducing conduct problems among children exposed to intimate partner violence: a randomized clinical trial examining effects of Project Support (Jouriles 2009) – additional information required [personal communication]. Email to: EN Jouriles 23 January 2018.

* Jouriles EN, McDonald R, Rosenfield D, Stephens N, Corbitt-Shindler D, Miller PC. Reducing conduct problems among children exposed to intimate partner violence: a randomized clinical trial examining effects of Project Support. *Journal of Consulting and Clinical Psychology* 2009;**77**(4):705-17. [DOI: 10.1037/a0015994] [PMID: 19634963]

Markie-Dadds 2006 *{published data only}*

Hogg E. A controlled evaluation of an enhanced self-directed behavioural family intervention for parents of children with conduct problems in rural and remote areas (Markie-Dadds 2006) – additional information required [personal communication]. Email to: MR Sanders 23 January 2018.

* Markie-Dadds C, Sanders MR. A controlled evaluation of an enhanced self-directed behavioural family intervention for parents of children with conduct problems in rural and remote areas. *Behaviour Change* 2006;**23**(1):55-72. [DOI: 10.1375/bech.23.1.55]

McCabe 2009 *{published data only}*

McCabe K, Yeh M, Lau A, Argote CB. Parent–child interaction therapy for Mexican Americans: results of a pilot randomized clinical trial at follow-up. *Behavior Therapy* 2012;**43**(3):606-18. [DOI: 10.1016/j.beth.2011.11.001] [PMCID: PMC7194395] [PMID: 22697448]

* McCabe K, Yeh M. Parent–child interaction therapy for Mexican Americans: a randomized clinical trial. *Journal of Clinical Child & Adolescent Psychology* 2009;**38**(5):753-9. [DOI: 10.1080/15374410903103544] [PMID: 20183659]

Mersky 2016 *{published data only}*

Mersky JP, Topitzes J, Grant-Savela SD, Brondino MJ, McNeil CB. Adapting parent–child interaction therapy to foster care. *Research on Social Work Practice* 2016;**26**(2):157-67. [DOI: 10.1177/1049731514543023]

Nicholson 1999 *{published data only}*

Hogg E. Randomized controlled trial of behavioral family intervention for the treatment of child behavior problems in stepfamilies (Nicholson 1999) – additional information required [personal communication]. Email to: JM Nicholson 23 January 2018.

* Nicholson JM, Sanders MR. Randomized controlled trial of behavioral family intervention for the treatment of child behavior problems in stepfamilies. *Journal of Divorce & Remarriage* 1999;**30**(3-4):1-23. [DOI: 10.1300/J087v30n03_01]

Nicholson JM. Randomized controlled trial of behavioral family intervention for the treatment of child behavior problems in stepfamilies (Nicholson 1999) – additional information required [personal communication]. Email to: E Hogg 25 January 2018.

Parent 2022 *{published data only}***10.1111/jcpp.13554**

* Parent J, Anton MT, Loiselle R, Highlander A, Bresland N, Forehand R, et al. A randomized controlled trial of technology-enhanced behavioural parent training: sustained parent skill use and child outcomes at follow-up. *Journal of Child Psychology and Psychiatry* 2022;**63**(9):992-1001. [DOI: 10.1111/jcpp.13554] [PMCID: PMC9177891 (available on 2023-09-01)] [PMID: 34888861]

References to studies excluded from this review

Basile 1993 *{published data only}*

Basile VC. Effects of Physical Exercise and a Mastery Task on Disruptive Behaviors and Self-Concept [PhD thesis]. Long Island (NY): Hofstra University, 1993.

Costantino 1984 *{published data only}*

Costantino G, Malgady RG, Rogler LH. Cuentos Folkloricos as a therapeutic modality with Puerto Rican children. *Hispanic Journal of Behavioral Sciences* 1984;**6**(2):169-78. [DOI: 10.1177/07399863840062007]

Eckenrode 2010 *{published data only}*

Eckenrode J, Campa M, Luckey DW, Henderson CR Jr, Cole R, Kitzman H, et al. Long-term effects of prenatal and infancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial [erratum in: Archives of Pediatrics & Adolescent Medicine 2010;164(5):424]. *Archives of Pediatrics & Adolescent Medicine* 2010;**164**(1):9-15. [DOI: 10.1001/archpediatrics.2009.240] [PMID: 20048236]

Eddy 2003 *{published data only}*

Eddy JM, Reid JB, Stoolmiller M, Fetrow RA. Outcomes during middle school for an elementary school-based preventive intervention for conduct problems: follow-up results from a randomized trial. *Behavior Therapy* 2003;**34**(4):535-52. [DOI: 10.1016/S0005-7894(03)80034-5]

Graham 2015 *{published data only}*

Graham S, Taylor A, Hudley C. A motivational intervention for African American boys labeled as aggressive. *Urban Education* 2015;**50**(2):194-224. [DOI: 10.1177/0042085914549364]

Oden 1982 *{published data only}*

Oden S. The applicability of social skills training research. *Child & Youth Services* 1982;**5**(3-4):75-90. [DOI: 10.1300/J024v05n03_06]

References to studies awaiting assessment

Bloom 1980 *{published data only}*

Bloom RM. The Impact of Psycholinguistic Intervention on Variables Associated with Behavioral Disorders in a School Age Population [EdD thesis]. Detroit (MI): Wayne State University, 1980.

Greene 1999 *{published data only}***CN-00656011**

Greene R. Operant parent management training versus cognitive behavioral therapy for symptoms and behavioral problems in children with bipolar disorder. Stanley Foundation Research Awards – 1999 Research Award Recipients 1999.

Jones 1991 *{published data only}*

Jones JJ. The Effects of a Multicomponent Cognitive-Behavioral Group Intervention on the Generalization of Self-Control Behaviors Across Settings [PhD thesis]. Austin (TX): University of Texas, 1991.

Robbins 1988 *{published data only}*

Robbins BD. The Effectiveness of Movement Strategies in Reducing Physically Aggressive Behaviors in Five- to Twelve-Year-Old Children Hospitalized for Severe Conduct Disorders [PhD thesis]. New York (NY): New York University, 1988.

Walker 1984 *{published data only}*

Walker JM. A Study of the Effectiveness of Social Learning Family Therapy for Reducing Aggressive Behavior in Boys [PhD thesis]. Terre Haute (IN): Indiana State University, 1984.

Additional references

Achenbach 1991

Achenbach TM, editor. Manual for the Teacher Report Form and 1991 Profile. Burlington (VT): University of Vermont, Department of Psychiatry, 1991.

Achenbach 2001

Achenbach TM, Rescorla LA, editor(s). Manual for the ASEBA School-Age Forms & Profiles. Burlington (VT): University of Vermont, 2001.

APA 2013

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 5th edition. Arlington (VA): American Psychiatric Association, 2013.

Arnold 1993

Arnold DS, O'Leary SG, Wolff LS, Acker MM. The Parenting Scale: a measure of dysfunctional parenting in discipline situations. *Psychological Assessment* 1993;**5**(2):137-44. [DOI: 10.1037/1040-3590.5.2.137]

Bakker 2017

Bakker MJ, Greven CU, Buitelaar JK, Glennon JC. Practitioner review: psychological treatments for children and adolescents with conduct disorder problems – a systematic review and meta-analysis. *Journal of Child Psychology and Psychiatry* 2017;**58**(1):4-18. [DOI: 10.1111/jcpp.12590] [PMID: 27501434]

Barlow 2016

Barlow J, Bergman H, Kornør H, Wei Y, Bennett C. Group-based parent training programmes for improving emotional and behavioural adjustment in young children. *Cochrane Database of Systematic Reviews* 2016, Issue 8. Art. No: CD003680. [DOI: 10.1002/14651858.CD003680.pub3] [PMCID: PMC6797064] [PMID: 27478983]

Barrera 1981

Barrera M Jr, Sandler IN, Ramsay TB. Preliminary development of a scale of social support: studies on college students. *American Journal of Community Psychology* 1981;**9**(4):435-47. [DOI: 10.1007/BF00918174]

Bayley 2006

Bayley N, editor. Bayley Scales of Infant and Toddler Development (Bayley-III). 3rd edition. San Antonio (TX): Pearson Psychological Corporation, 2006.

Beauchaine 2005

Beauchaine TP, Webster-Stratton C, Reid MJ. Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: a latent growth curve analysis. *Journal of Consulting and Clinical Psychology* 2005;**73**(3):371-88. [DOI: 10.1037/0022-006X.73.3.371] [PMID: 15982136]

Beck 1979

Beck AT, Rush AJ, Shaw BF, Emery G, editor(s). Cognitive Therapy of Depression. New York (NY): Guilford Press, 1979.

Beck 1996

Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment* 1996;**67**(3):588-97. [DOI: 10.1207/s15327752jpa6703_13] [PMID: 8991972]

Borrelli 2005

Borrelli B, Speinwall D, Emst D, Bellg AJ, Czajkowski S, Breger R, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of Consulting and Clinical Psychology* 2005;**73**(5):852-60. [DOI: 10.1037/0022-006X.73.5.852] [PMID: 16287385]

Boyce 2002

Boyce WT, Essex MJ, Woodward HR, Measelle JR, Ablow JC, Kupfer DJ. The confluence of mental, physical, social and academic difficulties in middle childhood. I: exploring the "headwaters" of early life morbidities. *Journal of the American Academy of Child & Adolescent Psychiatry* 2002;**41**(5):580-7. [DOI: 10.1097/00004583-200205000-00016] [PMID: 12014791]

Burns 2000

Burns GL, Patterson DR. Factor structure of the Eyberg Child Behavior Inventory: a parent rating scale of oppositional defiant behavior towards adults, inattentive behavior, and conduct problems. *Journal of Clinical Child Psychology* 2000;**29**(4):569-77. [DOI: 10.1207/S15374424JCCP2904_9] [PMID: 11126634]

Chacko 2012

Chacko A, Wymbs BT, Chimiklis A, Wymbs FA, Pelham WE. Evaluating a comprehensive strategy to improve engagement to group-based behavioral parent training for high risk families of children with ADHD. *Journal of Abnormal Child Psychology* 2012;**40**(8):1351-62. [DOI: 10.1007/s10802-012-9666-z] [PMCID: PMC3821981] [PMID: 22802072]

Chacko 2016

Chacko A, Jensen SA, Lowry LS, Cornwell M, Chimklis A, Chan E, et al. Engagement in behavioral parent training: review of the literature and implications for practice. *Clinical Child and Family Psychology Review* 2016;**19**(3):204-15. [DOI: 10.1007/s10567-016-0205-2] [PMID: 27311693]

Chamberlain 1987

Chamberlain P, Reid JB. Parent observation and report of child symptoms. *Behavioral Assessment* 1987;**9**(1):97-109. [URL: psycnet.apa.org/record/1988-12602-001]

Chronis 2007

Chronis AM, Lahey BB, Pelham WE Jr, Williams SH, Baumann BL, Kipp H, et al. Maternal depression and early positive parenting predict future conduct problems in young children with attention-deficit/hyperactivity disorder. *Developmental Psychology* 2007;**43**(1):70-82. [DOI: 10.1037/0012-1649.43.1.70] [PMID: 17201509]

Chronis-Tuscano 2011

Chronis-Tuscano A, O'Brien KA, Johnston C, Jones HA, Clarke TL, Raggi VL, et al. The relation between maternal ADHD symptoms & improvement in child behaviour following brief behavioral parent training is mediated by change in negative parenting. *Journal of Abnormal Child Psychology* 2011;**39**(7):1047-57. [DOI: 10.1007/s10802-011-9518-2] [PMCID: PMC3715311] [PMID: 21537894]

Collishaw 2004

Collishaw S, Maughan B, Goodman R, Pickles A. Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry* 2004;**45**(8):1350-62. [DOI: 10.1111/j.1469-7610.2004.00335.x] [PMID: 15482496]

Conners 1994

Conners CK. The Conners Rating Scales: use in clinical assessment, treatment planning and research. In: Maruish ME, editors(s). The Use of Psychological Testing for Treatment Planning and Outcome Assessment. Hillsdale (NJ): Lawrence Erlbaum, 1994:467-97.

Conners 1998

Conners CK, Sitarenios G, Parker JD, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology* 1998;**26**(4):257-68. [DOI: 10.1023/a:1022602400621] [PMID: 9700518]

Copeland 2009

Copeland WE, Shanahan L, Costello EJ, Angold A. Childhood and adolescent psychiatric disorders as predictors of young adult disorders. *Archives of General Psychiatry* 2009;**66**(7):764-72. [DOI: 10.1001/archgenpsychiatry.2009.85] [PMCID: PMC2891142] [PMID: 19581568]

Dadds 1991

Dadds MR, Powell MB. The relationship of interparental conflict and global marital adjustment to aggression, anxiety, and immaturity in aggressive and nonclinic children. *Journal of Abnormal Child Psychology* 1991;**19**(5):553-67. [DOI: 10.1007/BF00925820] [PMID: 1770185]

Dadds 2006

Dadds MR, Hawes D, editor(s). Integrated Family Intervention for Child Conduct Problems: a Behaviour-Attachment-Systems Intervention for Parents. Bowen Hills (QLD): Australian Academic Press, 2006.

Derogatis 1976

Derogatis LR, Rickels K, Rock AF. The SCL-90 and the MMPI: a step in the validation of a new self-report scale. *British Journal of Psychiatry* 1976;**128**:280-9. [DOI: 10.1192/bjp.128.3.280] [PMID: 1252693]

Dretzke 2009

Dretzke J, Davenport C, Frew E, Barlow J, Stewart-Brown S, Bayliss S, et al. 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* 2009;**3**(1):7. [DOI: 10.1186/1753-2000-3-7] [PMCID: PMC2660289] [PMID: 19261188]

Drugli 2010

Drugli MB, Fossum S, Larsson B, Morch WT. Characteristics of young children with persistent conduct problems 1 year after treatment with the Incredible Years program. *European Child & Adolescent Psychiatry* 2010;**19**(7):559-65. [DOI: 10.1007/s00787-009-0083-y] [PMID: 20012877]

Egger 1997

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629-34. [DOI: 10.1136/bmj.315.7109.629] [PMCID: PMC2127453] [PMID: 9310563]

Epstein 1983

Epstein NB, Baldwin LM, Bishop DS. The McMaster Family Assessment Device. *Journal of Marital and Family Therapy* 1983;**9**(2):171-80. [DOI: 10.1111/j.1752-0606.1983.tb01497.x]

Epstein 2015

Epstein RA, Fonnesbeck C, Potter S, Rizzone KH, McPheeters M. Psychosocial interventions for child disruptive behaviors: a meta-analysis. *Pediatrics* 2015;**136**(5):947-60. [DOI: 10.1542/peds.2015-2577] [PMID: 26482672]

Erskine 2014

Erskine HE, Ferrari AJ, Polanczyk GV, Moffit TE, Murray CJ, Vos T, et al. The global burden of conduct disorder and attention-deficit/hyperactivity disorder in 2010. *Journal of Child Psychology and Psychiatry* 2014;**55**(4):328-36. [DOI: 10.1111/jcpp.12186] [PMID: 24447211]

Essex 2002

Essex MJ, Boyce WT, Goldstein LH, Armstrong JM, Kraemer HC, Kupfer DJ, MacArthur Assessment Battery Working Group. The confluence of mental, physical, social and academic difficulties in middle childhood. II: developing the MacArthur Health and Behavior Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry* 2002;**41**(5):588-603. [DOI: 10.1097/00004583-200205000-00017] [PMID: 12014792]

Evans 2017

Evans SC, Burke JD, Roberts MC, Fite PJ, Lochman JE, de la Peña FR, et al. Irritability in child and adolescent psychopathology: an integrative review for ICD-11. *Clinical Psychology Review* 2017;**53**:29-45. [DOI: 10.1016/j.cpr.2017.01.004] [PMID: 28192774]

Eyberg 1978

Eyberg SM, Ross AW. Assessment of child behavior problems: the validation of a new inventory. *Journal of Clinical Child Psychology* 1978;**7**(2):113-6. [DOI: 10.1080/15374417809532835]

Fairchild 2019

Fairchild G, Hawes DJ, Frick PJ, Copeland WE, Odgers CL, Franke B, et al. Conduct disorder. *Nature Reviews. Disease Primers* 2019;**5**(1):43. [DOI: 10.1038/s41572-019-0095-y] [PMID: 31249310]

Fernandez 2011

Fernandez MA, Butler AM, Eyberg SM. Treatment outcome for low socioeconomic status African American families in parent–child interaction therapy: a pilot study. *Child & Family Behavior Therapy* 2011;**33**(1):32-48. [DOI: 10.1080/07317107.2011.545011]

Fisher 2015

Fisher AJ, Bosley HG. Personalized assessment and treatment of depression. *Current Opinion in Psychology* 2015;**4**:67-74. [DOI: 10.1016/j.copsyc.2015.03.031]

Fossum 2009

Fossum S, Handegård BH, Martinussen M, Mørch W-T. Psychosocial interventions for disruptive and aggressive behaviour in children and adolescents: a meta-analysis.

European Child & Adolescent Psychiatry 2009;**17**:438-51. [DOI: 0.1007/s00787-008-0686-8] [PMID: 18427863]

Frick 2004

Frick PJ. Inventory of Callous-Unemotional Traits. Unpublished rating scale. New Orleans (LA): University of New Orleans, 2004.

Frick 2012

Frick PJ, Nigg JT. Current issues in the diagnosis of attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. *Annual Review of Clinical Psychology* 2012;**8**:77-107. [DOI: 10.1146/annurev-clinpsy-032511-143150] [PMCID: PMC4318653] [PMID: 22035245]

Frick 2014

Frick PJ, Ray JV, Thornton LC, Kahn RE. Annual research review: a developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. *Journal of Child Psychology and Psychiatry* 2014;**55**(6):532-48. [DOI: 10.1111/jcpp.12152] [PMID: 24117854]

Frick 2016

Frick PJ. Current research on conduct disorder in children and adolescents. *South African Journal of Psychology* 2016;**46**(2):160-74. [DOI: 10.1177/0081246316628455]

Furlong 2012

Furlong M, McGilloway S, Bywater T, Hutchings J, Smith SM, Donnelly M. Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database of Systematic Reviews* 2012, Issue 2. Art. No: CD008225. [DOI: 10.1002/14651858.CD008225.pub2] [PMID: 22336837]

Gardner 2016

Gardner F, Pickles A, Beecham J, Hutchings J, Scott S, Landau S, et al. How far could widespread dissemination of parenting programmes improve child antisocial behaviour and reduce social inequalities? Combining datasets from trials in different communities to establish for whom programmes are effective and cost effective. www.nets.nihr.ac.uk/projects/phr/12307004 (accessed 1 August 2016).

Gardner 2019

Gardner F, Leijten P, Melendez-Torres GJ, Landau S, Harris V, Mann J, et al. The earlier the better? Individual participant data and traditional meta-analysis of age effects of parenting interventions. *Child Development* 2019;**90**(1):7-19. [DOI: 10.1111/cdev.13138] [PMID: 30216433]

Gerard 1994

Gerard AB, editor. Parent–Child Relationship Inventory (PCRI): Manual. Los Angeles (CA): Western Psychological Services, 1994.

Gibaud-Wallston 1978

Gibaud-Wallston J, Wandersman LP. Development and utility of the Parenting Sense of Competency Scale. 86th Annual Convention of the American Psychological Association; 1978 Aug 28-Sept 1; Toronto (ON).

Goodman 1997

Goodman R. The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry* 1997;**38**(5):581-6. [DOI: 10.1111/j.1469-7610.1997.tb01545.x] [PMID: 9255702]

GRADEpro GDT [Computer program]

GRADEpro GDT. Version accessed 26 September 2022. Hamilton (ON): McMaster University (developed by Evidence Prime). Available at gradepro.org.

Görtz-Dorten 2010

Görtz-Dorten A, Döpfner M, editor(s). Fragebogen zum Aggressiven Verhalten von Kindern (FAVK) [German Questionnaire on Aggressive Behaviour of Children]. Göttingen (Germany): Hogrefe, 2010.

Hawes 2005

Hawes DJ, Dadds MR. The treatment of conduct problems in children with callous-unemotional traits. *Journal of Consulting and Clinical Psychology* 2005;**73**(4):737-41. [DOI: 10.1037/0022-006X.73.4.737] [PMID: 16173862]

Hawes 2014

Hawes DJ, Price MJ, Dadds MR. Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: a comprehensive review. *Clinical Child and Family Psychology Review* 2014;**17**(3):248-67. [DOI: 10.1007/s10567-014-0167-1] [PMID: 24748077]

Higgins 2011

Higgins JP, Altman DG, Sterne JA. Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook/archive/v5.1/.

Higgins 2021

Higgins JP, Li T, Deeks JJ. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch V, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Holland 1997

Holland D, Dadds MR, editor(s). The Diagnostic Interview Schedule for Children, Adolescents, and Parents. Brisbane (QLD): Griffith University, 1997.

Horowitz 1979

Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. *Psychosomatic Medicine* 1979;**41**(3):209-18. [DOI: 10.1097/00006842-197905000-00004] [PMID: 472086]

Hutchings 2012

Hutchings J, Bywater T, Williams ME, Lane E, Whitaker CJ. Improvements in maternal depression as a mediator of child

behaviour change. *Psychology* 2012;**3**(9a):795-801. [DOI: 10.4236/psych.2012.329120]

Högström 2015

Högström J, Enebrink P, Melin B, Ghaderi A. Eighteen-month follow-up of internet-based parent management training for children with conduct problems and the relation of homework compliance to outcome. *Child Psychiatry and Human Development* 2015;**46**(4):577-88. [DOI: 10.1007/s10578-014-0498-7] [PMID: 25236326]

ICD 2023

World Health Organization. ICD-11. International Classification of Diseases 11th Revision. icd.who.int/en (accessed prior to 3 February 2023).

Jambroes 2016

Jambroes T, Jansen LM, Vermeiren RR, Doreleijers TA, Colins OF, Popma A. The clinical usefulness of the new LPE specifier for subtyping adolescents with conduct disorder in the DSM 5. *European Child & Adolescent Psychiatry* 2016;**25**(8):891-902. [DOI: 10.1007/s00787-015-0812-3] [PMID: 26725044]

Kaminski 2008

Kaminski JW, Valle LA, Filene JH, Boyle CL. A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology* 2008;**36**(4):567-89. [DOI: 10.1007/s10802-007-9201-9] [PMID: 18205039]

Kim-Cohen 2003

Kim-Cohen J, Caspi A, Moffit TE, Harrington H, Milne BJ, Poulton R. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry* 2003;**60**(7):709-17. [DOI: 10.1001/archpsyc.60.7.709] [PMID: 12860775]

Klahr 2014

Klahr AM, Burt SA. Evaluation of the known behavioural heterogeneity in conduct disorder to improve its assessment and treatment. *Journal of Child Psychology and Psychiatry* 2014;**55**(12):1300-10. [DOI: 10.1111/jcpp.12268] [PMID: 24889762]

Koerting 2013

Koerting J, Smith E, Knowles MM, Latter S, Elsey H, McCann DC, et al. Barriers to, and facilitators of, parenting programmes for childhood behaviour problems: a qualitative synthesis of studies of parents' and professionals' perceptions. *European Child & Adolescent Psychiatry* 2013;**22**(11):653-70. [DOI: 10.1007/s00787-013-0401-2] [PMCID: PMC3826057] [PMID: 23564207]

Leijten 2018

Leijten P, Gardner F, Landau S, Harris V, Mann J, Hutchings J, et al. 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* 2018;**59**(2):99-109. [DOI: 10.1111/jcpp.12781] [PMID: 28696032]

Leijten 2019

Leijten P, Gardner F, Melendez-Torres GJ, Weeland J, Hutchings J, Landau S, et al. Co-occurring change in children's conduct problems and maternal depression: latent class individual participant data meta-analysis of the Incredible Years parenting program. *Development and Psychopathology* 2019;**31**(5):1851-62. [DOI: 10.1017/S0954579419001068] [PMID: 31370916]

Leijten 2022

Leijten P, Melendez-Torres GJ, Gardner F. Research review: the most effective parenting program content for disruptive child behavior – a network meta-analysis. Journal of Child Psychology and Psychiatry 2022;**63**(2):132-42. [DOI: 10.1111/jcpp.13483] [PMID: 34240409]

Llewellyn 2010

Llewellyn G, Bundy A, Mayes R, McConnell D, Emerson E, Brentnall J. Development and psychometric properties of the Family Life Interview. *Journal of Applied Research in Intellectual Disabilities* 2010;**23**(1):52-62. [DOI: 10.1111/j.1468-3148.2009.00545.x]

Lochman 2004

Lochman JE, Wells KC. The coping power program for preadolescent aggressive boys and their parents: outcome effects at the 1-year follow-up. *Journal of Consulting and Clinical Psychology* 2004;**72**(4):571-8. [DOI: 10.1037/0022-006X.72.4.571] [PMID: 5301641]

Lochman 2014

Lochman JE, Baden RE, Boxmeyer CL, Powell NP, Qu L, Salekin KL, et al. Does a booster intervention augment the preventive effects of an abbreviated version of the coping power program for aggressive children? *Journal of Abnormal Child Psychology* 2014;**42**(3):367-81. [DOI: 10.1007/s10802-013-9727-y] [PMCID: PMC3690167] [PMID: 23417235]

Loeber 1985

Loeber R, Schmaling KB. Empirical evidence for overt and covert patterns of antisocial conduct problems: a metaanalysis. *Journal of Abnormal Child Psychology* 1985;**13**(2):337-53. [DOI: 10.1007/BF00910652] [PMID: 4008760]

Lundahl 2006

Lundahl B, Risser HJ, Lovejoy MC. A meta-analysis of parent training: moderators and follow-up effects. *Clinical Psychology Review* 2006;**26**(1):86-104. [DOI: 10.1016/j.cpr.2005.07.004] [PMID: 16280191]

McCabe 2005

McCabe KM, Yeh M, Garland AF, Lau AS, Chavez G. The GANA Program: a tailoring approach to adapting parent child interaction therapy for Mexican Americans. *Education and Treatment of Children* 2005;**28**(2):111-29. [URL: psycnet.apa.org/record/2006-06834-002]

McCabe 2012

McCabe K, Yeh M, Lau A, Argote CB. Parent–child interaction therapy for Mexican Americans: results of a pilot randomized clinical trial at follow-up. *Behavior Therapy* 2012;**43**(3):606-18.

[DOI: 10.1016/j.beth.2011.11.001] [PMCID: PMC7194395] [PMID: 22697448]

McDonald 2006

McDonald R, Jouriles EN, Skopp NA. Reducing conduct problems among children brought to women's shelters: intervention effects 24 months following termination of services. *Journal of Family Psychology* 2006;**20**(1):127-36. [DOI: 10.1037/0893-3200.20.1.127] [PMID: 16569097]

McKay 2020

McKay K, Kennedy E, Senior R, Scott S, Hill J, Doolan M, et al. Informing the personalisation of interventions for parents of children with conduct problems: a qualitative study. *BMC Psychiatry* 2020;**20**(1):513. [DOI: 10.1186/s12888-020-02917-1] [PMCID: PMC7576822] [PMID: 33081742]

McMahon 2003

McMahon RJ, Forehand RL, editors. Helping the NonCompliant Child: Family-Based Treatment for Oppositional Behavior. 2nd edition. New York (NY): Guilford Press, 2003.

Moffit 2009

Moffit TE, Scott S. Conduct disorders of childhood and adolescence. In: Rutter M, Bishop DV, Pine DS, Scott S, Stevenson J, Taylor E, et al, editors(s). Rutter's Child and Adolescent Psychiatry. 5th edition. Oxford (UK): Blackwell Publishing Ltd, 2009:543-64. [DOI: 10.1002/9781444300895.ch35]

Moffitt 2008

Moffitt TE, Arseneault L, Jaffee SR, Kim-Cohen J, Koenen KC, Odgers CL, et al. Research review: DSM-V conduct disorder: research needs for an evidence base. *Journal of Child Psychology and Psychiatry* 2008;**49**(1):3-33. [DOI: 10.1111/j.1469-7610.2007.01823.x] [PMCID: PMC2822647] [PMID: 18181878]

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine* 2009;**6**(7):e1000097. [DOI: 10.1371/journal.pmed1000097] [PMCID: PMC2707599] [PMID: 19621072]

MRC 2017

Medical Research Council. Strategy for lifelong mental health research; April 2017. www.mrc.ac.uk/documents/pdf/strategy-for-lifelong-mental-health-research/ (accessed prior to 6 February 2023).

Mullen 1995

Mullen EM, editor. Mullen Scales of Early Learning. Circle Pines (MN): American Guidance Service, 1995.

Muratori 2016

Muratori P, Lochman JE, Lai E, Milone A, Nocentini A, Pisano S, et al. Which dimension of parenting predicts the change of callous unemotional traits in children with disruptive behavior disorder? *Comprehensive Psychiatry* 2016;**69**:202-10. [DOI: 10.1016/j.comppsych.2016.06.002] [PMID: 27423362]

Murphy 2013

Murphy M, Fonagy P. Chapter 10. Mental health problems in children and young children; 24 October 2013. assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/252660/33571_2901304_CMO_Chapter_10.pdf (accessed 6 February 2023).

Ng 2016

Ng MY, Weisz JR. Annual research review: building a science of personalized intervention for youth mental health. *Journal of Child Psychology and Psychiatry* 2016;**57**(3):216-36. [DOI: 10.1111/jcpp.12470] [PMCID: PMC4760855] [PMID: 26467325]

NICE 2017

National Institute for Health and Care Excellence. Antisocial behaviour and conduct disorders in children and young people: recognition and management (CG158); 27 March 2013, updated 19 April 2017. www.nice.org.uk/guidance/cg158 (accessed prior to 6 February 2023).

NIMH 1985

National Institute of Mental Health. CGI (Clinical Global Impression Scale). *Psychopharmacological Bulletin* 1985;**21**:839-44.

Odgers 2007

Odgers CL, Caspi A, Broadbent JM, Dickson N, Hancox RJ, Harrington H, et al. Prediction of differential adult health burden by conduct problem subtypes in males. *Archives of General Psychiatry* 2007;**64**(4):476-84. [DOI: 10.1001/archpsyc.64.4.476] [PMID: 17404124]

Pardini 2012

Pardini DA, Byrd AL. Perceptions of aggressive conflicts and others' distress in children with callous-unemotional traits: 'I'll show you who's boss, even if you suffer and I get in trouble'. *Journal of Child Psychology and Psychiatry* 2012;**53**(3):283-91. [DOI: 10.1111/j.1469-7610.2011.02487] [PMCID: PMC3822527] [PMID: 22066467]

Pasalich 2016

Pasalich DS, Witkiewitz K, McMahon RJ, Pinderhughes EE, Conduct Problems Prevention Research Group. Indirect effects of the fast track intervention on conduct disorder symptoms and callous-unemotional traits: distinct pathways involving discipline and warmth. *Journal of Abnormal Child Psychology* 2016;**44**(3):587-97. [DOI: 10.1007/s10802-015-0059-y] [PMCID: PMC4961049] [PMID: 26242993]

Peters 2005

Peters S, Calam R, Harrington R. Maternal attributions and expressed emotion as predictors of attendance at parent management training. *Journal of Child Psychology and Psychiatry* 2005;**46**(4):436-48. [DOI: 10.1111/j.1469-7610.2004.00365.x] [PMID: 15819652]

Pilling 2013

Pilling S, Gould N, Whittington C, Taylor C, Scott S, Guideline Development Group. Recognition, intervention, and management of antisocial behaviour and conduct disorders in

children and young people: summary of NICE-SCIE guidance. *BMJ* 2013;**346**:f1298. [DOI: 10.1136/bmj.f1298] [PMID: 23535256]

Power 1993

Power TG, editor. Parenting Dimensions Inventory (PDI): a Research Manual. Houston (TX): University of Houston, 1993.

Procidiano 1983

Procidiano ME, Heller K. Measures of perceived social support from friends and from family: three validation studies. *American Journal of Community Psychology* 1983;**11**(1):1-24. [DOI: 10.1007/BF00898416] [PMID: 6837532]

Quay 1983

Quay HC, Peterson DR, editor(s). Manual for the Revised Behavior Problem Checklist. Coral Gables (FL): University of Miami, 1983.

Reed 2019

Reed GM, First MB, Kogan CS, Hyman SE, Gureje O, Gaebel W, et al. Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. *World Psychiatry* 2019;**18**(1):3-19. [DOI: 10.1002/wps.20611] [PMCID: PMC6313247] [PMID: 30600616]

Review Manager 2020 [Computer program]

Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

RevMan Web 2022 [Computer program]

Review Manger Web (RevMan Web). Version 4.24.0. The Cochrane Collaboration , 2020. Available at revman.cochrane.org.

Reyno 2006

Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalizing behavior problems – a meta-analytic review. *Journal of Child Psychology and Psychiatry* 2006;**47**(1):99-111. [DOI: 10.1111/j.1469-7610.2005.01544.x] [PMID: 16405646]

Rivenbark 2018

Rivenbark JG, Odgers CL, Caspi A, Harrington H, Hogan S, Houts RM, et al. The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *Journal of Child Psychology and Psychiatry* 2018;**59**(6):703-10. [DOI: 10.1111/jcpp.12850] [PMCID: PMC5975095] [PMID: 29197100]

Sanders 1982

Sanders MR, Dadds MR. The effects of planned activities and child management procedures in parent training: an analysis of setting generality. *Behavior Therapy* 1982;**13**(4):452-61. [DOI: 10.1016/S0005-7894(82)80007-5]

Sanders 1993

Sanders MR, Dadds MR, editor(s). Behavioural Family Intervention (Psychology Practitioner Guidebooks). Needham Heights (MA): Allyn & Bacon, 1993.

Schünemann 2021

Schünemann HJ, Higgins JP, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook/archive/v6.2.

Scott 2009

Scott S, Dadds MR. Practitioner review: when parent training doesn't work: theory-driven clinical strategies. *Journal of Child Psychology and Psychiatry* 2009;**50**(12):1441-50. [DOI: 10.1111/j.1469-7610.2009.02161.x] [PMID: 19754503]

Scott 2012

Scott S, O'Conner TG. An experimental test of differential susceptibility to parenting among emotionally-dysregulated children in a randomized controlled trial for oppositional behavior. *Journal of Child Psychology and Psychiatry* 2012;**53**(11):1184-93. [DOI: 10.1111/j.1469-7610.2012.02586.x] [PMID: 22881827]

Scott 2015

Scott S. Oppositional and conduct disorders. In: Thapar A, Pine DS, Leckman JF, Scott S, Snowling MJ, Taylor EA, editors(s). Rutter's Child and Adolescent Psychiatry. 6th edition. Chichester (UK): Wiley-Blackwell, 2015:913-30.

Scott 2016

Scott S. Commentary: finding out the best way to tailor psychological interventions for children and families – a commentary on Ng and Weisz (2016). *Journal of Child Psychology and Psychiatry* 2016;**57**(3):237-40. [DOI: 10.1111/jcpp.12543] [PMID: 26889891]

Shelleby 2014

Shelleby EC, Shaw DS. Outcomes of parenting interventions for child conduct problems: a review of differential effectiveness. *Child Psychiatry and Human Development* 2014;**45**(5):628-45. [DOI: 10.1007/s10578-013-0431-5] [PMCID: PMC4082479] [PMID: 24390592]

Silberg 2015

Silberg J, Moore AA, Rutter M. Age of onset and the subclassification of conduct/dissocial disorder. *Journal of Child Psychology and Psychiatry* 2015;**56**(7):826-33. [DOI: 10.1111/jcpp.12353] [PMCID: PMC4417077] [PMID: 25359313]

Smeets 2015

Smeets KC, Leeijen AA, van der Molen MJ, Scheepers FE, Buitelaar JK, Rommelse NN. Treatment moderators of cognitive behavior therapy to reduce aggressive behavior: a meta-analysis. *European Child & Adolescent Psychiatry* 2015;**24**(3):255-64. [DOI: 10.1007/s00787-014-0592-1] [PMID: 25138144]

Sonuga-Barke 2002

Sonuga-Barke EJ, Daley D, Thompson M. Does maternal ADHD reduce the effectiveness of parent training for preschool

children's ADHD? *Journal of the American Academy of Child and Adolescent Psychiatry* 2002;**41**(6):696-702. [DOI: 10.1097/00004583-200206000-00009] [PMID: 12049444]

Sourander 2016

Sourander A, McGrath PJ, Ristkari T, Cunningham C, Huttunen J, Lingley-Pottie P, et al. Internet-assisted parent training intervention for disruptive behavior in 4-year-old children: a randomized clinical trial. *JAMA Psychiatry* 2016;**73**(4):378-87. [CLINICALTRIALS.GOV: NCT01750996] [DOI: 10.1001/jamapsychiatry.2015.3411] [PMID: 26913614]

Sprafkin 1997

Sprafkin J, Gadow KD, editor(s). Early Childhood Inventory-4 Norms Manual. Stony Brook (NY): Checkmate Plus, 1997.

Strayhorn 1988

Strayhorn JM, Weidman CS. A Parent Practices Scale and its relation to parent and child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry* 1988;**27**(5):613-8. [DOI: 10.1097/00004583-198809000-00016] [PMID: 3182627]

Sukhodolsky 2004

Sukhodolsky DG, Kassinove H, Gorman BS. Cognitive-behavioural therapy for anger in children and adolescents: a meta-analysis. *Aggression and Violent Behavior* 2004;**9**(3):247-69. [DOI: 10.1016/j.avb.2003.08.005]

Tully 1999

Tully LA, Sanders MR, Pollard GE, Baade PD, Heywood AH, Lynch ME, et al. A survey of parenting practices in Queensland: implications for mental health promotion. *Health Promotion Journal of Australia* 1999;**9**(2):105-14. [URL: search.informit.com.au/documentSummary;dn=459104446634945;res=IELHEA]

Viding 2005

Viding E, Blair RJ, Moffitt TE, Plomin R. Evidence for substantial genetic risk for psychopathy in 7-year-olds. *Journal of Child Psychology and Psychiatry* 2005;**46**(6):592-7. [DOI: 10.1111/j.1469-7610.2004.00393.x] [PMID: 15877765]

Viding 2020

Viding E, McCrory E. Disruptive behavior disorders: the challenge of delineating mechanisms in the face of heterogeneity. *American Journal of Psychiatry* 2020;**177**(9):811-7. [DOI: 10.1176/appi.ajp.2020.20070998] [PMID: 2867523]

Waschbusch 2002

Waschbusch DA. A meta-analytic examination of comorbid hyperactive-impulsive-attention problems and conduct problems. *Psychological Bulletin* 2002;**128**(1):118-50. [DOI: 10.1037/0033-2909.128.1.118] [PMID: 11843545]

Wilkinson 2016

Wilkinson S, Waller R, Viding E. Practitioner review: involving young people with callous unemotional traits in treatment – does it work? A systematic review. *Journal of Child Psychology and Psychiatry* 2016;**57**(5):552-65. [DOI: 10.1111/jcpp.12494] [PMID: 26686467]

Woolfenden 2001

Woolfenden SR, Williams K, Peat J. Family and parenting interventions in children and adolescents with conduct disorder and delinquency aged 10–17. *Cochrane Database of Systematic Reviews* 2001, Issue 2. Art. No: CD003015. [DOI: 10.1002/14651858.CD003015] [PMCID: PMC7004256] [PMID: 11406062]

Zisser 2012

Zisser AR, Eyberg SM. Maternal ADHD: parent–child interactions and relations with child disruptive behavior. *Child & Family Behavior Therapy* 2012;**34**(1):33-52. [DOI: 10.1080/07317107.2012.654450]

References to other published versions of this review

Kennedy 2017

Kennedy EM, French L, Roberts C. Personalised interventions for subgroups of children with conduct problems. *Cochrane Database of Systematic Reviews* 2017, Issue 8. Art. No: CD012746. [DOI: 10.1002/14651858.CD012746]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Bor 2002

Study characteristics	
Methods	Allocation: RCT
	Blindness: assumed that parents and therapists not blinded
	Duration: 15- to 17-week treatment and 1-year follow-up
	Setting: Australia, within a community outreach campaign targeting disadvantaged families

Bor 2002 *(Continued)*

Participants	Subgroup: children with co-occurring disruptive behaviour and attentional/hyperactive difficulties Sample size: 87 Age: children: enhanced behavioural family intervention: M 40.41 (SD 3.80) months, standard behavioural family intervention: M 39.86 (SD 3.34) months, WL: M 42.81 (SD 3.81) months; mothers: enhanced behavioural family intervention: M 28.41 (SD 4.21) years, standard behavioural family intervention: M 30.21 (SD 4.69) years, WL: M 29.72 (SD 4.57) years; fathers: enhanced behavioural family intervention: M 31.54 (SD 6.23) years, standard behavioural family intervention: M 33.65 (SD 7.89) years, WL: M 33.03 (SD 5.51) years Sex: 68% boys, 32% girls Inclusion criteria: target child aged 36–48 months; mothers rated their child's behaviour in the elevated range on the ECBI Intensity score or Problem score; child showed no evidence of developmental disorder (e.g. language disorder, autism) or significant health impairment; child was not currently having regular contact with another professional or agency or taking medication for behavioural problems; and parents were not currently receiving therapy for psychological problems, were not intellectually disabled, and reported they were able to read the newspaper without assistance. In addition, all families had ≥ 1 of the following family adversity factors: maternal depression (BDI score ≥ 20); relationship conflict (PPC score ≥ 5); single parent household; or low gross family income (< 345 Australian dollars per week) or low occupational prestige (i.e. major income earner on the Power, Privilege and Prestige Scale ≥ 5). Mothers reported presence of ≥ 6 symptoms of inattention or hyperactivity–impulsivity in a clinical diagnostic interview based on DSM-IV criteria for ADHD. Mothers rated their child's behaviour as > 90th percentile on ECBI Inattentive Behavior subscale. Pearson correlation indicated that these 2 measures of children's attentional/hyperactivity problems were correlated (P < 0.0001) Exclusion criteria: not reported Comorbidities: not reported
Interventions	1. EBFI: comprised 3 elements: PST, and CST. The parent training component consisted of Standard Triple P as per the SBFI intervention. The PST and CST constituted 2 evidence-based adjunctive interventions designed to address the family risk factors of marital conflict and parental adjustment, respectively. Parents in the EBFI intervention received the intensive behavioural parent training component as described previously for the SBFI condition (i.e. 17 child management strategies and planned activities training) as well as partner support and coping skills. As in the SBFI condition, each family received *Every Parent* and a workbook, *Every Parent's Supplementary Workbook*. The adjunctive interventions were delivered through a combination of within-session exercises and homework assignments, and tailored to the needs of each family. Although all the content of each module was covered with each family, the amount of time spent on active skills training varied across families. The findings obtained from the initial assessment guided practitioners in determining which areas of each adjunctive module needed to be practiced within sessions. Completers of this intervention were those families that completed the content of each of the modules. a. Sample size: 26. b. Duration of treatment: although the programmes were intended to be completed via weekly sessions, because of various reasons such as illness and public/school holidays, it typically took families 17 weeks to complete the EBFI programme. c. Timing and delivery: parents received 14 hours of intervention via 12 appointments completed over 17 weeks. Families allocated to the intervention attended 60- to 90-minute sessions with a practitioner on an individual basis in local community health and neighbourhood centres. 2. SBFI: involved teaching parents 17 core child management strategies. T10 strategies were designed to promote children's competence and development, and 7 strategies were designed to help parents manage misbehaviour. In addition, parents were taught a 6-step planned activities routine to enhance the generalisation and maintenance of parenting skills. Consequently, parents were taught to apply parenting skills to a broad range of target behaviours in both home and community settings with the target child and all relevant siblings. a. Sample size: 29. b. Duration of treatment: it typically took families 17 weeks to complete the programme.

Bor 2002 *(Continued)*

	c. Timing and delivery: families allocated to the intervention attended 60- to 90-minute sessions with a practitioner on an individual basis in local community health and neighbourhood centres. 3. WL: families received no treatment and had no contact with the research team for 15 weeks. They completed the postassessment, participated in treatment, and took no further part in the study. a. Sample size: 32. b. Duration of treatment: not applicable. c. Timing and delivery: not applicable.
Outcomes	1. Parenting skills and knowledge: observations of mother and child behaviour, measured using the FOS. Trained observers rated negative parent behaviour 2. Improvement in child conduct problems: ECBI Intensity, Problem, Inattention, Oppositional Defiant, and Conduct Disorder subscales, completed by parents 3. Improvement in child conduct problems: PDR mean Daily and mean Target subscales, completed by parents 4. Parenting skills and knowledge: PS Total score, for parents 5. Parenting skills and knowledge: PSOC Total score, for parents 6. Family functioning: PPC Problem subscale, completed by parents 7. DASS Total score, for parents Completed at pre-intervention, postintervention, and 1-year follow-up
Funding	The Triple P project is an ongoing study conducted at the School of Psychology, The University of Queensland. The study is supported by grants from Queensland Health and the National Health and Medical Research Council (941044, 971099).
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random sequence generation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: methods of treatment allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: measures relied on parent report (including the primary outcomes of PDR and ECBI) (p 576).
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: only 63/87 participants who completed treatment were included in the analysis. Quote: "Differential attrition occurred across conditions based on mothers' reports of parenting conflict … Mothers who did not complete the intervention rated their child's behavior on [the ECBI problem score] as more problematic than those who did complete the program." (p 585)
Selective reporting (reporting bias)	Low risk	Comment: all outcomes were reported (p 579).

Bor 2002 *(Continued)*

Other risk of bias: lack of adherence to treatment manual	Low risk	Quote: "Analysis of [adherence] checklists in each condition indicated that 100% of the practitioners discussed all the content material specified for that condition with each family and gave out all the required homework assignments." (p 579)
Other risk of bias: group differences	Low risk	Quote: "There was no significant difference across conditions on any measure at preintervention." (p 579)

Dadds 1992

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 8-week treatment and 6-month follow-up Setting: Queensland, Australia
Participants	Subgroup: isolated single parents Sample size: 22 Age: parents: M 32.8 years (group 1: M 31.9 (SD 2.4) years; group 2: M 32.4 (SD 2.8) years; children: M 54.8 months (group 1: M 54.2 (SD 11.2) months; group 2: M 55.2 (SD 10.5)) Sex: 15 boys, 7 girls (group 1: 8 boys, 3 girls; group 2: 7 boys, 4 girls) Inclusion criteria: availability of a person to function as an ally for the course of the treatment programme and commit to 8-week treatment duration and be available to attend an initial session and terminating session before and after the 6-week training programme; target child met the DSM-III-R criteria for oppositional or CD as confirmed by standardised clinical intake interview; child's behaviour problem could not be associated with organic pathology, and no psychiatric pathology apart from the conduct problem was evident; no family member was undergoing other psychological treatment; willingness to complete self-report and home observation procedures Exclusion criteria: not reported Comorbidities: not reported
Interventions	1. Child management training + ally (Group 1; CMT + ally-specific task): group 1 consisted of parents and their allies, people nominated by the parents to act as allies or support people for the 8-week duration of contact. Allies were not required to attend the clinic for the child management programme but were permitted to participate if desired. An initial session was conducted with both parents and allies present. Ally-specific tasks to perform, which were outlined and modelled by the therapist, as well as written instructions provided in a booklet. The role of the ally was designed primarily as supporting the parent, rather than assisting in parenting per se. Allies were not expected to attend child management training sessions and acquire the skills being taught to the parent, but rather to function as a backup person during the consequences of skill implementation. Parents received child management training in mixed groups. As well as receiving technique training, mothers were encouraged to problem-solve for and with each other. In these sessions, mothers were encouraged to act as "therapists" and discuss alternative solutions to remaining problems. A terminating session with allies present was conducted for group 1, during which termination with allies was completed. Allies were asked to provide feedback to the therapist regarding what aspects of their participation they found most positive or least useful. The therapist attempted to reinforce allies for their involvement and point out their important role in the project. At this point allies' commitment to the programme was terminated, but they were encouraged to continue to act in this capacity if they desired. Additionally, parents were

Dadds 1992 *(Continued)*

encouraged to continue to recruit the type of assistance from others that they had been taught during the AST condition of treatment.
   a. Sample size: 11.
   b. Duration of treatment: 8 weeks.
   c. Timing and delivery: all interviews and training sessions with parents and allies were conducted in clinical facilities within the Department of Psychology at the University of Queensland. Child management training for both groups involved 6 training sessions provided by Therese A McHugh and another psychologist practicing in the area of child psychopathology.
2. Child management training (CMT; group 2): parents received contact with the therapist during group discussion sessions. In the first session parents were involved in a general discussion about issues and special problems for the single parent. Parents then received child management training in mixed groups.
   a. Sample size: 11.
   b. Duration of treatment: 8 weeks.
   c. Timing and delivery: all interviews and training sessions with parents and allies were conducted in clinical facilities within the Department of Psychology at the University of Queensland. Child management training for both groups involved 6 training sessions provided by Therese A McHugh and another psychologist practicing in the area of child psychopathology.

Outcomes	1. Improvement in child conduct problems: a. PDR completed by parents b. RBPC Conduct Disorder subscale, completed by parents 2. Personalised treatment outcomes: a. BDI completed by parents b. ISSB completed by parents c. PSS-Fa completed by parents d. PSS-Fr, completed by parents 3. Parenting skills and knowledge: FOS observers coded parental aversive commands All completed at pre-intervention, postintervention, and 6-month follow-up.
Funding	Not reported
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment condition.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: multiple outcome measures relied on parent report (including the primary outcome measures of PDR and RBPC) (p 254).
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: all participants appeared to complete treatment.

Dadds 1992 *(Continued)*

		Quote: "At follow-up, 11 families responded to treatment (50%), and 11 families did not (50%)" (p 256).
Selective reporting (reporting bias)	Low risk	Comment: all outcome measures were reported (p 255).
Other risk of bias: lack of adherence to treatment manual	Unclear risk	Comment: no information given about implementation of training.
Other risk of bias: group differences	Low risk	Quote: "Treatment groups were not significantly different on any of the demographic measures." (p 253)

Dadds 2019

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 13-week treatment and 3-month follow-up Setting: New South Wales, Australia
Participants	Subgroup: rural families Sample size: 133 Age: overall: M 6 years; AccessEI: M 6.79 (SD 1.68) years; FTF: M 6.75 (SD 1.92) years Sex: AccessEI: 77.6% boys; FTF: 81.8% boys Inclusion criteria: child met full diagnosis for ODD or CD according to DSM-IV or displayed subclinical symptoms warranting interventions for either disorder. Comorbidities of concurrent ADHD, anxiety, mood disorder, or autism spectrum disorder of mild severity (including Asperger's disorder) were allowed where they were secondary to the primary referral for ODD or CD. If child was on medication for emotional or behavioural concerns (e.g. methylphenidate, dexamphetamine), regimen was stabilised prior to treatment commencement and no further changes were planned. Child aged 3–9 years at pretreatment assessment; parental access to a computer and internet during intervention; family lived outside a major city in Australia; and willingness to travel to Sydney and attend assessment Exclusion criteria: child with developmental delay as indicated by IQ 65; IQ scores gathered using WISC–III, WPPSI-III, or WPPSI-IV depending on the child's age (IQ tests conducted by registered psychologist within last 2 years); child had a major neurological disorder, concurrent diagnosis of a psychotic disorder, or primary moderate-to-severe autism spectrum disorder; parent had significant untreated severe mental health problems or substance abuse disorder; parent had immediate risk of suicide, violence, abuse, or neglect requiring crisis response or child protection notification; family were engaged in ongoing concurrent psychological intervention for the child Comorbidities: not reported
Interventions	1. AccessEI: based on Integrated Family Intervention for Child Conduct Problems, a fully manualised parent training programme with a substantial evidence base for both FTF and more recently, using a therapist-assisted online version. It focuses on training parents to reduce child conduct problems and improve the parent–child relationship via well-established parenting strategies based on social learning, attachment, and family systems theories. AccessEI consisted of 6 interactive and educational video modules as well as sessions with a clinician delivered via secure video-conferencing. The prerecorded video modules included psychoeducation, strategies, and role plays. Participants were re-

Dadds 2019 *(Continued)*

quested to watch each video module before participating in an individualised video-conference session with their therapist within a secure platform.
   a. Sample size: 67.
   b. Duration of treatment: not reported.
   c. Timing and delivery: 6 interactive and educational video modules as well as approximately 6–10 sessions with a clinician delivered via secure video-conferencing. The prerecorded video modules were presented by a clinical psychologist. The 6 video modules were 7–19 minutes in duration (total duration 1 hour and 14 minutes). Therapists scheduled weekly 50- to 60-minute video-conference sessions. 9 clinical psychologists, as well as 2 provisionally registered psychologists undertaking postgraduate training in clinical psychology, conducted the treatment across both modalities.
2. FTF: consisted of intensive parent management training delivered by a clinical psychologist. Parenting strategies were taught through active skills training including modelling, rehearsal, and feedback. No videos were viewed during the FTF intervention. Treatment content and implementation was reviewed in a follow-up telephone call after treatment delivery.
   a. Sample size: 66.
   b. Duration of treatment: delivered over 1 week. Treatment was reviewed in a follow-up telephone call about 2–4 weeks after treatment delivery.
   c. Timing and delivery: for most families, 4 × 1.5-hour sessions were scheduled. 9 clinical psychologists, as well as 2 provisionally registered psychologists undertaking postgraduate training in clinical psychology, conducted the treatment across both modalities.

Outcomes	1. Improvement in child conduct problems: DISCAP Clinical symptom severity of CD and ODD, completed with parents 2. Improvement in child conduct problems: CPRS-R Oppositional Behavior subscale, completed by mother 3. Improvement in child conduct problems: SDQ Total Difficulties score, completed by mother Completed at pre-intervention, postintervention, and 3-month follow-up
Funding	Supported by grant from the National Health and Medical Research Council
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Participating families were randomly allocated to either AccessEI or FTF using a block randomized, sequentially numbered, opaque sealed envelopes system … Block sizes of four (two of each treatment condition) and six (three of each treatment condition) envelopes were used. Each block was thoroughly shuffled so that order within the block was random. A coin-toss method was used to allocate the order of the blocks in the sequence." (p 709)
Allocation concealment (selection bias)	Low risk	Quote: "Participating families were randomly allocated to either AccessEI or FTF using a block randomized, sequentially numbered, opaque sealed envelopes system… Each block was thoroughly shuffled so that order within the block was random. A coin-toss method was used to allocate the order of the blocks in the sequence." (p 709)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists were aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: all outcome measures relied on parent report (including primary outcome measures of DSM diagnoses, CPRS-R Oppositional Behavior subscale, and SDQ Total score).

Dadds 2019 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "All analyses were checked for ITT samples using multiple imputation methods." (p 713)
Selective reporting (reporting bias)	Low risk	Comment: all outcome measures were reported (p 712).
Other risk of bias: lack of adherence to treatment manual	Low risk	Comment: mean adherence to treatment manual above 80% in both conditions, no significant differences between groups that are likely to be substantial enough to affect outcomes and adequate sample size.
Other risk of bias: group differences	Low risk	Quote: "Therapist ratings of implementation fidelity for proportion of content covered in treatment ranged from 50% to 100% per family. There was no significant difference between treatment groups (AccessEI: M = 94% SD = 13%; FTF: M = 94% SD = 11%)." (p 712) Quote: "To examine equivalence of families randomized to AccessEI and FTF groups at pretreatment, t tests for continuous variables and chi-square tests for categorical variables were conducted across pretreatment measures and sociodemographic variables. The results showed no significant group differences on any measure indicating that randomization resulted in comparable groups at baseline." (p 712)

Greene 2004

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: up to 16-week treatment and 4-month follow-up Setting: outpatient mental health clinic specialising in the treatment of disruptive behaviour disorders at a university teaching hospital; USA
Participants	Subgroup: affectively dysregulated children with ODD Sample size: 47 Age: children: collaborative problem-solving: M 7.4 (SD 0.40) years, PT: M 6.8 (SD 0.45) years Sex: 32 boys, 15 girls (collaborative problem-solving: 18 boys, 10 girls; PT: 14 boys, 5 girls) Inclusion criteria: met full diagnostic criteria for ODD; none met full diagnostic criteria for CD at the time of enrolment in the study (many had subthreshold features of CD); had at least subthreshold features of juvenile bipolar disorder or major depression (defined as more than half of the symptoms needed to meet criteria for the diagnosis) Exclusion criteria: estimated full-scale IQ < 80 or were actively suicidal or homicidal on entry into study Comorbidities: subthreshold or full major depression – PT: 12 (63.2%), collaborative problem-solving: 17 (60.7%). Subthreshold or full bipolar disorder – PT: 15 (78.9%), collaborative problem-solving: 18 (64.3%); ADHD – PT: 13 (68.4%), collaborative problem-solving: 18 (64.3%); anxiety disorder (≥ 1) – PT: 8 (42.1%); collaborative problem-solving: 11 (39.3%)
Interventions	1. CPS: families received a model of psychosocial treatment. The CPS approach aims to help adults achieve the following treatment goals: (a) understand the cognitive factors that may contribute to aggressive outbursts, most notably in the domains of emotion regulation, frustration tolerance, prob-

Greene 2004 *(Continued)*

lem solving, and adaptability skills; (b) become cognisant of 3 basic strategies (known as the baskets framework) for handling unmet expectations, including imposition of adult will, CPS, and removing the expectation; (c) recognise the impact of each of these 3 strategies on adult–child interactions; and (d) become proficient, along with their children, at CPS as a means of resolving disagreements and defusing potentially conflictual situations so as to reduce the likelihood of aggressive outbursts. Although CPS is manualised, session content and duration were not circumscribed to facilitate greater matching of therapeutic ingredients to the needs of individual children and families.
   a. Sample size: 30.
   b. Duration of treatment: range of treatment sessions was 7–16 weeks, and the mean length of treatment was 11 weeks.
   c. Timing and delivery: treatment sessions were attended primarily by parents, with identified children included at the discretion of the therapist.
2. PT: families received a 10-week behaviour management programme, consisting of: (a) discussing and educating parents about the causes of children's defiant behaviour, (b) instructing parents on positive attending through use of special time, (c) training parents to use attending skills to increase compliant behaviour, (d) increasing the effectiveness of parental commands, (e) implementing a contingency management programme, (f) using the time-out procedure, (g) managing children's behaviour in public places, and (h) using a daily school–home report card. This treatment programme is manualised, with specified weekly session content.
   a. Sample size: 20.
   b. Duration of treatment: 10-week behaviour management programme.
   c. Timing and delivery: weekly treatment sessions in this condition were attended primarily by parents, with identified children included as indicated by the training manual.

Outcomes	1. Parenting skills and knowledge: PCRI completed by mothers and fathers at pretreatment and post-treatment 2. Parental stress: PSI completed by mothers and fathers at pretreatment and post-treatment 3. Improvement in child conduct problems: ODD Rating Scales, completed by mothers and fathers at pretreatment, post-treatment, and 4-month follow-up 4. Improvement in child conduct problems: CGI completed by the therapist at pretreatment, post-treatment, and 4-month follow-up
Funding	Theodore and Vada Stanley Foundation
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists were aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: some outcome measures were completed by parents (including a primary outcome measure of ODD Rating Scale).
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "28 children completed treatment in the CPS condition, and 19 children completed treatment in the PT condition." (p 1158)

Greene 2004 *(Continued)*

		Quote: "Of the children who completed treatment, 87% (PT, *n* = 16; CPS, *n* = 25) were available for follow-up assessment at 4-month posttreatment." (p 1158) Quote: "Data for 6 children (3 in each treatment condition) could not be collected at 4-month follow-up. We used last-observation-carried-forward methodology to account for these missing participants; means for each relevant variable were virtually unchanged, and comparisons of group differences were unaltered." (p 1161)
Selective reporting (reporting bias)	High risk	Comment: all outcome measures reported; however, only 3/7 subscales of the PCRI reported, and authors only reported mother-reported data due to frequently incomplete father-reported data.
Other risk of bias: lack of adherence to treatment manual	Low risk	Quote: "Data indicate that the PT condition was characterised largely by PT specific interventions, with very little inclusion of content relevant to CPS, and that CPS was characterized exclusively by CPS-specific interventions, with no inclusion of content relevant to PT" (p 1160). Quote: "CPS is a flexible intervention regarding order of content delivery." (p 1160)
Other risk of bias: group differences	Low risk	Quote: "There were no significant differences between the two treatment groups in any demographic variables, past or current GAF scores, or rates of diagnostic comorbidity. There were also no significant differences … in rates of children who were receiving psychotropic medication at pretreatment … or posttreatment." (p 1160)

Görtz-Dorten 2019

Study characteristics	
Methods	Allocation: RCT Blindness: assumed children, parents, teachers, and therapists not blinded Duration: 24-week treatment Setting: Cologne, Germany
Participants	Subgroup: boys with ODD/CD and peer-related aggression Sample size: 101 initially randomised, 91 included in analyses Age: 6–12 years; THAV: M 8.88 (SD 1.95) years; PLAY M 8.63 (SD 1.74) years Sex: boys Inclusion criteria: boys aged 6–12 years with IQ ≥ 80 (in Culture Fair Intelligence Test) and ICD-10 diagnosis of CD (F91), mixed disorder of conduct and emotions (F92), or hyperkinetic CD (F90.1) using the semi-structured interview for Disruptive Behavior Disorders (ODD, CD) (DCL-DBD) of the DISYPS-II. Peer-related aggressive behaviour had to cause persistent impairments in relationships with other children (clinical rating in the semi-structured interview) and child had to have a high symptom score (Stanine score ≥ 7) on SCL-DBD Total score of the DISYPS-II at pre-assessment Exclusion criteria: presence of a primary comorbid disorder (e.g. autism) according to judgement of clinician, planned change in medication in a child receiving psychotropic medication, other child psychotherapy, and parents who do not speak German

Görtz-Dorten 2019 *(Continued)*

	Comorbidities: not reported
Interventions	1. THAV: a social competence training programme developed in Germany. THAV is a CBT intervention for children aged 6–12 years with peer-related overt aggressive behaviour. It provides individualised treatment for problem-maintaining factors in specific daily life situations, which each respective child has experienced in previous weeks. Depending on the problem-maintaining factors specific to each individual, THAV aims to modify social-cognitive information processing, impulse control, social problem-solving, social skills, and social interactions in these situations. It combines patient-, parent-, teacher-, and peer-focused interventions. Patient-focused interventions are the main component, and parent-, teacher-, or peer-focused interventions are added according to the individual needs of the patient. a. Sample size: 50. b. Duration of treatment: 24 weeks. c. Timing and delivery: THAV comprises 24 weekly child sessions (lasting 45 minutes each) and additional sessions or shorter contacts with parents. Treatment and control interventions were carried out by 13 experienced child therapists or therapists in training. The same therapists administered the treatment and control interventions. The therapists received weekly group supervision by a senior child therapist. 2. PLAY: comprised educational group play, with 3–5 children in each group. Techniques to activate resources and the opportunity to practice prosocial interactions in groups were utilised. During the group sessions, social play interactions and projects were offered that aimed to develop co-operative interaction or to provide the opportunity to practice socially competent ways of solving conflicts. Children were supported to solve conflicts and to develop co-operative interactions, and were praised for socially competent behaviour and for their own general competencies. Parents attended 2 parent group sessions during which they received psychoeducation on appropriate general parenting strategies. However, these general parenting strategies were not tailored to the specific problems of the child and the parents were not trained to implement these techniques in their daily parenting behaviour. a. Sample size: 51; 41 included in analyses. b. Duration of treatment: 24 weeks. c. Timing and delivery: each group received 12 fortnightly sessions (lasting 90 minutes each). Parents attended 2 parent group sessions (90 minutes each). Treatment and control interventions were carried out by 13 experienced child therapists or therapists in training. The same therapists administered the treatment and control interventions. The therapists received weekly group supervision by a senior child therapist.
Outcomes	1. Improvement in child conduct problems: SCL-DBD ODD, CD, and Prosocial Behaviour subscales, completed by parents and teachers 2. Improvement in child conduct problems: CBCL Externalizing subscale, completed by parents 3. Improvement in child conduct problems: TRF Externalized subscale, completed by teachers 4. Improvement in child conduct problems: SPST Aggressive Behaviour and Socially Competent Behaviour subscales, rated by clinicians 5. Improvement in child conduct problems: FAVK Peer-related Aggression subscale, completed by parent and teacher Completed at pre- and postintervention
Funding	School of Child and Adolescent Behavior Therapy at the University Hospital Cologne
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "patients were randomized (50 THAV; 51 PLAY; block randomization with a block size of 4 and random selections from all 6 permutations)." (p 167)

Görtz-Dorten 2019 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Comment: awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents, teachers, and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: all outcome measures relied on parent, teacher, or clinician report (including primary measures of the parent- and teacher-rated DISYPS-II Symptom Checklist for Disruptive Behavior Disorders (assessing ODD and CD), the CBCL Externalizing, the TFR Externalized, and the Social Problem-Solving Test (Aggressive Behaviour).
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Missing values were imputed using the expectation maximization procedure, assuming a missing at random pattern." (p 169)
Selective reporting (reporting bias)	High risk	Comment: all primary and secondary outcome measures detailed in the methods were reported in the results. However, in the results section, there were subscales added that were not described in the methods (e.g. DISYPS-II Symptom Checklist for Disruptive Behavior Disorders – ADHD). Furthermore, there were changes from protocol stage, with new measures added in the study (DISYPS-II Symptom Checklist for Disruptive Behavior Disorders, the Inventory for Callous-Unemotional Traits), and others not reported (Symptom Checklist for Oppositional Defiant and Conduct Disorder, Individual Problem Checklist).
Other risk of bias: lack of adherence to treatment manual	Low risk	Quote: "therapists indicated that they spent 88% of the total treatment time on specific modules of the THAV treatment program … Therapists were supervised regularly and treatment integrity was rated globally as good to excellent by the supervisor." (p 169) Comment: mean adherence to treatment manual above 80% and rated as good to excellent and adequate sample size.
Other risk of bias: group differences	Unclear risk	Comment: no information reported regarding group differences.

Jones 2014

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 12-week treatment Setting: North Central, North Carolina
Participants	Subgroup: low-income families Sample size: 22 Age: children: overall: M 5.67 (SD 1.72) years; TE-HNC: M 5.57 (SD 1.27); HNC: M 5.75 (SD 2.12); carers: overall: M 36.73 (SD 8.81) years; TE-HNC: M 35 (SD 5.92) years; HNC: M 38.25 (SD 10.95) years

Jones 2014 *(Continued)*

	Sex: children: overall: 53% boys; TE-HNC: 57% boys; HNC: 50% boys; carers: overall: 87% women; TE-HNC: 71% women; HNC: 100% women Inclusion criteria: low-income families (i.e. adjusted gross income < 150% of federal poverty limit; child aged 3–8 years; and child exhibited disruptive behaviours in clinical range as evidenced by meeting or exceeding clinical cutoffs on the carer-report of ECBI Severity or Intensity subscales Exclusion criteria: child developmental or physical disability that precluded use of HNC skills; carer current diagnosis of substance abuse/dependence, mood, or psychotic disorder; family involvement with Department of Social Services related to abuse/neglect Comorbidities: not reported
Interventions	1. TE-HNC: a smartphone-enhanced version of 1 evidence-based BPT programme, Helping the Noncompliant Child. Smartphone-enhancements: (a) a 3-minute skills video for each of the HNC skills, including psychoeducation, as well as modelling of the skill by parent–child dyads; (b) daily surveys of skill practice and progress that are used to guide mid-week calls and weekly sessions (e.g. problem-solving more suitable home practice times if a parent indicates a failure to practice on a daily survey); (c) mid-week video-calls during which therapists reinforce caregivers for progress and problem-solve obstacles to practice; (d) weekly videotaped home practice, which provided a "window" for therapists to use during the session to provide feedback regarding skill development; and (5) text reminders regarding the relevance of home practice, the mid-week call, and session attendance, as well as reinforcing messages regarding progress. As such, TE-HNC capitalises on the capacity for smartphones to push HNC content to the caregiver, rather than relying on the caregiver to access the content, a proven strategy with other low-income clients. a. Sample size: 11. b. Duration of treatment: 12 weeks. c. Timing and delivery: 12 weekly sessions, delivered by master's level graduate students. 2. HNC: all families received the standard, 2-phase HNC programme. Following an orientation session, caregiver–child dyads start Phase I, Differential Attention, in which caregivers learn to increase the frequency and range of social attention to the child and to reduce the frequency of competing verbal behaviour. A primary goal is to reduce the coercive cycle of parent–child interaction by (re)establishing a positive and mutually reinforcing parent–child relationship. In the context of "Child's Game". The caregiver is taught to: (a) increase the frequency and range of positive attention; (b) eliminate instructions, questions, and criticisms; and (c) ignore minor inappropriate behaviour. In Phase II, Compliance Training, caregivers are taught the difference between unclear and clear instructions, to give the "Clear Instruction" sequence, and to use a non-physical discipline procedure, "Time-Out", for occasions of non-compliance and other inappropriate behaviour that cannot be ignored. a. Sample size: 11. b. Duration of treatment: 12 weeks. c. Timing and delivery: 12 weekly sessions, delivered by master's level graduate students.
Outcomes	1. Improvement in child conduct problems: ECBI – Intensity and Problem subscales, completed by carers at pre-intervention and postintervention 2. Engagement and decreased dropout: engagement measured by session attendance and mid-week check-in call availability throughout intervention
Funding	Provided by NIMH 1R34MH082956 (ClinicalTrials.gov Identifier: NCT01367847)
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.

Jones 2014 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: some outcome measures relied on parent report (including primary outcome measure of ECBI).
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "The four families (HNC = 2; TE-HNC = 2) who dropped out of the study notified project staff prior to discontinuing participation and each cited a major health (e.g., organ transplant) or family (e.g., divorce) stressor necessitating dropout. Given the pilot nature of the project, we considered only complete data (i.e., data from participants available at pre- and post- treatment)." (p 9) Comment: 21% attrition, no ITT analysis conducted.
Selective reporting (reporting bias)	Low risk	Comment: all outcomes were reported.
Other risk of bias: lack of adherence to treatment manual	Unclear risk	Comment: no information regarding adherence to treatment manual.
Other risk of bias: group differences	High risk	Quote: "Randomization failed to yield equivalent groups on an established correlate of BPT drop out, baseline disruptive behaviors. Caregivers randomized to TE-HNC were more likely to report higher levels of child disruptive behaviors on the Intensity Subscale in the TE-HNC group at baseline (M[Mean]ECBI Intensity: TE-HNC = 148.86; HNC = 131.50)" (p 9).

Jouriles 2001

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 16-month treatment Setting: 3 area shelters that offered refuge to battered women and their dependent children; Houston-Galveston, Texas, US
Participants	Subgroup: children exposed to intimate partner violence Sample size: 36 Age: children: M 5.67 (SD 1.88) years Sex: children: 26 boys, 10 girls Inclusion criteria: mother had to report the occurrence of ≥ 1 physically violent act from a male partner during previous 12 months; have ≥ 1 child aged 4–9 years who met DSM-IV criteria for ODD or CD (families identified as eligible for participation at screening interview were followed after shelter departure

Jouriles 2001 *(Continued)*

	to ensure eligibility for study). For families to remain eligible for participation after shelter departure mother and target child had to be living together in same household; had to be in a residence in which former partner was not a member; residence had to be within 50 miles of shelter from which they departed; residence had to be sufficiently safe for project staff to visit regularly Exclusion criteria: mother or the target child were exhibiting symptoms of serious mental illness (e.g. psychosis, autism) Comorbidities: not reported
Interventions	1. Intervention: multicomponent family intervention comprising 2 primary components: (a) providing mothers and children with social and instrumental support and mothers with problem-solving skills and (b) teaching mothers to use certain child management skills designed to help reduce their children's conduct problems. Families were assigned to an intervention team consisting of a trained therapist and ≥ 1 advanced undergraduate or postbaccalaureate students. Therapists worked primarily with the mothers (e.g. providing support and facilitating the development of problem-solving skills, teaching child management skills), while the students served as mentors for the children (e.g. providing positive support and serving as prosocial models). The principal component of this intervention was teaching mothers to effectively use a particular set of child management skills. Although manualised, the intervention was sufficiently flexible to allow adaptation to the needs of each family. That is, therapists systematically assessed each mother's beliefs, practices, and knowledge about parenting, each child's behaviour patterns, and the relationships amongst family members. The intervention was then tailored to meet the family's specific needs, with a focus on using the child management skills to address these needs. a. Sample size: 18. b. Duration of treatment: sessions were conducted in the family's home, beginning after shelter departure and continuing for up to 8 months. c. Timing and delivery: the intervention was designed to include weekly sessions of 1–1.5 hours. Intervention was structured so that it could be delivered in a flexible manner, but it was stopped for all families after 8 months. The mean number of sessions for families who completed the intervention was 23. 6 clinical psychology graduate students and 1 clinical psychologist served as therapists for the families. 2. Existing services: families were contacted monthly either in person or by telephone. They encouraged families to use existing community or shelter services. That is, no restrictions were placed on families' receipt of services from other sources, and indeed, the researchers encouraged them to make use of the resources available to them. Except for immediate safety concerns, the families received no clinical services through the programme or from therapists who were delivering clinical services through the programme. a. Sample size: 18. b. Duration of treatment: not applicable. c. Timing and delivery: not applicable.
Outcomes	1. Improvement in child conduct problems: CBCL Externalizing subscale, completed by mothers 2. Family functioning: direct observation of mothers' child management skills, rated by coders blinded to treatment allocation 3. Personalised treatment outcomes: SCL-90-R, completed by mothers Completed at 5 time points (every 4 months over 16 months)
Funding	Grants from Hogg Foundation for Mental Health, Texas Higher Education Coordinating Broads, and Grant 53380 from the National Institute of Mental Health
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement

Jouriles 2001 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation (particularly as control group received only community services).
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: some outcome measures relied on parent report (including primary outcome measure of CBCL).
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Assessments were repeated every 4 months … resulting in a total of five assessments … Thirty-one of the families participated in all five assessments. 33 participants participated in three of the five assessments. 3 families participated in 2 or fewer of the assessments." (p 778) Comment: 14% did not complete all assessments.
Selective reporting (reporting bias)	Low risk	Comment: all outcomes were reported (p 779).
Other risk of bias: lack of adherence to treatment manual	Low risk	Quote: "100% of the child management skills were presented in all the completed therapy cases." (p 779)
Other risk of bias: group differences	Low risk	Quote: "The groups did not differ at this initial assessment on any of the demographic variables, screening measures, or measured outcome variables." (p 779)

Jouriles 2009

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 20-month treatment Setting: urban and suburban DV shelters in US
Participants	Subgroup: children exposed to intimate partner violence Sample size: 66 Age: mothers: project support: M 29.8 (SD 6.2) years, existing services: M 29.1 years (SD 4.2) years; children: aged 4–9 years Sex: children: project support: 58.8% boys; existing services: 41.2% boys Inclusion criteria: eligibility determined at 3 time points: during an in-shelter screen, after shelter departure but prior to the first assessment conducted at the family's postshelter residence, and during the first in-home assessment. As part of the in-shelter screen, mothers completed the CTS2. Mother had to have reported experiencing ≥ 1 act of physical IPV from male partner during previous 12 months;

Jouriles 2009 *(Continued)*

	mothers participated in a structured clinical interview to assess whether their children met criteria for ODD or CD from DSM–IV; mother had ≥ 1 child aged 4–9 years who met criteria for ODD or CD; family must not have been currently receiving services targeting the child conduct problems. For families with > 1 eligible child, the oldest was the target child for data collection purposes Exclusion criteria: mothers participated in a brief structured interview about past psychiatric diagnoses (their own and their children's), hospitalisations, suicide attempts, mental health services, prescription medications, and substance use. If significant psychiatric symptoms or substance use was evident or reported, interviewer consulted the supervising psychologist to determine whether the problems were likely to interfere with the family's ability to participate in the project. If families were judged ineligible because of this, project staff assisted them in obtaining appropriate help during their shelter residence. Families considered ineligible at this point if they could not be located, they had moved > 50 miles (80.5 km) from project location, mother's abusive partner lived with family following shelter departure; family declined participation. During first assessment, authors repeated aspects of in-shelter screen for child conduct problems, psychiatric illness, and substance abuse. At conclusion of assessment, families were deemed ineligible if the target child no longer met DSM–IV criteria for ODD or CD, assessment revealed problems (psychiatric illness, substance abuse) not detected during in-shelter screen that would interfere with family's ability to participate in project, family no longer wanted to participate. Families excluded for other reasons as well (e.g. responding randomly to assessment questions) Comorbidities: not reported
Interventions	1. Project support: families received a family intervention that included 2 primary components: (a) teaching mothers child management skills and (b) providing instrumental and emotional support to mothers. The child management skills were detailed in a manual that specified the particular skills to be taught, scenarios for practice role plays using the skills, and homework assignments. It included 12 child management skills (e.g. listening to your child, praising, reprimanding). The skills were presented in sequence; the initial skills focused on improving the quality of the mother–child relationship and increasing prosocial child behaviour, and the latter skills focused on reducing problematic behaviour. Therapists worked primarily with the mothers, although children were brought into sessions for evaluating mothers' use of skills and children's responses to the skills. The skills were taught to mothers through didactic instruction accompanied by written materials, role plays, in vivo practice, corrective feedback, between session homework assignments, and mastery checks. a. Sample size: 32. b. Duration of treatment: for up to 8 months following shelter departure. Families received a mean of 20 (SD 9, range 2–40) home-based treatment sessions during the 8-month period. c. Timing and delivery: weekly home visits by project staff and training in child management skills. 8 Master's level clinicians and 1 clinical psychologist served as therapists. Therapists received extensive training in the content and techniques of the intervention. Specifically, therapists in training read and discussed the treatment manual and background materials on behavioural parent training with the clinical supervisor (a clinical psychologist), attended weekly group supervision meetings to learn from discussions and supervision of ongoing cases, and were required to complete a mastery test (3 role plays) to evaluate their proficiency in delivering the intervention. 2. Existing services: project staff attempted to contact families monthly. Contacts were structured so that these families could receive instrumental and emotional support services similar to those provided to Project Support families. In addition, no restrictions were placed on comparison families' receipt of services from other sources; families were encouraged them to make use of community resources. a. Sample size: 34. b. Duration of treatment: 8-month period following shelter departure. Families had a mean 3.7 (SD 2.7, range 0–9) contacts with project staff. c. Timing and delivery: monthly contacts, in person or telephone.
Outcomes	1. Improvement in child conduct problems: CBCL Externalizing subscale, completed by mothers at baseline, 8 months postintervention, and 20 months' (final) follow-up 2. Improvement in child conduct problems: ECBI Intensity subscale, completed by mothers at baseline, 8 months postintervention, and 20 months' (final) follow-up 3. Improvement in child conduct problems: observational data on oppositional child behaviour, collected and coded by coders blinded to treatment allocation at baseline, postintervention, and 16 months' follow-up

Jouriles 2009 *(Continued)*

4. Family functioning: PDI Consistency subscale, completed by mothers at baseline, 8 months postintervention, and 20 months' (final) follow-up
5. Personalised treatment outcomes: CTS2 Physical Aggression and Psychological Aggression subscales, completed by mothers at baseline, 8 months postintervention, and 20 months' (final) follow-up
6. Family functioning: mother expressed negative affect and behaviour, collected and coded by coders blinded to treatment allocation at baseline, postintervention, and 16 months' follow-up
7. Personalised treatment outcomes: SCL-90-R completed by mothers at baseline, 8 months postintervention, and 20 months' (final) follow-up
8. Personalised treatment outcomes: IES, completed by mothers at baseline, 8 months postintervention, and 20 months' (final) follow-up

Funding	This research was supported by the following grants to Ernest N. Jouriles: Grant MH-53380, awarded by the National Institute of Mental Health, and Grant 2005-JW-BX-K017, by the Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice. Points of view or opinions in this document are those of the authors and do not represent the official position or policies of the U.S. Department of Justice or other federal agencies.
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "… a separate randomization code for each of the six participating shelters, using a random numbers table." (p 708)
Allocation concealment (selection bias)	Low risk	Quote: "The project coordinator for the evaluation study developed a separate randomization code for each of the six participating shelters, using a random numbers table. The site coordinator for a particular shelter (the project staff person who was responsible for organizing and managing the screening and assessments schedules for that shelter) was informed of the group assignment prior to the first assessment, and mothers were informed of the condition to which she was assigned by the site coordinator after the first assessment was completed." (p 708) Comment: co-ordinators responsible for screening and enrolment unaware of allocation sequence.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: some outcome measures relied on parent report (including primary outcome measures of CBCL and ECBI).
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: 56 participants were retained at or beyond 4th assessment. ITT analyses were undertaken.
Selective reporting (reporting bias)	High risk	Comment: all outcome measures were reported, but measures were not reported for every time point measured.
Other risk of bias: lack of adherence to treatment manual	High risk	Quote: "Among the 32 families assigned to the Project Support condition, an average of 51% (SD 16) of session time was dedicated to the child management skills."

Jouriles 2009 *(Continued)*

Other risk of bias: group differences	High risk	Quote: "Fifteen families (47%) received instruction on all 12 child management skills, 23 (72%) received instruction on at least 8 child management skills, and 26 (81%) received instruction on at least 4 child management skills." (p 710)

Markie-Dadds 2006

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: not reported Setting: rural and remote areas of Western Australia
Participants	Subgroup: parents of children with conduct problems in rural and remote areas Sample size: 41 Age: children: enhanced self-directed condition: M 47.21 (SD 10.19) months, self-directed condition: M 47.27 (SD 9.84) months, WL: M 46.17 (SD 13.29) months; mothers: enhanced self-directed condition: M 34.00 (SD 4.37) years, self-directed condition: M 31.53 (SD 3.46) years, WL: M 30.67 (SD 9.68) years; fathers: enhanced self-directed condition: M 37.38 (SD 6.92) years, self-directed condition: M 33.87 (SD 3.76) years, WL: 35.83 (SD 5.94) years Sex: 76% boys Inclusion criteria: standardised telephone interview used to ensure families who responded to the outreach campaign met following criteria: target child aged 2–6 years; mothers reported they were concerned about their child's behaviour in response to a specific question; child showed no evidence of developmental disorder (e.g. autism) or significant health impairment; child not currently having regular contact with another health professional or agency or taking medication for behavioural problems; parents not currently receiving therapy for psychological problems, not intellectually disabled, and reported they were able to read the newspaper without assistance; mothers rated child's behaviour in the elevated range on the ECBI Intensity score (≥ 127) or Problem score (≥ 11) Exclusion criteria: not reported Comorbidities: not reported
Interventions	1. ESD: families received a 10-unit self-directed programme comprising *Every Parent* and *Every Parent's Workbook*. This positive parenting programme was designed to help parents acquire a variety of skills known to influence children's development. The programme involved teaching parents 17 core child-management strategies. 10 of the strategies are designed to promote children's competence and development and 7 strategies are designed to help parents manage misbehaviour. In addition, parents were taught a 6-step planned activities routine to enhance the generalisation and maintenance of parenting skills. Weekly telephone consultations were arranged, aimed at encouraging parents' problem-solving skills. When parents reported problems with implementing suggested parenting strategies, they were prompted to refer back to the written material provided rather than rely on the practitioner for solutions. Parents were prompted to self-monitor their own and their child's behaviour, to self-select goals and specific behaviours for change, to select strategies to use, to identify their own and their child's strengths and areas for improvement, and to select contingent rewards for themselves and their child. Discussions were restricted to behaviour problems of the target child and elaboration of concepts nominated by the parents as not being well understood. a. Sample size: 14. b. Duration of treatment: not reported. c. Timing and delivery: weekly telephone consultations, with each call lasting 30 minutes or less, with a mean of 20 minutes and a range of 5–30 minutes. Calls were delivered by 1 clinical psychologist.

Markie-Dadds 2006 *(Continued)*

2. Self-directed: families received the same 10-unit self-directed programme comprising *Every Parent* and *Every Parent's Workbook* and the parents were taught the same 6-step planned activities routine to enhance the generalisation and maintenance of parenting skills.
   a. Sample size: 15.
   b. Duration of treatment: not reported.
   c. Timing and delivery: families received a 10-unit self-directed programme comprising *Every Parent* and *Every Parent's Workbook*. Contact with the research team was minimal and there were no face-to-face meetings.
3. WL: families received no treatment and had no contact with the research team for 12 weeks. These families completed the postintervention measures and then received the programme of their choice, namely Enhanced Self-Directed Triple P or Self-Directed Triple P. These families took no further part in the study.
   a. Sample size: 12.
   b. Duration of treatment: not applicable.
   c. Timing and delivery: not applicable.

Outcomes	1. Improvement in child conduct problems: ECBI Intensity and Problem subscales 2. Improvement in child conduct problems: PDR Problem and Target subscales 3. Parental mental health and stress: DASS subscales 4. Parenting skills and knowledge: PS Laxness, Over-reactivity, and Verbosity subscales 5. Parenting skills and knowledge: PSOC Satisfaction and Efficacy subscales 6. Family functioning: PPC Problem and Intensity subscales Completed by both mothers and fathers at pre-intervention, postintervention and 6-month follow-up in the 2 treatment groups, and at pre-intervention and postintervention in WL group
Funding	Not reported
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: some measures relied on parent report (including primary outcome measures of ECBI and PDR).
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "98% families completed pre and postassessment." (p 63)
Selective reporting (reporting bias)	High risk	Comment: outcome measures completed by both mothers and fathers. However, father-reported outcomes were not presented because there were no significant differences on father reports. Therefore, only mother-reported data were provided.

Markie-Dadds 2006 *(Continued)*

Other risk of bias: lack of adherence to treatment manual	Unclear risk	Comment: no information given on treatment fidelity.
Other risk of bias: group differences	Low risk	Quote: "No significant differences among the three groups on any of the demographic characteristics prior to intervention." (p 58)

McCabe 2009

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: not reported Setting: community mental health clinic in US
Participants	Subgroup: Mexican American families Sample size: 58 Age: female primary carers: M 32.2 (SD 8.1) years, male primary carers: M 35.0 (SD 10.3) years; children: PCIT: M 48.9 (SD 9.2) months, Guiando a Niños Activos: M 54.3 (SD 11.6) months, TAU: M 55.1 (SD 15.3) months Sex: children: PCIT: 73.7% boys, Guiando a Niños Activos: 76.2% boys, TAU: 61.1% boys Inclusion criteria: parent identified child as Mexican American and aged 3–7 years; child received score greater than clinical cutpoint on ECBI Intensity subscale; neither parent nor child was participating in any other psychosocial treatment targeting the child's behaviour problems simultaneously Exclusion criteria: not reported Comorbidities: not reported
Interventions	1. PCIT: parents are taught skills to establish a nurturing and secure relationship with their child while increasing their child's prosocial behaviour and decreasing negative behaviour. Therapists actively coach parents and terminate when parents demonstrate mastery of the skills and their child's behaviour is within half an SD of the normative mean on the ECBI Intensity Scale. a. Sample size: 19. b. Duration of treatment: unlimited number of sessions. Therapists were bilingual practicum students and professional psychology doctoral programmes. PCIT and GANA therapists were provided with 40 hours of training on their respective approach by the principal investigator, who was also responsible for supervising both conditions. c. Timing and delivery: timing not reported. 2. GANA: GANA retained the core features of PCIT but modified its delivery to optimise cultural fit for MA families. The GANA programme involved tailoring the delivery of the programme based on a cultural assessment of the family. Cultural concepts were referenced throughout treatment so that the programme could be presented in ways that were congruent with the parents' belief system. Other adaptations to the programme included (a) framing programme as an educational skill building; (b) increasing orientation to therapy; (c) increasing session time for rapport building; (d) translating, simplifying, and adding representations of MA families in written handouts; and (e) implementing an engagement protocol. a. Sample size: 21. b. Duration of treatment: therapists were allowed an unlimited number of sessions. Therapists were bilingual practicum students and professional psychology doctoral programmes. PCIT and GANA

McCabe 2009 *(Continued)*

	therapists were provided with 40 hours of training on their respective approach by the principal investigator, who was also responsible for supervising both conditions. c. Timing and delivery: timing not reported. 3. TAU: families were assigned to therapists without training in PCIT at the same clinic. The 3 TAU therapists described their orientations as 'person-centred cognitive behavioural,' 'trauma focused cognitive behavioural,' and 'family systems' and were allowed complete freedom in the approaches they used. a. Sample size: 18. b. Duration of treatment: therapists were allowed an unlimited number of sessions. Therapists in the TAU condition were supervised by either an LCSW or a Doctorate level clinical psychologist employed by the community mental health clinic. c. Timing and delivery: timing not reported.
Outcomes	Improvement in child conduct problems: ECBI Intensity and Problem subscales, completed by parents at pre-intervention and postintervention; change from baseline reported 1. Improvement in child conduct problems: CBCL Externalizing subscale, completed by parents at pre-intervention and postintervention; change from baseline reported 2. Improvement in child conduct problems: ECBI ODD and CD subscales, completed by parents at pre-intervention and postintervention; change from baseline reported 3. Parenting skills and knowledge: PPS completed by parents at pre-intervention and postintervention; change from baseline reported 4. Personalised treatment outcomes: PSS Total score, completed by parents at pre-intervention and postintervention; change from baseline reported 5. Engagement and decreased dropout: session attendance, reported for all 3 interventions 6. Engagement and decreased dropout: dropout rates, reported for Guiando a Ninos Activos and PCIT
Funding	National Institute of Mental Health grant K01MH01924 to Kristen McCabe
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Quote: "Assessors and families were blind to assignment." (p 755) Comment: families participated in the treatment, therefore they were unlikely to be blinded.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Quote: "Assessors and families were blind to assignment." (p 755) Comment: families participated in the treatment, therefore they were unlikely to be blinded and the primary outcomes relied on parent report (ECBI and CBCL)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "All results reflect intent-to-treat analyses." (p 756)

McCabe 2009 *(Continued)*

Selective reporting (reporting bias)	Low risk	Comment: all outcome measures were reported (p 756).
Other risk of bias: lack of adherence to treatment manual	Low risk	Quote: "Detailed session checklists are included in the PCIT treatment manual, and 82% of items were present." (p 756) Comment: > 80% adherence.
Other risk of bias: group differences	Low risk	Quote: "The three conditions did not differ significantly on any demographic variable, number of sessions attended …, or attrition." (p 756)

Mersky 2016

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 18-month treatment (from September 2011 to March 2013) Setting: US
Participants	Subgroup: foster families Sample size: 91 Age: M 4.6 years Sex: 54% girls Inclusion criteria: aged 3–6 years; placed in a licenced, non-relative foster home; in the clinical range for externalising problems using ECBI according to foster parent ratings Exclusion criteria: children with intellectual, physical, or pervasive developmental disabilities such as autism, deafness, or blindness; cases nearing adoption or reunification (to reduce attrition due to predictable placement change). Only 1 eligible child per foster family was enrolled to reduce threats to group equivalence such as diffusion and burden. Children were not excluded for prior or current receipt of mental health and psychosocial services or psychotropic medication Comorbidities: not reported
Interventions	1. PCIT; group 2: PCIT was implemented using 2 novel modalities: group-based training and telephone consultation. The principles and skills learned in the PCIT workshops were reinforced with periodic telephone consultation and daily homework exercises, activities that aimed to promote rapport between clinician and caregiver, increase treatment adherence, and generalise skills to the home environment. Day 1 of the workshop commenced with didactic instruction provided by a lead PCIT clinician who introduced parents to CDI, the first of 2 PCIT phases. At this time, children engaged in activities with childcare providers in separate rooms. Afterward, children and caregivers were reunited in a group setting to practice CDI exercises facilitated by PCIT trained graduate students enrolled in a Title IV-E child welfare training programme. On a rotating basis, each parent–child dyad was directed to a private clinical room to engage in CDI with the lead clinician for 20 minutes. A parent from a different family also joined the clinician to watch the session, providing an occasion for observational learning. After completing the session, the observing parent moved to the clinical room to engage in active coaching with her foster child while the outgoing parent transitioned to the observation room to watch the session. This rotational process continued throughout the day, interspersed with respite periods, until all parent–child dyads completed ≥ 2 clinical coaching sessions. The training day then adjourned with a group discussion to consolidate knowledge and skills. Day 2 of the workshop resembled the first day in structure, commencing with didactic instruction, followed by experiential exercis-

Mersky 2016 *(Continued)*

es, and ending with group discussion. Day 2 focused on PDI, the second PCIT phase during which the caregiver learns behaviour management and positive discipline skills. During the closing group discussion, parents were prepared for the next phase of the intervention involving in-home practice (i.e. homework) coupled with telephone consultation provided by PCIT clinicians. Following usual PCIT protocol, homework was used to bolster group training, promote overlearning and mastery, and help to ensure that skills were applied in the home.

a. Sample size: 19.
b. Duration of treatment: 3 days of group training and 14 weeks of in-home services.
c. Timing and delivery: a typical workshop included 4–8 parent–child dyads. After completing the 2-day workshop, caregivers from both experimental conditions were asked to complete daily homework exercises and engage in PCIT telephone consultation for 8 weeks. Telephone contact was scheduled weekly for 4 weeks followed by biweekly contact for another 4 weeks. Parent–child dyads attended a 1-day booster training focused on PDI skills, which tend to be more difficult than CDI skills to master. Participants completed 6 more weeks of homework and biweekly telephone consults.

2. PCIT; group 1: same as group 2 but with different durations.
   a. Sample size: 39.
   b. Duration of treatment: 2 days of group training and 8 weeks of in-home services.
   c. Timing and delivery: a typical workshop included 4–8 parent–child dyads. After completing the 2-day workshop, caregivers from both experimental conditions were asked to complete daily homework exercises and engage in PCIT phone consultation for 8 weeks. Telephone contact was scheduled weekly for 4 weeks followed by biweekly contact for another 4 weeks.
3. TAU (group 0): foster parents continued to receive their usual services, including case management and standard parent training. Foster children also continued to receive standard care options designated by their case plan, including medication and other mental health services such as play therapy. After completing their final assessment at 14 weeks' postbaseline, WL controls were eligible to attend PCIT workshops.
   a. Sample size: 33.
   b. Duration of treatment: not applicable.
   c. Timing and delivery: not applicable.

Outcomes	1. Improvement in child conduct problems: ECBI Intensity and Problem subscales 2. Improvement in child conduct problems: CBCL Externalizing subscale Completed by foster parents at baseline, 8 weeks postbaseline; and 14 weeks postbaseline
Funding	Support provided by the National Institutes of Health, National Institute of Child Health and Human Development, Award No. 1R15HD067829-01A1
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias)	High risk	Comment: measures relied on parent report (primary outcomes of ECBI and CBCL).

Mersky 2016 *(Continued)*
All outcomes

Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "24–26%" attrition (p 161). Quote: "All analyses were run in SAS 9.3 under intention-to-treat assumption-s." (p 161)
Selective reporting (reporting bias)	Low risk	Comment: all outcome measures were reported.
Other risk of bias: lack of adherence to treatment manual	Unclear risk	Comment: no implementation information reported.
Other risk of bias: group differences	High risk	Quote: "Study groups were largely equivalent …, tests revealed two significant differences. Significant differences between groups were controlled for in analysis of intervention effects." (p 161)

Nicholson 1999

Study characteristics	
Methods	Allocation: RCT Blindness: assumed parents and therapists not blinded Duration: 10-week treatment and 6-month follow-up Setting: Australia
Participants	Subgroup: step-families Sample size: 60 Age: children: M 9.6 (SD 1.5) years, parents: M 34.2 (SD 5.2) years, step-parents: M 35.7 (SD 6.9) years Sex: 64.3% boys Inclusion: child aged 7–12 years, child reported to display significant oppositional or conduct behaviour problems defined on child receiving a total raw score > 40 on either parent or step-parent completed CBCL and displaying disturbed behaviour for ≥ 6 months, including ≥ 5 symptoms for ODD or 3 symptoms for CD; child resided ≥ 5 days per week with presenting parent–step-parent couple, included legally remarried and cohabiting couples, stepmothers, stepfathers, and blended step-families Exclusion: not reported Comorbidities: not reported
Interventions	1. Therapist-directed programme: intervention for addressing child behaviour problems in stepfamilies in which standard behavioural family intervention approaches were supplemented by components that focused on co-operative parenting and conflict resolution skills. The intervention aimed to reduce the amount of problem behaviour displayed by children, by providing parents and stepparents with knowledge and skills in 5 key areas. The informational content of therapist-directed and self-directed programmes was identical. All participating couples received extensive written materials, with descriptions and monitoring sheets for recommended home-based activities. Therapy sessions had a standardised format and content as outlined in the detailed treatment manual provided to therapists and involved review of homework assignments, discussion and identification of skills deficits, and active skills training. a. Sample size: 20.

Nicholson 1999 *(Continued)*

b. Duration of treatment: 10 weeks.
c. Timing and delivery: couples met on a weekly or fortnightly basis in the clinic with the therapist assigned to their family and sessions typically lasted 1.5–2 hours.

2. Self-directed programme: the informational content of therapist-directed and self-directed programmes was identical; see above. However, in the initial contact with the therapist, the therapist explained the format, aims, and content of the programme, and provided the first module to take home. Each module matched a corresponding therapist-directed intervention session, and both programmes used the same readings and activities.
   a. Sample size: 20.
   b. Duration of treatment: 10 weeks.
   c. Timing and delivery: the intervention was delivered as a self-directed programme, completed by participants at home. Couples receiving the self-directed programme had an initial contact of 15–30 minutes' duration with the therapist assigned to their family. The programme was divided up into modules mailed to the family on a weekly or fortnightly basis, to facilitate the couple's ability to complete the programme in a systematic, paced manner.
3. WL.
   a. Sample size: 20.
   b. Duration of treatment: not applicable.
   c. Timing and delivery: not applicable.

Outcomes	1. Improvement in child conduct problems: CBCL Total subscale, completed by parent and step-parent 2. Improvement in child conduct problems: PDR, completed by parent and step-parent 3. Improvement in child conduct problems: interviewer ratings of clinically significant levels of symptoms for ODD and CD. Parents and step-parents were jointly interviewed by a trained psychologist to determine presence of ODD/CD symptoms 4. Personalised treatment outcomes: CDI, completed by children 5. Personalised treatment outcomes: CMA – Revised version, completed by children 6. Personalised treatment outcomes: SEI Child form, completed by children 7. Personalised treatment outcomes: PPC Parent and Step-parent report, completed by parent and step-parent Completed at pre-intervention, postintervention, and 6-month follow-up
Funding	National Health and Medical Research Council grant (920182) and Public Health Research and Development Council post-doctoral fellowship (954213)
Notes	Additional information: no additional information to report

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: method of random allocation not reported.
Allocation concealment (selection bias)	Unclear risk	Comment: method or awareness of random allocation not reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: reasonable to assume that parents and therapists would have been aware of treatment allocation.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: primary outcome measures, CBCL and PDR, relied on parent report.

Nicholson 1999 *(Continued)*

		Comment: "Clinical interviewer aware of treatment condition" (p 6) who completed the primary outcome measure of interviewer ratings of clinically significant levels of symptoms for OD and CD.
Incomplete outcome data (attrition bias) All outcomes	High risk	Quote: "30% of families dropped out before the completion of intervention." (p 6) Quote: "Outcome data are reported for the 42 families for whom pre- and postintervention assessment data were obtained." (p 10)
Selective reporting (reporting bias)	Low risk	Comment: all outcome measures were reported.
Other risk of bias: lack of adherence to treatment manual	Unclear risk	Comment: information regarding treatment integrity not reported.
Other risk of bias: group differences	Unclear risk	Comment: information about group differences or other treatments not reported.

Parent 2022

Study characteristics	
Methods	Allocation: RCT Blindness: assumed that parents and therapists not blinded Duration: 3-month and 6-month follow-up Setting: low-income families, US
Participants	Subgroup: children with clinically significant problem behaviours from low-income families Sample size: 101 Age: HNC: M 4.28 (SD 1.17) years, TE-HNC: M 4.13 (SD 1.19) years Sex: 54.9% boys, 45.1% girls Inclusion criteria: children with clinically significant problem behaviours (ECBI Problem > 15 or Intensity > 131); recruited via advertisements and flyers distributed at non-profit organisations, local schools, agencies serving low-income families, and word-of-mouth Exclusion criteria: significant developmental or physical impairment (or both) that prohibited use of HNC; if carers had a current mood, psychotic, substance use disorder (or a combination of these), or a pending or prior (or both) substantiated child abuse/neglect case Comorbidities: not reported
Interventions	1. Families received HNC, which is a therapist-delivered, criteria-based (i.e. therapists conduct weekly observation and coding of skill use to determine progression through skills and programme completion) BPT intervention for children with behaviour disorders. HNC included weekly face-to-face therapy sessions, as well as a brief midweek telephone check-in. HNC consists of 2 phases: Differential Attention and Compliance Training. When parents progress to Phase II (i.e. Compliance Training), they continue to practice Phase I skills to maintain skill proficiency. a. Sample size: 54. b. Duration of treatment: not reported.

Parent 2022 *(Continued)*

c. Timing and delivery: HNC included weekly face-to-face therapy sessions (60 minutes), as well as a brief midweek telephone check-in.

2. TE-HNC group received the full HNC protocol augmented with a digital companion, 'Tantrum Tamers'. Tantrum Tamers is a HIPAA-compliant, interactive system that allowed therapists to monitor caregiver activity on the mobile application, as well as tailor the focus and pace of treatment based on parent practice and progress. The Tantrum Tamers application included: (a) daily surveys of skills practice, (b) weekly video-recorded home practice, (c) daily text reminders for skill practice and appointments; (d) video calls with the family midweek to problem-solve obstacles; and (e) skills video series to model new skills and share with other caregivers. Additionally, a homework checklist was added to remind caregivers of daily and weekly assignments. TE-HNC families had access to all app functionality during treatment. After families completed treatment, they had access to a limited range of content, including the skills video series and surveys with automated feedback through their 3-month follow-up assessment when they returned their study phones. From 3- to 6-month follow-ups, TE-HNC parents had access to a programme blog, which provided content intended to remind caregivers of programme content.
   a. Sample size: 47.
   b. Duration of treatment: not reported.
   c. Timing and delivery: TE-HNC included weekly face-to-face therapy sessions (60 minutes), as well as a brief midweek phone check-in, augmented with the Tantrum Tamers digital companion.

Outcomes	1. Improvement in child conduct problems: parent-reported child problem behaviour. At all waves, carers completed the ECBI 2. Personalised treatment outcomes: observed parenting and child compliance
Funding	This study was funded by the National Institute of Mental Health (NIMH) R01MH100377 (D.J.J.; ClinicalTrials.gov Identifier: NCT02191956). Other support was provided by NIMH R21MH113887 (D.J.J.; ClinicalTrials.gov Identifier: NCT03597789), the National Institute of Child Health and Human Development (R.L.; T32HD007377), the National Science Foundation (A.H.; DGE-1650116), the National Institute of Minority Health and Health Disparities (J.P., PI; R01MD015401), and the National Institute on Drug Abuse (M.H.; T32DA043449).
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: eligible families randomised 1:1 to either HNC or TE-HNC.
Allocation concealment (selection bias)	Unclear risk	Comment: did not report whether they had used any process to conceal the random allocation.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Comment: master's-level therapists treated families in both groups.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: assessors of the primary outcome measures within each study were carers, teachers, or clinicians who were not blinded to the group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: overall, rates of retention by group were similar across follow-ups. Patterns of missingness did not differ by treatment condition; child age, sex, race/ethnicity; carer age, ethnicity/race; or family economic stress (all P > 0.05). Further, random patterns of missingness along with a non-significant Little's missing completely at random (MCAR) test, v2(239) = 267.64, P > 0.05,

Parent 2022 *(Continued)*

		suggest that the mechanism of missingness was MCAR and support use of multiple imputation and FIML for primary analyses (p 4).
Selective reporting (reporting bias)	Low risk	Comment: within-group effect sizes used to examine magnitude in change in child problem behaviours and observed parenting comparing pretreatment and post-treatment, before 3-month follow-up, and before 6-month follow-up, controlling for the correlation between each time point, for each group. FIML techniques were used for inclusion of all available data. Primary outcomes were child behaviour problems and secondary outcomes were parenting skills.
Other risk of bias: lack of adherence to treatment manual	Low risk	No other biases were identified regarding adherence to the treatment manual.
Other risk of bias: group differences	Low risk	No other biases were identified regarding group differences.

ADHD: attention deficit hyperactivity disorder; BDI: Beck Depression Inventory; BPT: Behavioral Parent Training; CBCL: Child Behaviour Checklist; CBT: cognitive behavioural therapy; CD: conduct disorder; CDI: Child-Directed Interaction; CGI: Clinical Global Impression Scale; CMA: Child Manifest Anxiety Scale; CPRS-R: Conners' Parent Rating Scale – Revised; CPS: Collaborative Problem Solving; CST: Coping Skills Training; CTS2: Conflict Tactics Scale – Revised; DASS: Depression Anxiety Stress Scale; DISCAP: Diagnostic Interview Schedule For Children, Adolescents, and Parents; DISYPS-II: German Diagnostic System For Children And Adolescents; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders – 4th Edition; DSM: Diagnostic And Statistical Manual of Mental Disorders; DV: domestic violence; EBFI: Enhanced Behavioural Family Intervention; ECBI: Eyberg Child Behavior Inventory; ECI: Early Childhood Inventory; ESD: Enhanced Self-Directed Triple P; FAVK: Questionnaire for Aggressive Behavior of Children; FIML: full information maximum likelihood; FOS: Family Observation Schedule; FTF: face-to-face; GAF: Global Assessment of Functioning; GANA: Guiando A Niños Activos; HNC: Helping the Noncompliant Child; ICD: International Classification of Diseases and Related Health Problems; IES: Impact Event Scales; IPV: interpersonal violence; IQ: intelligence quotient; ISSB: Inventory of Socially Supportive Behaviours; ITT: intention-to-treat; LCSW: Licenced Clinical Social Worker; M: mean; MA: Mexican American; MCAR: missing completely at random; n: number; ODD: oppositional defiant disorder; PCIT: Extended Parent–Child Interaction Therapy; PCIT: Parent–Child Interaction Therapy; PCRI: Parent–Child Relationship Inventory; PDI: Parenting Dimensions Inventory; PDR: Parent Daily Report; PLAY: Group Play; PPC: Parent Problem Checklist; PPS: Parenting Practices Scale; PS: Parenting Scale; PSI: Parenting Stress Index; PSOC: Parenting Sense of Competency Scale; PSS-Fa: Perceived Social Support From Family; PSS-Fr: Perceived Social Support From Friends; PSS: Parenting Stress Scale; PST: Partner Support Training; PT: parent training; RBPC: Revised Behaviour Problem Checklist; RCT: randomised controlled trial; SAS: Statistical Analysis System; SCL-90-R: Symptom Checklist-90 – Revised; SCL-DBD: German Symptom Checklist For Disruptive Behavior Disorder; SD: standard deviation; SDQ: Strengths And Difficulties Questionnaire; SEI: Coopersmith Self-Esteem Inventory; SPST: Social Problem-Solving Test; TAU: treatment as usual; TE-HNC: Technology-Enhanced Helping the Noncompliant Child; THAV: Treatment Program For Children With Aggressive Behavior; TRF: Teacher Report Form; WISC-III: Wechsler Intelligence Scale for Children, 3rd Edition; WL: waitlist control group.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Basile 1993	Study did not set out to measure our prespecified outcomes.
Costantino 1984	Study did not set out to measure our prespecified outcomes.
Eckenrode 2010	Study did not set out to measure our prespecified outcomes.
Eddy 2003	Study did not set out to measure our prespecified outcomes.
Graham 2015	Study did not set out to measure our prespecified outcomes.
Oden 1982	Study did not set out to measure our prespecified outcomes.

Characteristics of studies awaiting classification *[ordered by study ID]*

Bloom 1980

Methods	Allocation: randomly assigned Blindness: not reported Duration: not reported Setting: not reported
Participants	Subgroup: children with concomitant language disorder and low socioeconomic status Sample size: 60 Age: 7 to 11 years and 11 months Sex: 50% boys and 50% girls Inclusion criteria: behavioural disorder, language impairment, aged 7–11 years and 11 months, either sex (an equal distribution of boys and girls were randomly assigned to 4 groups), IQ in the range of the normal population, low socioeconomic status. Exclusion criteria: not reported
Interventions	1. Structured language therapy programme a. Sample size: 15 b. Duration of treatment: not reported c. Timing and delivery: timing not reported; 6 teachers of the speech and language impaired 2. Control group I: no language therapy a. Sample size: 15 b. Duration of treatment: not reported c. Timing and delivery: not reported 3. Control group II: received informal support; consisted of games and recreational activities a. Sample size: 15 b. Duration of treatment: not reported c. Timing and delivery: timing not reported; 5 social workers 4. Comparison group: no language therapy as their language abilities were normal a. Sample size: 15 b. Duration of treatment: not reported c. Timing and delivery: not reported
Outcomes	1. Relationship between language disorder and academic achievement – the Wide Range Achievement Test 2. Relationship between language disorder and self-esteem – Coopersmith Self-Esteem Inventory 3. Relationship between language disorder and symptoms of behavioural disorder – The Jesness Inventory 4. The effects of language therapy on variables associated with behavioural disorder – The Jesness Inventory Timing of outcome assessment: not reported for any measure
Notes	Additional information: no additional information to report

Greene 1999

Methods	Allocation: not reported Blindness: not reported Duration: not reported Setting: not reported
Participants	Subgroup: children with bipolar disorder Sample size: not reported Age: not reported Sex: not reported Inclusion criteria: not reported Exclusion criteria: not reported
Interventions	1. Operant parent management training a. Sample size: not reported b. Duration of treatment: not reported c. Timing and delivery: not reported 2. Cognitive behavioural therapy a. Sample size: not reported b. Duration of treatment: not reported c. Timing and delivery: not reported
Outcomes	Not reported
Notes	Additional information: no additional information to report

Jones 1991

Methods	Allocation: randomised Blindness: not reported Duration: not reported Setting: not reported
Participants	Subgroup: impulsive male children Sample size: not reported Age: 9–12 years Sex: boys Inclusion criteria: not reported

Jones 1991 *(Continued)*

	Exclusion criteria: not reported
Interventions	1. Multicomponent cognitive-behavioural group training programme: with the addition of brief teacher and parent training a. Sample size: not reported b. Duration of treatment: 10 weeks c. Timing and delivery: not reported 2. Control group: no additional information reported a. Sample size: not reported b. Duration of treatment: not reported c. Timing and delivery: not reported
Outcomes	1. Problem-solving skills, measured at pre- and postintervention. Measure not reported 2. Behavioural adjustment, measured at pre- and postintervention. Measure not reported
Notes	Additional information: no additional information to report

Robbins 1988

Methods	Allocation: randomly assigned Blindness: not reported Duration: June 1986 to September 1987; 2-week baseline phase, 6-week treatment phase, 2-week follow-up phase Setting: children's inpatient unit of a major New York City psychiatric hospital, US
Participants	Subgroup: children hospitalised for severe CDs Sample size: 14 Age: 5–12 years Sex: not reported Inclusion criteria: not reported Exclusion criteria: not reported
Interventions	1. Movement: groups employed 3 specifically designed movement strategies: change in context, re-combination and substitution, and leader intervention a. Sample size: 8 b. Duration of treatment: 6 weeks c. Timing and delivery: group met twice-weekly for 45 minutes 2. Control: no additional information a. Sample size: 6 b. Duration of treatment: 6 weeks

Robbins 1988 *(Continued)*

c. Timing and delivery: group met twice-weekly for 45 minutes

Outcomes	1. General level of pathology and Anger/Hostility Standard Score (CBI) completed at pre- and postintervention by staff members of the hospital's Biometrics Department who had no involvement in the study 2. Aggression Incident Frequency completed twice-weekly throughout the study by raters from the hospital's Child Life Department who were unaware of the group to which the child was assigned
Notes	Additional information: no additional information to report

Walker 1984

Methods	Allocation: randomly assigned Blindness: not reported Duration: not reported Setting: incoming families to the Family Counseling Project at Indiana State University, US
Participants	Subgroup: boys Sample size: 43 (we assume 6 participants dropped out due to the discrepancies between the sample size, and the subsequent experimental and control participant numbers) Age: second to sixth grade Sex: boys Inclusion criteria: not reported Exclusion criteria: not reported
Interventions	1. Social learning treatment: time-limited family counselling model, troubled families a. Sample size: 20 b. Duration of treatment: not reported c. Timing and delivery: not reported 2. Waitlist control: no additional information reported a. Sample size: 17 b. Duration of treatment: not reported c. Timing and delivery: not reported
Outcomes	1. Negative behaviour: PDR completed by parents at pre- and postintervention 2. Aggression: ACBC completed by parents at pre- and postintervention 3. Conflict: FES completed by parents at pre- and postintervention 4. Target behaviour: DBC completed by classroom teacher at pre- and postintervention

Walker 1984 *(Continued)*

Notes	Additional information: no additional information to report

ACBC: Achenbach Child Behavior Checklist; CBI: Children's Behavior Inventory; DBC: Daily Behavior Checklist; FES: Family Environment Scale; IQ: intelligence quotient; PDR: Parent Daily Report.

DATA AND ANALYSES

Comparison 1. Personalised intervention versus non-personalised intervention

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Improvement in child conduct problems: Eyberg Child Behavior Inventory (ECBI), Problem (short term)	6	278	Mean Difference (IV, Random, 95% CI)	-3.04 [-6.06, -0.02]
1.2 Improvement in child conduct problems: ECBI, Intensity (short term)	6	278	Mean Difference (IV, Random, 95% CI)	-6.25 [-16.66, 4.15]
1.3 Improvement in child conduct problems: ECBI, Problem (medium term)	3	186	Mean Difference (IV, Random, 95% CI)	-5.08 [-6.03, -4.14]
1.4 Improvement in child conduct problems: ECBI, Intensity (medium term)	3	186	Mean Difference (IV, Fixed, 95% CI)	-8.98 [-12.95, -5.01]
1.5 Improvement in child conduct problems: Child Behaviour Checklist (CBCL), Externalising (short term)	3	189	Mean Difference (IV, Random, 95% CI)	-2.19 [-6.97, 2.59]
1.6 Improvement in child conduct problems: Parent Daily Report (PDR), mean Daily (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.7 Improvement in child conduct problems: PDR, mean Daily (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.8 Improvement in child conduct problems: PDR, mean Target (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.9 Improvement in child conduct problems: PDR, mean Target (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.10 Improvement in child conduct problems: ECBI, Problem (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.11 Improvement in child conduct problems: ECBI, Intensity (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.12 Improvement in child conduct problems: ECBI, Oppositional Defiant (OD) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.13 Improvement in child conduct problems: ECBI, OD (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.14 Improvement in child conduct problems: ECBI, Conduct Disorder (CD) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.15 Improvement in child conduct problems: ECBI, CD (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.16 Improvement in child conduct problems: clinical symptom severity of oppositional defiant disorder (ODD)/CD (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.17 Improvement in child conduct problems: clinical symptom severity of ODD/CD (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.18 Improvement in child conduct problems: ODD/CD diagnosis (short term)	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.19 Improvement in child conduct problems: ODD/CD diagnosis (medium term)	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.20 Improvement in child conduct problems: Conners' Parent Rating Scale – Revised (CPRS-R) oppositional behaviour (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.21 Improvement in child conduct problems: CPRS-R oppositional behaviour (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.22 Improvement in child conduct problems: Strengths and Difficulties Questionnaire Total (SDQ-T) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.23 Improvement in child conduct problems: SDQ-T (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.24 Improvement in child conduct problems: PDR, Total (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.25 Improvement in child conduct problems: PDR, Total (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.26 Improvements in child conduct problems: Symptom Checklist For Disruptive Behavior Disorder (SCL-DBD), ODD – parent (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.27 Improvements in child conduct problems: SCL-DBD, ODD - teacher (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.28 Improvements in child conduct problems: SCL-DBD, CD – parent (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.29 Improvements in child conduct problems: SCL-DBD, CD – teacher (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.30 Improvement in child conduct problems: Teacher Report Form (TRF), Externalized Behaviour (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.31 Improvement in child conduct problems: Social Problem-Solving Test (SPST) aggressive behaviour (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.32 Improvement in child conduct problems: CBCL, Externalising (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.33 Improvement in child conduct problems: PDR, Target (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.34 Improvement in child conduct problems: PDR, Target (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.35 Improvement in child conduct problems: PDR, Problem (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.36 Improvement in child conduct problems: PDR, Problem (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.37 Improvement in child conduct problems: CBCL, Externalising (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.38 Personalised treatment outcomes: ECBI, Inattention (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.39 Personalised treatment outcomes: ECBI, Inattention (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.40 Personalised treatment outcomes: Beck Depression Inventory (BDI) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.41 Personalised treatment outcomes: BDI (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.42 Personalised treatment outcomes: Inventory of Socially Supportive Behaviours (ISSB) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.43 Personalised treatment outcomes: ISSB (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.44 Personalised treatment outcomes: Perceived Social Support from Family (PSS-Fa) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.45 Personalised treatment outcomes: PSS-Fa (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

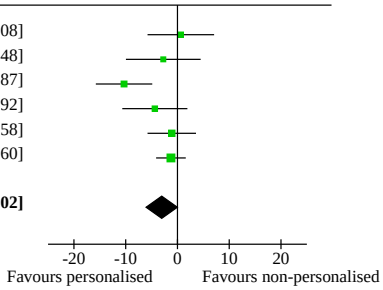
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.46 Personalised treatment outcomes: Perceived Social Support from Friends (PSS-Fr) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.47 Personalised treatment outcomes: PSS-Fr (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.48 Personalised treatment outcomes: Questionnaire for Aggressive Behavior of Children (FAVK) Peer Aggression – parent (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.49 Personalised treatment outcomes: FAVK Peer Aggression – teacher (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.50 Personalised treatment outcomes: SCL-DBD, prosocial behaviour - parent (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.51 Personalised treatment outcomes: SCL-DBD, prosocial behaviour - teacher (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.52 Personalised treatment outcomes: SPST, socially competent behaviour (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.53 Parenting skills and knowledge: Family Observation System (FOS), Negative Parent Behaviour (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.54 Parenting skills and knowledge: FOS, Negative Parent Behaviour (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.55 Parenting skills and knowledge: Parenting Scale (PS) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.56 Parenting skills and knowledge: PS (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.57 Parenting skills and knowledge: Parenting Sense of Competency (PSOC) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.58 Parenting skills and knowledge: PSOC (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.59 Parenting skills and knowledge: FOS, Parental Correct Implementation (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.60 Parenting skills and knowledge: FOS, Parental Correct Implementation (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.61 Parenting skills and knowledge: PS, Laxness (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.62 Parenting skills and knowledge: PS, Laxness (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.63 Parenting skills and knowledge: PS, Over-reactivity (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.64 Parenting skills and knowledge: PS, Over-reactivity (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.65 Parenting skills and knowledge: PS, Verbosity (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.66 Parenting skills and knowledge: PS, Verbosity (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.67 Parenting skills and knowledge: PSOC, Satisfaction (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.68 Parenting skills and knowledge: PSOC, Satisfaction (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.69 Parenting skills and knowledge: PSOC, Efficacy (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.70 Parenting skills and knowledge: PSOC, Efficacy (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.71 Parenting skills and knowledge: Parenting Practices Scale (PPS) (pre–post)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.72 Family functioning: Parent Problem Checklist (PPC), Problem (short term)	2	64	Mean Difference (IV, Fixed, 95% CI)	0.30 [-1.23, 1.83]
1.73 Family functioning: PPC, Problem (long term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.74 Family functioning: PPC, Intensity (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.75 Family functioning: PPC, Problem (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.76 Family functioning: PPC, Intensity (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.77 Engagement and decreased dropout: session attendance (%)	3	156	Mean Difference (IV, Random, 95% CI)	0.97 [-0.32, 2.27]
1.78 Engagement and decreased dropout: mid-week call availability (%)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.79 Engagement and decreased dropout: dropout rates	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected

Analysis 1.1. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 1: Improvement in child conduct problems: Eyberg Child Behavior Inventory (ECBI), Problem (short term)

	Personalised			Non-personalised				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Bor 2002	10.46	10.88	15	9.81	7.76	21	13.4%	0.65 [-5.78 , 7.08]	
Jones 2014	6.14	5.67	8	8.88	8.17	7	11.6%	-2.74 [-9.96 , 4.48]	
Markie-Dadds 2006	2.15	3.36	13	12.47	10.15	15	16.2%	-10.32 [-15.77 , -4.87]	
McCabe 2009	7.35	9	21	11.72	11.06	19	13.8%	-4.37 [-10.66 , 1.92]	
Mersky 2016	10.8	8.587	19	11.9	8.4307	39	18.8%	-1.10 [-5.78 , 3.58]	
Parent 2022	12.39	7.9	47	13.65	6.6	54	26.2%	-1.26 [-4.12 , 1.60]	
Total (95% CI)			**123**			**155**	**100.0%**	**-3.04 [-6.06 , -0.02]**	

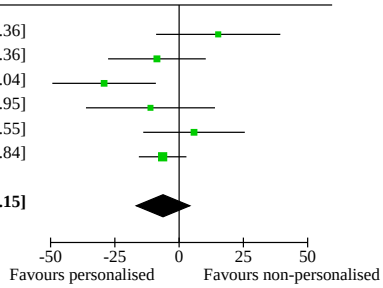
Heterogeneity: Tau² = 6.91; Chi² = 10.23, df = 5 (P = 0.07); I² = 51%
Test for overall effect: Z = 1.97 (P = 0.05)
Test for subgroup differences: Not applicable



Analysis 1.2. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 2: Improvement in child conduct problems: ECBI, Intensity (short term)

	Personalised			Non-personalised				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Bor 2002	132.07	38.31	15	116.86	33.68	21	12.3%	15.21 [-8.94 , 39.36]	
Jones 2014	83	15.34	8	91.63	21.25	7	16.5%	-8.63 [-27.62 , 10.36]	
Markie-Dadds 2006	100.69	17.41	13	129.87	35.12	15	15.4%	-29.18 [-49.32 , -9.04]	
McCabe 2009	84.3	34.4	21	95.44	45.2	19	11.7%	-11.14 [-36.23 , 13.95]	
Mersky 2016	133.8	36.3096	19	128	35.4091	39	15.8%	5.80 [-13.95 , 25.55]	
Parent 2022	103.84	23.7	47	110.25	23.6	54	28.3%	-6.41 [-15.66 , 2.84]	
Total (95% CI)			**123**			**155**	**100.0%**	**-6.25 [-16.66 , 4.15]**	

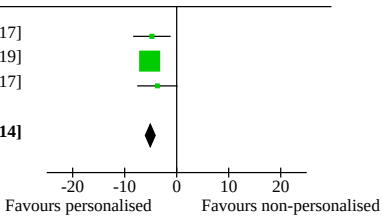
Heterogeneity: Tau² = 77.21; Chi² = 9.65, df = 5 (P = 0.09); I² = 48%
Test for overall effect: Z = 1.18 (P = 0.24)
Test for subgroup differences: Not applicable



Analysis 1.3. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 3: Improvement in child conduct problems: ECBI, Problem (medium term)

	Personalised			Non-personalised				Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Markie-Dadds 2006	3.23	3.61	13	8	5.76	14	6.9%	-4.77 [-8.37 , -1.17]	
Mersky 2016	5.4	2.03	19	10.6	1.38	39	87.2%	-5.20 [-6.21 , -4.19]	
Parent 2022	10.01	9.07	47	13.72	10.8	54	5.9%	-3.71 [-7.59 , 0.17]	
Total (95% CI)			**79**			**107**	**100.0%**	**-5.08 [-6.03 , -4.14]**	

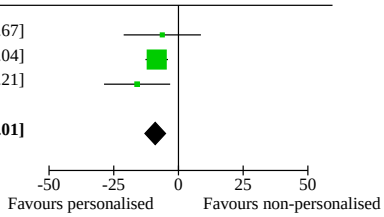
Heterogeneity: Tau² = 0.00; Chi² = 0.56, df = 2 (P = 0.75); I² = 0%
Test for overall effect: Z = 10.56 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 1.4. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 4: Improvement in child conduct problems: ECBI, Intensity (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	101.32	12.69	13	107.57	25.26	14	7.1%	-6.25 [-21.17 , 8.67]
Mersky 2016	118.2	8.74	19	126.6	5.98	39	83.2%	-8.40 [-12.76 , -4.04]
Parent 2022	102.37	30.1	47	118.35	35.39	54	9.7%	-15.98 [-28.75 , -3.21]
Total (95% CI)			**79**			**107**	**100.0%**	**-8.98 [-12.95 , -5.01]**

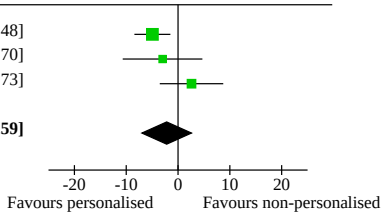
Heterogeneity: Chi² = 1.35, df = 2 (P = 0.51); I² = 0%
Test for overall effect: Z = 4.43 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 1.5. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 5: Improvement in child conduct problems: Child Behaviour Checklist (CBCL), Externalising (short term)

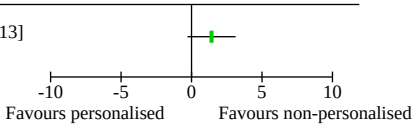
Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Görtz-Dorten 2019	14.52	7.26	50	19.47	9.24	41	46.0%	-4.95 [-8.42 , -1.48]
McCabe 2009	45.83	11.28	21	48.82	13.31	19	23.6%	-2.99 [-10.68 , 4.70]
Mersky 2016	22.3	11.3767	19	19.7	10.7414	39	30.4%	2.60 [-3.53 , 8.73]
Total (95% CI)			**90**			**99**	**100.0%**	**-2.19 [-6.97 , 2.59]**

Heterogeneity: Tau² = 9.79; Chi² = 4.42, df = 2 (P = 0.11); I² = 55%
Test for overall effect: Z = 0.90 (P = 0.37)
Test for subgroup differences: Not applicable



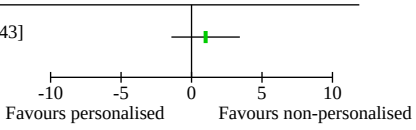
Analysis 1.6. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 6: Improvement in child conduct problems: Parent Daily Report (PDR), mean Daily (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	5.03	2.61	15	3.61	2.53	21	1.42 [-0.29 , 3.13]



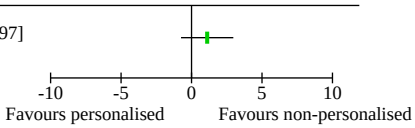
Analysis 1.7. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 7: Improvement in child conduct problems: PDR, mean Daily (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	5.55	3.61	13	4.55	3.18	19	1.00 [-1.43 , 3.43]



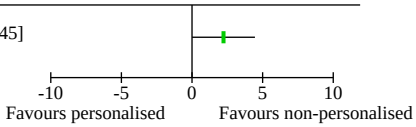
Analysis 1.8. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 8: Improvement in child conduct problems: PDR, mean Target (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	3.54	2.99	15	2.43	2.51	21	1.11 [-0.75 , 2.97]



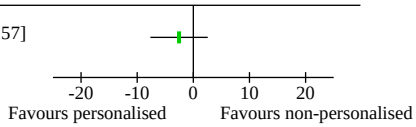
Analysis 1.9. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 9: Improvement in child conduct problems: PDR, mean Target (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	4.63	3.87	13	2.4	1.61	19	2.23 [0.01 , 4.45]



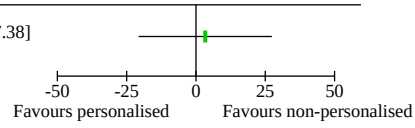
Analysis 1.10. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 10: Improvement in child conduct problems: ECBI, Problem (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	9.18	6.79	13	11.72	7.85	19	-2.54 [-7.65 , 2.57]



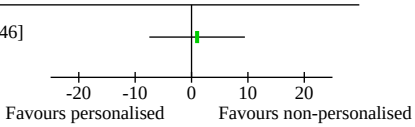
Analysis 1.11. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 11: Improvement in child conduct problems: ECBI, Intensity (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	126	35.67	13	122.65	31.57	19	3.35 [-20.68 , 27.38]



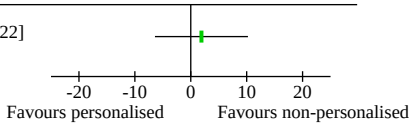
Analysis 1.12. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 12: Improvement in child conduct problems: ECBI, Oppositional Defiant (OD) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	39.27	12.41	15	38.28	13.3	21	0.99 [-7.48 , 9.46]



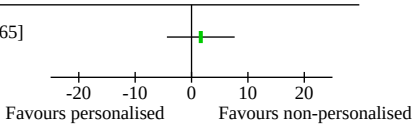
Analysis 1.13. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 13: Improvement in child conduct problems: ECBI, OD (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	39.91	11.37	13	38	12.37	19	1.91 [-6.40 , 10.22]	



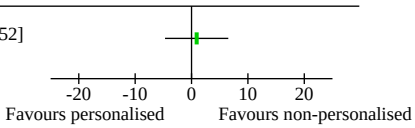
Analysis 1.14. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 14: Improvement in child conduct problems: ECBI, Conduct Disorder (CD) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	21.47	9.52	15	19.83	8.4	21	1.64 [-4.37 , 7.65]	



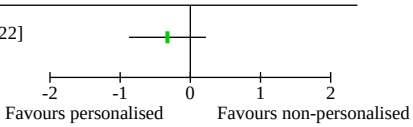
Analysis 1.15. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 15: Improvement in child conduct problems: ECBI, CD (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	21.91	7.83	13	21	8.13	19	0.91 [-4.70 , 6.52]	



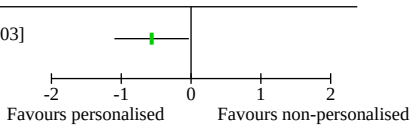
Analysis 1.16. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 16: Improvement in child conduct problems: clinical symptom severity of oppositional defiant disorder (ODD)/CD (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 2019	2.364	1.47	55	2.69	1.501	58	-0.33 [-0.87 , 0.22]	



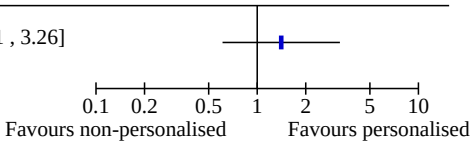
Analysis 1.17. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 17: Improvement in child conduct problems: clinical symptom severity of ODD/CD (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 2019	2.321	1.541	53	2.885	1.231	52	-0.56 [-1.10 , -0.03]	



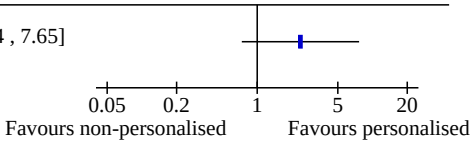
Analysis 1.18. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 18: Improvement in child conduct problems: ODD/CD diagnosis (short term)

Study or Subgroup	Personalised Events	Total	Non-personalised Events	Total	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI
Dadds 2019	15	55	6	31	1.41 [0.61 , 3.26]	



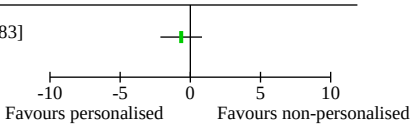
Analysis 1.19. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 19: Improvement in child conduct problems: ODD/CD diagnosis (medium term)

Study or Subgroup	Personalised Events	Total	Non-personalised Events	Total	Risk Ratio M-H, Fixed, 95% CI	Risk Ratio M-H, Fixed, 95% CI
Dadds 2019	13	53	3	29	2.37 [0.74 , 7.65]	



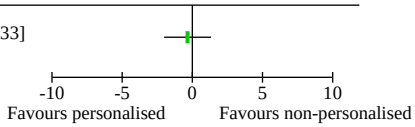
Analysis 1.20. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 20: Improvement in child conduct problems: Conners' Parent Rating Scale – Revised (CPRS-R) oppositional behaviour (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 2019	9	3.61	45	9.65	3.6	46	-0.65 [-2.13 , 0.83]	



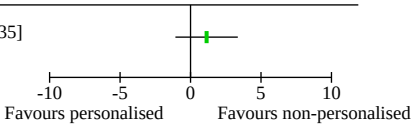
Analysis 1.21. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 21: Improvement in child conduct problems: CPRS-R oppositional behaviour (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 2019	9.44	4.01	45	9.78	4.11	46	-0.34 [-2.01 , 1.33]	



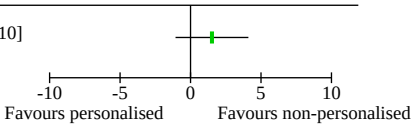
Analysis 1.22. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 22: Improvement in child conduct problems: Strengths and Difficulties Questionnaire Total (SDQ-T) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 2019	18.36	5.25	45	17.22	5.52	46	1.14 [-1.07 , 3.35]	



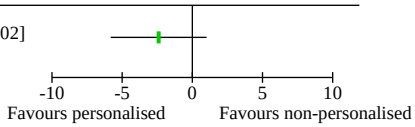
Analysis 1.23. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 23: Improvement in child conduct problems: SDQ-T (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 2019	18.49	6.92	45	16.97	5.57	46	1.52 [-1.06 , 4.10]	



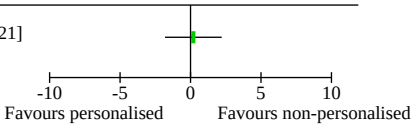
Analysis 1.24. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 24: Improvement in child conduct problems: PDR, Total (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 1992	6.1	2.2	10	8.5	5.3	11	-2.40 [-5.82 , 1.02]	



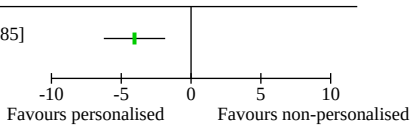
Analysis 1.25. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 25: Improvement in child conduct problems: PDR, Total (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 1992	6.3	1.7	10	6.1	2.9	11	0.20 [-1.81 , 2.21]	



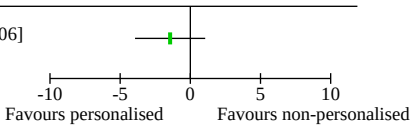
Analysis 1.26. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 26: Improvements in child conduct problems: Symptom Checklist For Disruptive Behavior Disorder (SCL-DBD), ODD – parent (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	9.36	4.86	50	13.41	5.67	41	-4.05 [-6.25 , -1.85]	



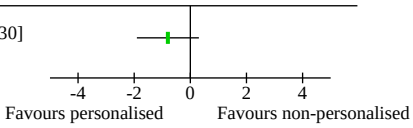
Analysis 1.27. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 27: Improvements in child conduct problems: SCL-DBD, ODD - teacher (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	10.53	5.22	50	11.97	6.66	41	-1.44 [-3.94 , 1.06]	



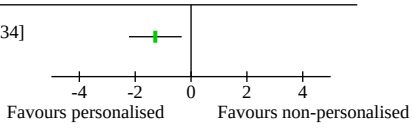
Analysis 1.28. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 28: Improvements in child conduct problems: SCL-DBD, CD – parent (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	2.88	2.4	50	3.68	2.88	41	-0.80 [-1.90 , 0.30]	



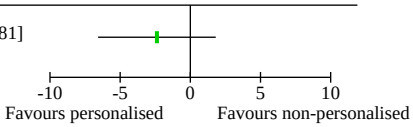
Analysis 1.29. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 29: Improvements in child conduct problems: SCL-DBD, CD – teacher (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	2.4	1.6	50	3.68	2.72	41	-1.28 [-2.22 , -0.34]	



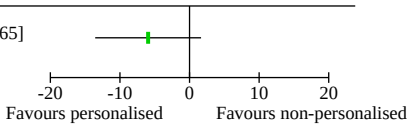
Analysis 1.30. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 30: Improvement in child conduct problems: Teacher Report Form (TRF), Externalized Behaviour (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	20.74	9.18	50	23.12	10.88	41	-2.38 [-6.57 , 1.81]	



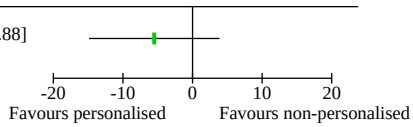
Analysis 1.31. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 31: Improvement in child conduct problems: Social Problem-Solving Test (SPST) aggressive behaviour (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	17.45	18	50	23.4	18.75	41	-5.95 [-13.55 , 1.65]	



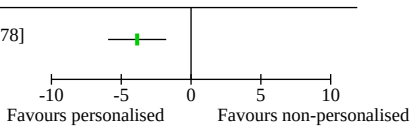
Analysis 1.32. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 32: Improvement in child conduct problems: CBCL, Externalising (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Jouriles 2001	54.5	11.5	13	60	14.7	17	-5.50 [-14.88 , 3.88]	



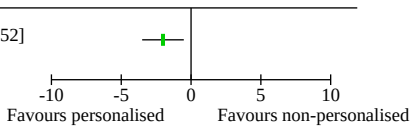
Analysis 1.33. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 33: Improvement in child conduct problems: PDR, Target (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	0.5	0.56	13	4.36	4.07	15	-3.86 [-5.94 , -1.78]	



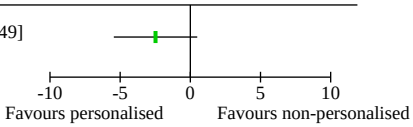
Analysis 1.34. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 34: Improvement in child conduct problems: PDR, Target (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	0.53	0.74	13	2.54	2.74	14	-2.01 [-3.50 , -0.52]	



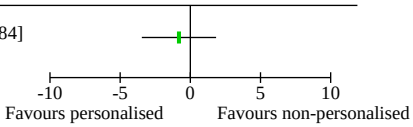
Analysis 1.35. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 35: Improvement in child conduct problems: PDR, Problem (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	3.76	2.76	13	6.24	5.07	15	-2.48 [-5.45 , 0.49]



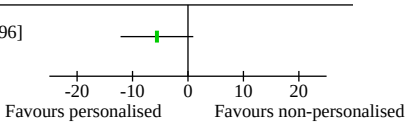
Analysis 1.36. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 36: Improvement in child conduct problems: PDR, Problem (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	3.63	3.26	13	4.44	3.75	14	-0.81 [-3.46 , 1.84]



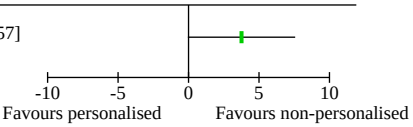
Analysis 1.37. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 37: Improvement in child conduct problems: CBCL, Externalising (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Mersky 2016	17.2	11.9434	19	22.8	11.0296	33	-5.60 [-12.16 , 0.96]



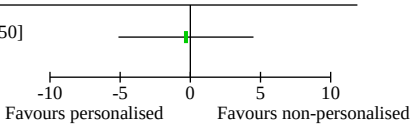
Analysis 1.38. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 38: Personalised treatment outcomes: ECBI, Inattention (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	18.13	5.58	15	14.37	5.97	21	3.76 [-0.05 , 7.57]



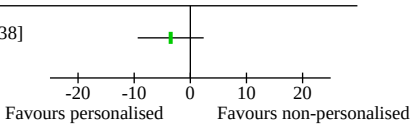
Analysis 1.39. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 39: Personalised treatment outcomes: ECBI, Inattention (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	15	6.97	13	15.31	6.59	19	-0.31 [-5.12 , 4.50]



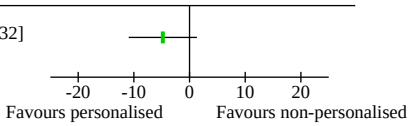
Analysis 1.40. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 40: Personalised treatment outcomes: Beck Depression Inventory (BDI) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Dadds 1992	7.5	6	10	11	7.7	11	-3.50 [-9.38 , 2.38]



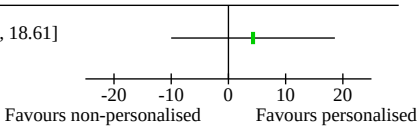
Analysis 1.41. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 41: Personalised treatment outcomes: BDI (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Dadds 1992	6.6	4.9	10	11.4	9	11	-4.80 [-10.92 , 1.32]



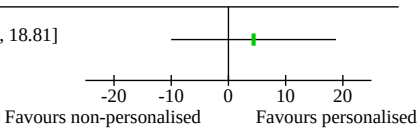
Analysis 1.42. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 42: Personalised treatment outcomes: Inventory of Socially Supportive Behaviours (ISSB) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Dadds 1992	67.9	19.99	10	63.6	12.1	11	4.30 [-10.01 , 18.61]



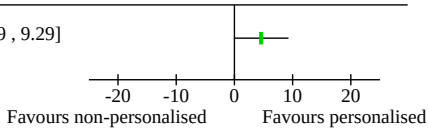
Analysis 1.43. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 43: Personalised treatment outcomes: ISSB (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Dadds 1992	73.1	17.2	10	68.7	16.4	11	4.40 [-10.01 , 18.81]



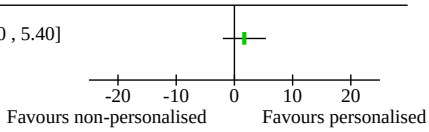
Analysis 1.44. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 44: Personalised treatment outcomes: Perceived Social Support from Family (PSS-Fa) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 1992	15.1	5.8	10	10.5	5.1	11	4.60 [-0.09 , 9.29]	



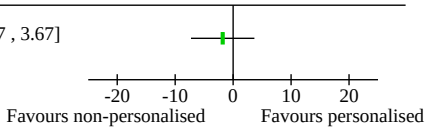
Analysis 1.45. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 45: Personalised treatment outcomes: PSS-Fa (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 1992	14	4.6	10	12.3	4	11	1.70 [-2.00 , 5.40]	



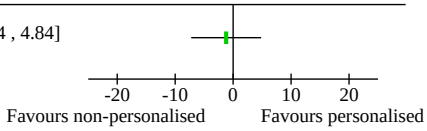
Analysis 1.46. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 46: Personalised treatment outcomes: Perceived Social Support from Friends (PSS-Fr) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 1992	8.7	6.8	10	10.5	5.9	11	-1.80 [-7.27 , 3.67]	



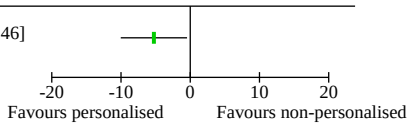
Analysis 1.47. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 47: Personalised treatment outcomes: PSS-Fr (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Dadds 1992	8.5	7	10	9.7	7.1	11	-1.20 [-7.24 , 4.84]	



Analysis 1.48. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 48: Personalised treatment outcomes: Questionnaire for Aggressive Behavior of Children (FAVK) Peer Aggression – parent (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	23.25	10	50	28.5	12.75	41	-5.25 [-10.04 , -0.46]	

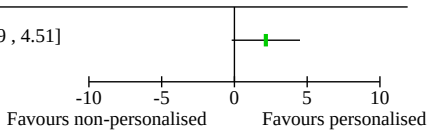


Analysis 1.49. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 49: Personalised treatment outcomes: FAVK Peer Aggression – teacher (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	27	12.5	50	30.5	17.25	41	-3.50 [-9.82 , 2.82]	

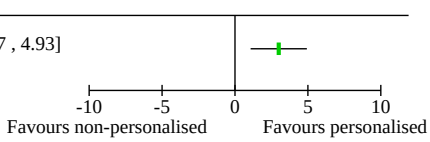
Analysis 1.50. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 50: Personalised treatment outcomes: SCL-DBD, prosocial behaviour - parent (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	22.92	4.92	50	20.76	6.24	41	2.16 [-0.19 , 4.51]	



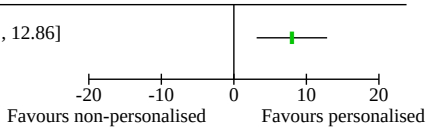
Analysis 1.51. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 51: Personalised treatment outcomes: SCL-DBD, prosocial behaviour - teacher (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	20.88	4.8	50	17.88	4.56	41	3.00 [1.07 , 4.93]	



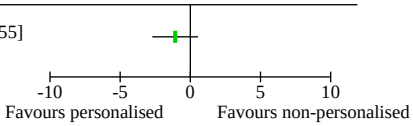
Analysis 1.52. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 52: Personalised treatment outcomes: SPST, socially competent behaviour (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Görtz-Dorten 2019	31.9	12.95	50	23.9	10.7	41	8.00 [3.14 , 12.86]



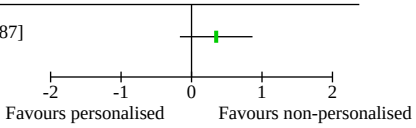
Analysis 1.53. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 53: Parenting skills and knowledge: Family Observation System (FOS), Negative Parent Behaviour (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	0.53	1.04	15	1.61	3.6	21	-1.08 [-2.71 , 0.55]



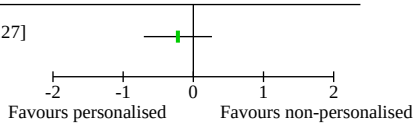
Analysis 1.54. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 54: Parenting skills and knowledge: FOS, Negative Parent Behaviour (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	0.56	0.9	13	0.21	0.36	19	0.35 [-0.17 , 0.87]



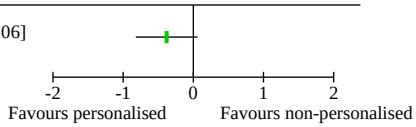
Analysis 1.55. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 55: Parenting skills and knowledge: Parenting Scale (PS) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	2.58	0.72	15	2.8	0.75	21	-0.22 [-0.71 , 0.27]



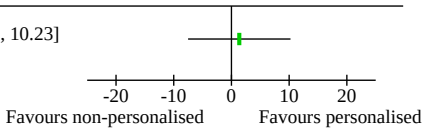
Analysis 1.56. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 56: Parenting skills and knowledge: PS (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	2.56	0.59	13	2.94	0.67	19	-0.38 [-0.82 , 0.06]



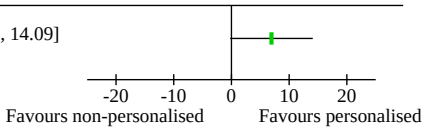
Analysis 1.57. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 57: Parenting skills and knowledge: Parenting Sense of Competency (PSOC) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	65.31	12.96	15	63.95	13.97	21	1.36 [-7.51 , 10.23]



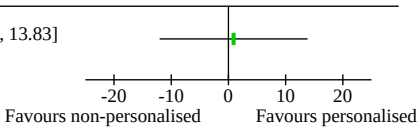
Analysis 1.58. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 58: Parenting skills and knowledge: PSOC (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	68.18	8.42	13	61.24	12.22	19	6.94 [-0.21 , 14.09]



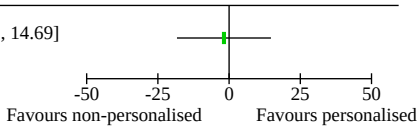
Analysis 1.59. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 59: Parenting skills and knowledge: FOS, Parental Correct Implementation (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Dadds 1992	71.8	16.7	10	70.9	13.1	11	0.90 [-12.03 , 13.83]



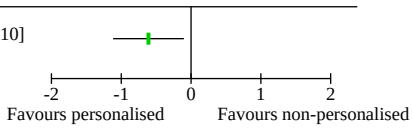
Analysis 1.60. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 60: Parenting skills and knowledge: FOS, Parental Correct Implementation (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Dadds 1992	67.2	24.5	10	69	10.9	11	-1.80 [-18.29 , 14.69]



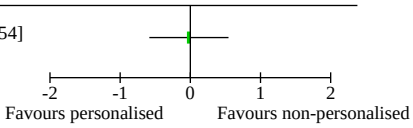
Analysis 1.61. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 61: Parenting skills and knowledge: PS, Laxness (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	1.94	0.71	13	2.55	0.65	15	-0.61 [-1.12 , -0.10]



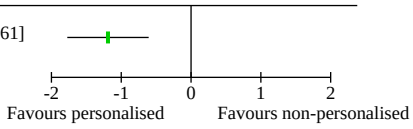
Analysis 1.62. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 62: Parenting skills and knowledge: PS, Laxness (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.45	0.88	13	2.47	0.57	14	-0.02 [-0.58 , 0.54]



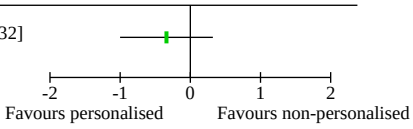
Analysis 1.63. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 63: Parenting skills and knowledge: PS, Over-reactivity (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.1	0.78	13	3.29	0.79	15	-1.19 [-1.77 , -0.61]



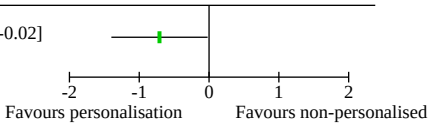
Analysis 1.64. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 64: Parenting skills and knowledge: PS, Over-reactivity (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.45	0.89	13	2.79	0.86	14	-0.34 [-1.00 , 0.32]



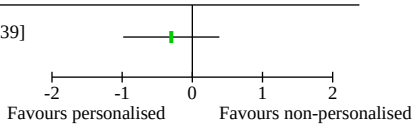
Analysis 1.65. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 65: Parenting skills and knowledge: PS, Verbosity (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.58	0.89	13	3.29	0.97	15	-0.71 [-1.40 , -0.02]



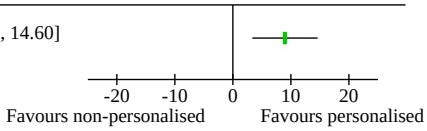
Analysis 1.66. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 66: Parenting skills and knowledge: PS, Verbosity (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.82	1.12	13	3.12	0.6	14	-0.30 [-0.99 , 0.39]



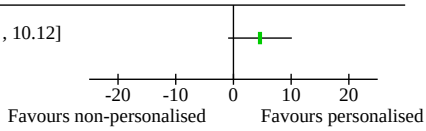
Analysis 1.67. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 67: Parenting skills and knowledge: PSOC, Satisfaction (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	43.62	6.61	13	34.64	8.54	15	8.98 [3.36 , 14.60]



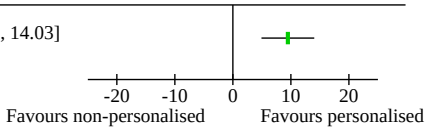
Analysis 1.68. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 68: Parenting skills and knowledge: PSOC, Satisfaction (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	41.46	6.94	13	36.86	7.69	14	4.60 [-0.92 , 10.12]



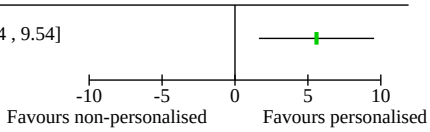
Analysis 1.69. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 69: Parenting skills and knowledge: PSOC, Efficacy (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	35.62	5.44	13	26.14	6.83	15	9.48 [4.93 , 14.03]



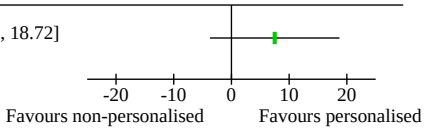
Analysis 1.70. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 70: Parenting skills and knowledge: PSOC, Efficacy (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	33.23	3.68	13	27.64	6.5	14	5.59 [1.64 , 9.54]	



Analysis 1.71. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 71: Parenting skills and knowledge: Parenting Practices Scale (PPS) (pre–post)

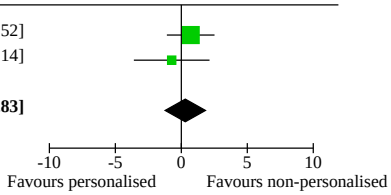
Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
McCabe 2009	156	15.84	21	148.5	19.9	19	7.50 [-3.72 , 18.72]	



Analysis 1.72. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 72: Family functioning: Parent Problem Checklist (PPC), Problem (short term)

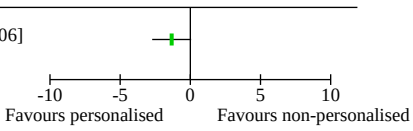
Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	3.56	3.05	15	2.85	2.19	21	71.6%	0.71 [-1.10 , 2.52]	
Markie-Dadds 2006	4	3.92	13	4.73	3.79	15	28.4%	-0.73 [-3.60 , 2.14]	
Total (95% CI)			**28**			**36**	**100.0%**	**0.30 [-1.23 , 1.83]**	

Heterogeneity: Chi² = 0.69, df = 1 (P = 0.40); I² = 0%
Test for overall effect: Z = 0.39 (P = 0.70)
Test for subgroup differences: Not applicable



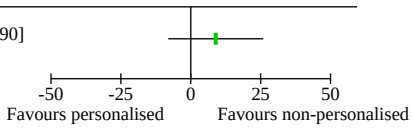
Analysis 1.73. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 73: Family functioning: PPC, Problem (long term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	2	1.55	13	3.33	2.45	19	-1.33 [-2.72 , 0.06]	



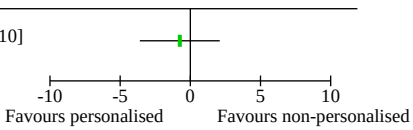
Analysis 1.74. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 74: Family functioning: PPC, Intensity (short term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	38.46	28.63	13	29.53	13.39	15	8.93 [-8.04 , 25.90]	



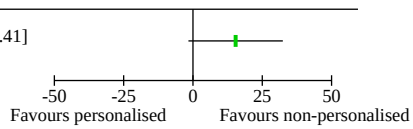
Analysis 1.75. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 75: Family functioning: PPC, Problem (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	3.82	3.25	13	4.57	4.26	14	-0.75 [-3.60 , 2.10]	



Analysis 1.76. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 76: Family functioning: PPC, Intensity (medium term)

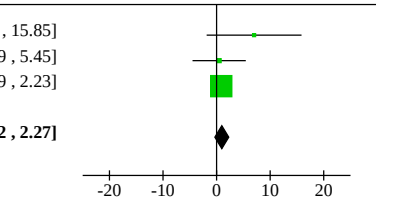
Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	43.46	28.72	13	28.07	12.95	14	15.39 [-1.63 , 32.41]	



Analysis 1.77. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 77: Engagement and decreased dropout: session attendance (%)

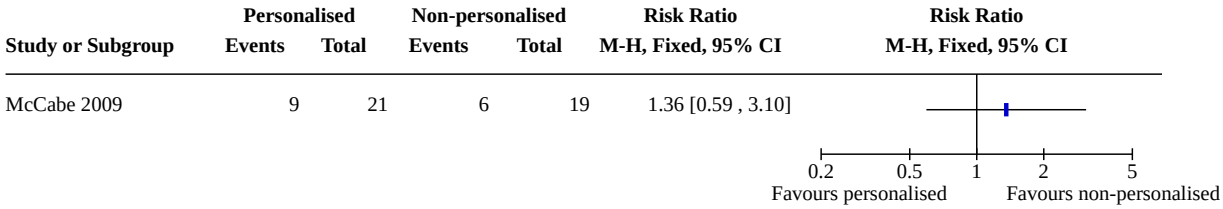
Study or Subgroup	Personalised Mean	SD	Total	Non-personalised Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
Jones 2014	97	5	8	90	11	7	2.1%	7.00 [-1.85 , 15.85]	
McCabe 2009	13.9	7.99	21	13.42	8.03	19	6.8%	0.48 [-4.49 , 5.45]	
Parent 2022	7.56	3.72	47	6.69	3.16	54	91.1%	0.87 [-0.49 , 2.23]	
Total (95% CI)			**76**			**80**	**100.0%**	**0.97 [-0.32 , 2.27]**	

Heterogeneity: Tau² = 0.00; Chi² = 1.84, df = 2 (P = 0.40); I² = 0%
Test for overall effect: Z = 1.48 (P = 0.14)
Test for subgroup differences: Not applicable



Analysis 1.78. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 78: Engagement and decreased dropout: mid-week call availability (%)

Analysis 1.79. Comparison 1: Personalised intervention versus non-personalised intervention, Outcome 79: Engagement and decreased dropout: dropout rates



Comparison 2. Personalised intervention versus waitlist control or treatment as usual (TAU)

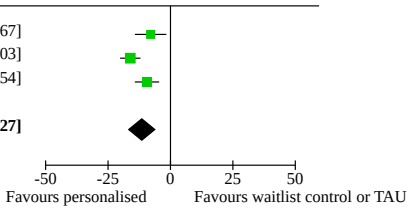
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Improvement in child conduct problems: Eyberg Child Behavior Inventory (ECBI), Problem (short term)	3	119	Mean Difference (IV, Random, 95% CI)	-11.50 [-16.73, -6.27]
2.2 Improvements in child conduct problems: ECBI, Intensity (short term)	3	119	Mean Difference (IV, Random, 95% CI)	-28.44 [-50.10, -6.78]
2.3 Improvement in child conduct problems: Parent Daily Report (PDR), Mean daily (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.4 Improvement in child conduct problems: PDR, Mean target (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.5 Improvement in child conduct problems: ECBI, Oppositional Defiant (OD) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.6 Improvement in child conduct problems: ECBI, Conduct Disorder (CD) (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.7 Improvements in child conduct problems: PDR, Problem (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.8 Improvements in child conduct problems: PDR, Target (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.9 Improvement in child conduct problems: ECBI, Problem (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.10 Improvement in child conduct problems: ECBI, Intensity (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.11 Improvement in child conduct problems: Child Behaviour Checklist (CBCL), Externalizing (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.12 Improvement in child conduct problems: CBCL, Externalising (medium term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.13 Personalised treatment outcomes: ECBI, Inattention (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.14 Parenting skills and knowledge: FOS, Negative parent behaviour (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.15 Parenting skills and knowledge: Parenting Scale (PS), Total (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.16 Parenting skills and knowledge: Parenting Sense of Competency (PSOC), Total (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.17 Parenting skills and knowledge: PS, Laxness (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.18 Parenting skills and knowledge: PS, Over-reactivity (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.19 Parenting skills and knowledge: PS, Verbosity (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.20 Parenting skills and knowledge: PSOC, Satisfaction (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.21 Parenting skills and knowledge: PSOC, Efficacy (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.22 Family functioning: Parent Problem Checklist (PPC), Problem (short term)	2	67	Mean Difference (IV, Random, 95% CI)	-1.82 [-3.53, -0.12]
2.23 Family functioning: PPC, Intensity (short term)	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected

Analysis 2.1. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 1: Improvement in child conduct problems: Eyberg Child Behavior Inventory (ECBI), Problem (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Bor 2002	10.46	10.88	15	18.42	8.07	27	28.5%	-7.96 [-14.25 , -1.67]
Markie-Dadds 2006	2.15	3.36	13	18.25	6.43	12	37.4%	-16.10 [-20.17 , -12.03]
Mersky 2016	10.8	8.587	19	20.2	8.6743	33	34.1%	-9.40 [-14.26 , -4.54]
Total (95% CI)			**47**			**72**	**100.0%**	**-11.50 [-16.73 , -6.27]**

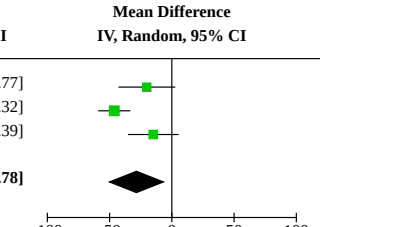
Heterogeneity: Tau² = 14.70; Chi² = 6.54, df = 2 (P = 0.04); I² = 69%
Test for overall effect: Z = 4.31 (P < 0.0001)
Test for subgroup differences: Not applicable



Analysis 2.2. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 2: Improvements in child conduct problems: ECBI, Intensity (short term)

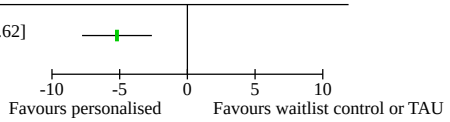
Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Bor 2002	132.07	38.31	15	152.12	31.91	27	29.7%	-20.05 [-42.87 , 2.77]
Markie-Dadds 2006	100.69	17.41	13	146.92	15.53	12	38.3%	-46.23 [-59.14 , -33.32]
Mersky 2016	133.8	36.3096	19	148.7	35.3291	33	31.9%	-14.90 [-35.19 , 5.39]
Total (95% CI)			**47**			**72**	**100.0%**	**-28.44 [-50.10 , -6.78]**

Heterogeneity: Tau² = 275.25; Chi² = 8.28, df = 2 (P = 0.02); I² = 76%
Test for overall effect: Z = 2.57 (P = 0.01)
Test for subgroup differences: Not applicable



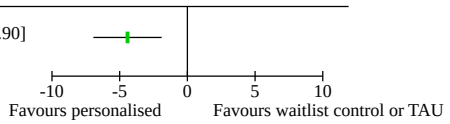
Analysis 2.3. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 3: Improvement in child conduct problems: Parent Daily Report (PDR), Mean daily (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	5.03	2.61	15	10.23	5.87	27	-5.20 [-7.78 , -2.62]



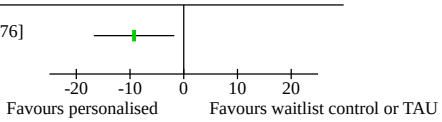
Analysis 2.4. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 4: Improvement in child conduct problems: PDR, Mean target (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Bor 2002	3.54	2.99	15	7.96	5.35	27	-4.42 [-6.94 , -1.90]



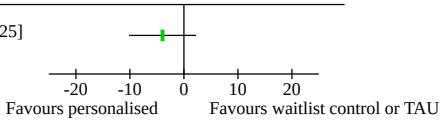
Analysis 2.5. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 5: Improvement in child conduct problems: ECBI, Oppositional Defiant (OD) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	38.27	12.41	15	47.52	10.81	27	-9.25 [-16.74 , -1.76]	



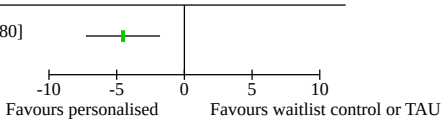
Analysis 2.6. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 6: Improvement in child conduct problems: ECBI, Conduct Disorder (CD) (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	21.47	9.52	15	25.43	10.38	27	-3.96 [-10.17 , 2.25]	



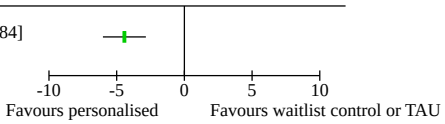
Analysis 2.7. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 7: Improvements in child conduct problems: PDR, Problem (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	3.76	2.76	13	8.29	4.04	12	-4.53 [-7.26 , -1.80]	



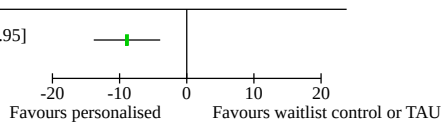
Analysis 2.8. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 8: Improvements in child conduct problems: PDR, Target (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	0.5	0.56	13	4.93	2.75	12	-4.43 [-6.02 , -2.84]	



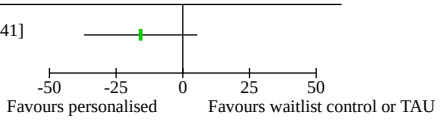
Analysis 2.9. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 9: Improvement in child conduct problems: ECBI, Problem (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Mersky 2016	5.4	8.8486	19	14.3	8.6168	33	-8.90 [-13.85 , -3.95]	



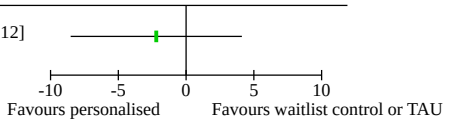
Analysis 2.10. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 10: Improvement in child conduct problems: ECBI, Intensity (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Mersky 2016	118.2	38.0968	19	134	36.6503	33	-15.80 [-37.01 , 5.41]	



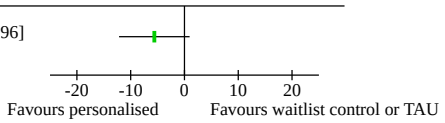
Analysis 2.11. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 11: Improvement in child conduct problems: Child Behaviour Checklist (CBCL), Externalizing (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Mersky 2016	22.3	11.3767	19	24.5	10.8572	33	-2.20 [-8.52 , 4.12]	



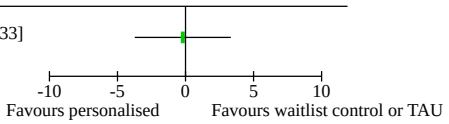
Analysis 2.12. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 12: Improvement in child conduct problems: CBCL, Externalising (medium term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Mersky 2016	17.2	11.9434	19	22.8	11.0296	33	-5.60 [-12.16 , 0.96]	



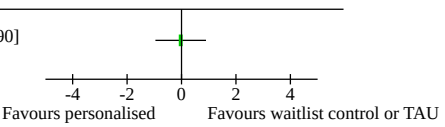
Analysis 2.13. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 13: Personalised treatment outcomes: ECBI, Inattention (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	18.13	5.58	15	18.33	5.62	27	-0.20 [-3.73 , 3.33]	



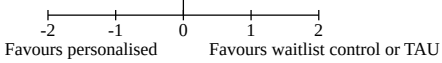
Analysis 2.14. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 14: Parenting skills and knowledge: FOS, Negative parent behaviour (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	0.53	1.04	15	0.56	2.03	27	-0.03 [-0.96 , 0.90]	



Analysis 2.15. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 15: Parenting skills and knowledge: Parenting Scale (PS), Total (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	2.58	0.72	15	3.36	0.71	27	-0.78 [-1.23 , -0.33]	

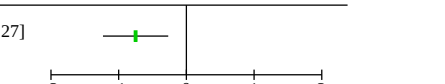


Analysis 2.16. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 16: Parenting skills and knowledge: Parenting Sense of Competency (PSOC), Total (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Bor 2002	65.31	12.96	15	53.24	12.6	27	12.07 [3.97 , 20.17]	

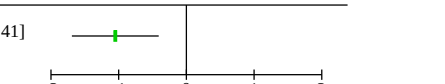
Analysis 2.17. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 17: Parenting skills and knowledge: PS, Laxness (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	1.94	0.71	13	2.69	0.51	12	-0.75 [-1.23 , -0.27]	



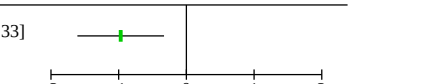
Analysis 2.18. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 18: Parenting skills and knowledge: PS, Over-reactivity (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.1	0.78	13	3.15	0.85	12	-1.05 [-1.69 , -0.41]	



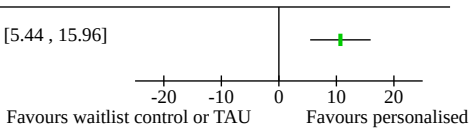
Analysis 2.19. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 19: Parenting skills and knowledge: PS, Verbosity (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	2.58	0.89	13	3.55	0.74	12	-0.97 [-1.61 , -0.33]	



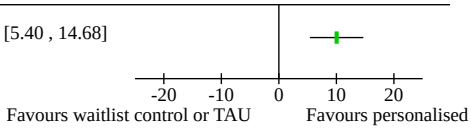
Analysis 2.20. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 20: Parenting skills and knowledge: PSOC, Satisfaction (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	43.62	6.61	13	32.92	6.78	12	10.70 [5.44 , 15.96]



Analysis 2.21. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 21: Parenting skills and knowledge: PSOC, Efficacy (short term)

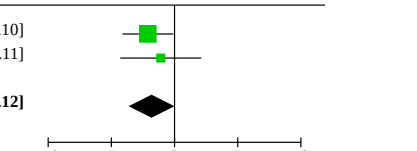
Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	35.62	5.44	13	25.58	6.32	12	10.04 [5.40 , 14.68]



Analysis 2.22. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 22: Family functioning: Parent Problem Checklist (PPC), Problem (short term)

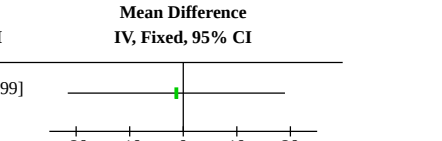
Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Bor 2002	3.56	3.05	15	5.67	3.43	27	71.7%	-2.11 [-4.12 , -0.10]
Markie-Dadds 2006	4	3.92	13	5.09	4.23	12	28.3%	-1.09 [-4.29 , 2.11]
Total (95% CI)			**28**			**39**	**100.0%**	**-1.82 [-3.53 , -0.12]**

Heterogeneity: Tau² = 0.00; Chi² = 0.28, df = 1 (P = 0.60); I² = 0%
Test for overall effect: Z = 2.09 (P = 0.04)
Test for subgroup differences: Not applicable



Analysis 2.23. Comparison 2: Personalised intervention versus waitlist control or treatment as usual (TAU), Outcome 23: Family functioning: PPC, Intensity (short term)

Study or Subgroup	Personalised Mean	SD	Total	Waitlist control Mean	SD	Total	Mean Difference IV, Fixed, 95% CI
Markie-Dadds 2006	38.46	28.63	13	39.75	22.97	12	-1.29 [-21.57 , 18.99]



APPENDICES

Appendix 1. Search strategies

Cochrane Central Register of Controlled Trials (CENTRAL) via Cochrane Register of Studies Online

Searched 17 May 2017 (3635 records)
Searched 26 March 2019 (1047 records)
Searched 25 June 2020 (546 records)
Searched 1 February 2022 (767 records)

#1MESH DESCRIPTOR Attention Deficit and Disruptive Behavior Disorders
#2MESH DESCRIPTOR Disruptive, Impulse Control, and Conduct Disorders
#3MESH DESCRIPTOR Conduct Disorder
#4((conduct* adj3 (difficult* or disorder* or disturb* or problem*))):TI,AB,KY
#5((difficult* or disorder* or disturb* or problem*) adj3 conduct):TI,AB,KY
#6((oppositional adj3 (defian* or disorder*))):TI,AB,KY
#7((disrupt* adj3 (behav* or disorder*))):TI,AB,KY
#8(((behav* or disorder*) adj3 disrupt*)):TI,AB,KY
#9((defian* adj3 (behav* or disorder*))):TI,AB,KY
#10(((behav* or disorder*) adj3 defian*)):TI,AB,KY
#11((impuls* adj3 (behav* or disorder*))):TI,AB,KY
#12(((behav* or disorder*) adj3 impuls*)):TI,AB,KY
#13((conduct adj3 (difficult* or disorder* or disturb* or problem*))):TI,AB,KY
#14((difficult* or disorder* or disturb* or problem*) adj3 conduct):TI,AB,KY
#15MESH DESCRIPTOR Aggression EXPLODE ALL TREES
#16((aggressiv* adj2 (behav* or disorder*))):TI,AB,KY
#17(((behav* or disorder*) adj2 aggressiv*)):TI,AB,KY
#18MESH DESCRIPTOR Child Behavior Disorders
#19MESH DESCRIPTOR Social Behavior Disorders
#20((behav* adj2 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*))):TI,AB,KY
#21(((agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*) adj2 behav*)):TI,AB,KY
#22#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21
#23((boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or pediatric* or paediatric* or preadolescen* or prepubert* or prepubescen* or preschool* or preteen* or pubert* or pubescen* or school* or toddler*)):TI,AB,KY
#24#22 AND #23 [NOTE:FINAL LINE 2017]
#2517/05/2017 TO 26/03/2019:CD
#26#24 AND #25 [NOTE:FINAL LINE 2019]
#27 26/03/2019 TO 25/06/2020:CD
#28#24 AND #27 [NOTE:FINAL LINE 2020]
#29 25/06/2020 TO 01/02/2022:CD
#30 #24 AND #29 [NOTE:FINAL LINE 2022]

MEDLINE Ovid

Searched 16 May 2017 (3222 records)
Searched 26 March 2019 (451 records)
Searched 25 June 2020 (305 records)
Searched 1 February 2022 (430 records)

1 "Attention Deficit and Disruptive Behavior Disorders"/
2 "disruptive, impulse control, and conduct disorders"/
3 conduct disorder/
4 (conduct$ adj3 (difficult$ or disorder$ or disturb$ or problem$)).tw,kf.
5 child behavior disorders/
6 Social Behavior Disorders/
7 (oppositional adj3 (defian$ or disorder$)).tw,kf.
8 (disrupt$ adj3 (behav$ or disorder$)).tw,kf.
9 (defian$ adj3 (behav$ or disorder$)).tw,kf.
10 (impuls$ adj3 (behav$ or disorder$)).tw,kf.
11 exp Aggression/
12 (aggressiv$ adj2 (behav$ or disorder$)).tw,kf.
13 (behav$ adj2 (agonistic or antisocial or anti-social or challeng$ or disorder$ or disturb$ or externali$ or problem$)).tw,kf.
14 or/1-13
15 Minors/
16 exp child/
17 (boy$1 or child$ or delinquen$ or girl$1 or graders or junior$1 or juvenile$1 or kindergarten or minors or p?ediatric$ or preadolescen$ or prepubert$ or prepubescen$ or preschool$ or preteen$ or pubert$ or pubescen$ or school$ or toddler$).tw,kf.
18 or/15-17
19 14 and 18
20 randomized controlled trial.pt.
21 controlled clinical trial.pt.
22 randomi#ed.ab.

23 placebo.ab.
24 clinical trials as topic.sh.
25 randomly.ab.
26 trial.ti.
27 or/20-26
28 exp animals/ not humans.sh.
29 27 not 28
30 19 and 29
31 remove duplicates from 30 [NOTE:FINAL LINE 2017]
32 limit 31 to ed=20170501-20190314 [NOTE:FINAL LINE 2019]
33 limit 31 to ed=20190314-20200611[NOTE:FINAL LINE 2020]
34 (202006* or 202007* or 202008* or 202009* or 202010* or 202011*or 202012* or 2021* or 2022*).dt,ez,da.
35 31 and 34 [NOTE:FINAL LINE 2022]

MEDLINE EPub Ahead of Print Ovid

Searched 16 May 2017 (142 records)
Searched 26 March 2019 (112 records after deduplication with previous records)
Searched 25 June 2020 (136 records after deduplication with previous records)
Searched 1 February 2022 (524 records after deduplication with previous records)

1 (conduct$ adj3 (difficult$ or disorder$ or disturb$ or problem$)).tw,kf.
2 (oppositional adj3 (defian$ or disorder$)).tw,kf.
3 (disrupt$ adj3 (behav$ or disorder$)).tw,kf.
4 (defian$ adj3 (behav$ or disorder$)).tw,kf.
5 (impuls$ adj3 (disorder$ or behav$)).tw,kf.
6 (aggressiv$ adj3 (disorder$ or behav$)).tw,kf.
7 (behav$ adj2 (agonistic or antisocial or anti-social or challeng$ or disorder$ or disturb$ or externali$ or problem$)).tw,kf.
8 or/1-7
9 (boy$1 or child$ or delinquen$ or girl$1 or graders or junior$1 or juvenile$1 or kindergarten or minors or p?ediatric$ or preadolescen$ or prepubert$ or prepubescen$ or preschool$ or preteen$ or pubert$ or pubescen$ or school$ or toddler$).tw,kf.
10 (random$ or placebo$ or trial).tw,kf.
11 8 and 9 and 10

MEDLINE In-Process and Other Non-Indexed Citations Ovid

Searched 16 May 2017 (365 records)
Searched 26 March 2019 (436 records after deduplication with previous records)
Searched 25 June 2020 (496 records after deduplication with previous records)
Searched 1 February 2022 (137 records after deduplication with previous records)

1 (conduct$ adj3 (difficult$ or disorder$ or disturb$ or problem$)).tw,kf.
2 (oppositional adj3 (defian$ or disorder$)).tw,kf.
3 (disrupt$ adj3 (behav$ or disorder$)).tw,kf.
4 (defian$ adj3 (behav$ or disorder$)).tw,kf.
5 (impuls$ adj3 (disorder$ or behav$)).tw,kf.
6 (aggressiv$ adj3 (disorder$ or behav$)).tw,kf.
7 (behav$ adj2 (agonistic or antisocial or anti-social or challeng$ or disorder$ or disturb$ or externali$ or problem$)).tw,kf.
8 or/1-7
9 (boy$1 or child$ or delinquen$ or girl$1 or graders or junior$1 or juvenile$1 or kindergarten or minors or p?ediatric$ or preadolescen$ or prepubert$ or prepubescen$ or preschool$ or preteen$ or pubert$ or pubescen$ or school$ or toddler$).tw,kf.
10 (random$ or placebo$ or trial).tw,kf.
11 8 and 9 and 10

Embase Ovid

Searched 16 May 2017 (3679 records)
Searched 26 March 2019 (598 records)
Searched 25 June 2020 (445 records)
Searched 1 February 2022 (650 records)

1 Conduct Disorder/
2 exp disruptive behavior/

3 impulse control disorder/
4 intermittent explosive disorder/
5 oppositional defiant disorder/
6 aggression/
7 (conduct$ adj3 (difficult$ or disorder$ or disturb$ or problem$)).tw,kw.
8 (disrupt$ adj3 (behav$ or disorder$)).tw,kw.
9 (defian$ adj3 (behav$ or disorder$)).tw,kw.
10 (oppositional adj3 (defian$ or disorder$)).tw,kw.
11 (impuls$ adj3 (behav$ or disorder$)).tw,kw.
12 ((aggressive$ or aggression) adj1 (behav$ or disorder$)).tw,kw.
13 (behav$ adj1 (agonistic or antisocial or anti-social or challeng$ or disorder$ or disturb$ or externali$ or problem$)).tw,kw.
14 or/1-13
15 "minor (person)"/
16 exp child/
17 (boy$1 or child$ or delinquen$ or girl$1 or graders or junior$1 or juvenile$1 or kindergarten or minors or p?ediatric$ or preadolescen$ or prepubert$ or prepubescen$ or preschool$ or preteen$ or pubert$ or pubescen$ or school$ or toddler$).tw,kw.
18 or/15-17
19 14 and 18
20 Randomized controlled trial/
21 controlled clinical trial/
22 randomi#ed.ab.
23 randomly.ab.
24 placebo.ab.
25 trial.ti.
26 or/20-25
27 19 and 26
28 exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/
29 human/ or normal human/ or human cell/
30 28 and 29
31 28 not 30
32 27 not 31
33 remove duplicates from 32 [NOTE:FINAL LINE 2017]
34 limit 33 to yr="2017 -Current" [NOTE:FINAL LINE 2019]
35 limit 33 to yr="2019-Current" [NOTE:FINAL LINE 2020]
36 limit 33 to yr="2020 -Current" [NOTE:FINAL LINE 2022]

APA PsycINFO Ovid

Searched 17 May 2017 (4957 records)
Searched 26 March 2019 (2498 records)
Searched 25 June 2020 (379 records)
Searched 1 February 2022 (650 records)

1 conduct disorder/
2 explosive disorder/
3 oppositional defiant disorder/
4 impulse control disorders/
5 aggressive behavior/
6 acting out/
7 externalization/
8 (conduct$ adj3 (difficult$ or disorder$ or disturb$ or problem$)).tw.
9 (oppositional adj3 (defian$ or disorder$)).tw.
10 (disrupt$ adj3 (behav$ or disorder$)).tw.
11 (defian$ adj3 (behav$ or disorder$)).tw.
12 (impuls$ adj3 (behav$ or disorder$)).tw.
13 (aggressiv$ adj2 (behav$ or disorder$)).tw.
14 exp behavior problems/
15 (behav$ adj2 (agonistic or antisocial or anti-social or challeng$ or disorder$ or disturb$ or externali$ or problem$)).tw.
16 or/1-15
17 ("100" or "160" or "180").ag.
18 (boy$1 or child$ or delinquen$ or girl$1 or graders or junior$1 or juvenile$1 or kindergarten or minors or p?ediatric$ or preadolescen$ or prepubert$ or prepubescen$ or preschool$ or preteen$ or pubert$ or pubescen$ or school$ or toddler$).tw.
19 17 or 18

20 16 and 19
21 clinical trials/
22 treatment effectiveness evaluation/
23 exp program evaluation/
24 psychotherapeutic outcomes/
25 followup studies/
26 longitudinal studies/
27 random$.tw.
28 ((clinical or control$) adj1 trial).tw.
29 ((control$ or experiment$ or treat$ or intervention) adj1 group$).tw.
30 or/21-29)
31 20 and 30
32 limit 31 to human [NOTE:FINAL LINE 2017]
33 limit 32 to up=20170501-20190304 [NOTE:FINAL LINE 2019]
34 limit 32 to up=20190304-20200622 [NOTE:FINAL LINE 2020]
35 limit 32 to up=20200622-20220124 [NOTE:FINAL LINE 2022]

CINAHLPlus EBSCOhost

Searched 17 May 2017 (3831 records)
Searched 26 March 2019 (810 records after deduplicating with previous records)
Searched 25 June 2020 (618 records)
Searched 3 February 2022 (711 records)

S1(MH "Child Behavior Disorders")
S2(MH "Disruptive Behavior")
S3(MH "Social Behavior Disorders")
S4(MH "Aggression")
S5 conduct* N3 (difficult* or disorder* or disturb* or problem*))
S6(oppositional N3 (defian* or disorder*))
S7(disrupt* N3 (behav* or disorder*))
S8(defian* N3 (behav* or disorder*))
S9(impuls* N3 (behav* or disorder*))
S10(aggressiv* N3 (behav* or disorder*))
S11(behav* N2 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*))
S12S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11
S13(MH "Child")
S14(MH "Child, Preschool")
S15(MH "Minors (Legal)")
S16boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or pediatric* or paediatric* or preadolescen* or prepubert* or prepubescen* or preschool* or
preteen* or pubert* or pubescen* or school* or toddler*
S17S13 OR S14 OR S15 OR S16
S18S12 AND S17
S19(MH "Clinical Trials+")
S20MH random assignment
S21PT randomized controlled trial
S22PT Clinical trial
S23((clinical or control*) N1 trial*)
S24(MH "Program Evaluation")
S25(MH "Treatment Outcomes")
S26(MH "Evaluation Research+")
S27((evaluat* or followup or follow-up) N1 study)
S28(random* OR RCT)
S29((control* or experiment* or treat* or intervention) N1 group*)
S30 S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29
S31 S18 AND S30 [NOTE:FINAL LINE 2017]
S32 EM 201705-
S33 S31 AND S32[NOTE:FINAL LINE 2019]
S34 EM 201903-
S35 S31 AND S34[NOTE:FINAL LINE 2020]
S36 EM 202006-
S37 S31 AND S36[NOTE:FINAL LINE 2022]

ERIC EBSCOhost

Searched 17 May 2017 (5091 records)
Searched 26 March 2019 (428 records)
Searched 25 June 2020 (235 records)
Searched 3 February 2022 (221 records)

S1DE "Behavior Problems"
S2DE "Behavior Disorders"
S3(DE "Aggression")
S4TI (conduct* N3 (difficult* or disorder* or disturb* or problem*)) OR AB(conduct* N3 (difficult* or disorder* or disturb* or problem*))
S5TI(oppositional N3 (defian* or disorder*)) OR AB(oppositional N3 (defian* or disorder*))
S6TI(disrupt* N3 (behav* or disorder*)) OR AB(disrupt* N3 (behav* or disorder*))
S7TI(defian* N3 (behav* or disorder*)) OR AB(defian* N3 (behav* or disorder*))
S8TI (impuls* N3 (behav* or disorder*)) OR AB(impuls* N3 (behav* or disorder*))
S9TI(aggressiv* N3 (behav* or disorder*)) OR AB(aggressiv* N3 (behav* or disorder*))
S10TI(behav* N2 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*)) OR AB(behav* N2 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*))
S11S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10
S12DE "Children" OR DE "Preschool Children" OR DE "Young Children" OR DE "Toddlers" OR DE "Preadolescents"
S13TI(boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or pediatric* or paediatric* or preadolescen* or prepubert* or prepubescen* or preschool* or preteen* or pubert* or pubescen* or school* or toddler*) OR AB(boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or pediatric* or paediatric* or preadolescen* or prepubert* or prepubescen* or preschool* or preteen* or pubert* or pubescen* or school* or toddler*)
S14S12 OR S13
S15S11 AND S14
S16DE "Evaluation Research" OR DE "Control Groups" OR DE "Experimental Groups" OR DE "Longitudinal Studies" OR DE "Followup Studies" OR DE "Program Effectiveness" OR DE "Program Evaluation"
S17TI (random* N1 (allocat* or assign*)) or AB(random* N1 (allocat* or assign*)) OR RCT
S18TI((clinical or control*) N1 trial*) OR AB((clinical or control*) N1 trial*)
S19TI((control* or experiment* or treat* or intervention) N1 group*) OR AB((control* or experiment* or treat* or intervention) N1 group*)
S20TI((evaluat* or followup or follow-up) N1 study) OR AB((evaluat* or followup or follow-up) N1 study)
S21S16 OR S17 OR S18 OR S19 OR S20
S22S15 AND S21 [NOTE:FINAL LINE 2017]
S23S15 AND S21 Limiters - Date Published: 20170101-20191231 [NOTE:FINAL LINE 2019]
S24S15 AND S21 Limiters - Date Published: 20190101- 20201231[NOTE:FINAL LINE 2020]
S25S15 AND S21 Limiters - Date Published: 20200101- 20221231[NOTE:FINAL LINE 2022]

Conference Proceedings Citation Index – Science (CPCI-S) and Conference Proceedings Citation Index – Social Science & Humanities Web of Science (CPCI-SS&H) Web of Science

Searched together 17 May 2017 (285 records)
Searched together 26 March 2019 (13 records)
Searched 25 June 2020 (7 records)
Searched 3 February 2022 (7 records)

# 18 #15 AND #10 [NOTE:FINAL LINE 2022]
Indexes=CPCI-S, CPCI-SSH Timespan=2020-2022
# 17 #15 AND #10 [NOTE:FINAL LINE 2020]
Indexes=CPCI-S, CPCI-SSH Timespan=2019-2020
# 16 #15 AND #10 [NOTE:FINAL LINE 2019]
Indexes=CPCI-S, CPCI-SSH Timespan=2017-2019
# 16 #15 AND #10 [NOTE:FINAL LINE 2017]
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 15 #14 OR #13 OR #12 OR #11
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 14 TS=((evaluat* or followup or follow-up) Near/1 study)
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 13 TS=((control* or experiment* or treat* or intervention) Near/1 group*)
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 12 TS= ((clinical or control*) near/1 trial*)
Indexes=CPCI-S, CPCI-SSH Timespan=All years

# 11 TS= (RANDOM*)
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# #9 AND #8
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 9 TS=(boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or pediatric* or paediatric* or preadolescen* or prepubert* or prepubescen* or preschool* or preteen* or pubert* or pubescen* or school* or toddler*)
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 8 #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 7 TS=(behav* NEAR/1 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 6 TS= (aggressiv* NEAR/1 (behav* or disorder*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 5 TS=(impuls* NEAR/1 (behav* or disorder*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# TS= (defian* NEAR/1(behav* or disorder*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years
#3 TS=(disrupt* NEAR/1 (behav* or disorder*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 2 TS=(oppositional NEAR/1 (defian* or disorder*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years
# 1 TS=(conduct NEAR/1 (difficult* or disorder* or disturb* or problem*))
Indexes=CPCI-S, CPCI-SSH Timespan=All years

Cochrane Database of Systematic Reviews (CDSR), part of the Cochrane Library

Searched 17 May 2017 (12 records)
Searched 27 March 2019 (2 records)
Searched 25 June 2020 (5 records)
Searched 3 February 2022 (7 records)

#1[mh ^"conduct disorder"]
#2[mh ^"Attention Deficit and Disruptive Behavior Disorders"]
#3[mh "disruptive, impulse control, and conduct disorders"]
#4[mh ^"child behavior disorders"]
#5(oppositional near/1 (defian* or disorder*)):ti
#6(disrupt* near/1 (behav* or disorder*)):ti
#7(defian* near/1 (behav* or disorder*)):ti
#8(impuls* near/1 (behav* or disorder*)):ti
#9(aggressiv* near/1 (behav* or disorder*)):ti
#10(behav* near/1 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*)):ti
#11(conduct* near/1 (difficult* or disorder* or disturb* or problem*)):ti
#12{or #1-#11}
#13(boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or preadolescen* or prepubert* or prepubescen* or preschool* or preteen* or pubert* or pubescen* or school* or toddler*):ti
#14#12 and #13 in Cochrane Reviews (Reviews and Protocols) [NOTE:FINAL LINE 2017]
#15#12 and #13 with Cochrane Library publication date from May 2017 to Mar 2019, in Cochrane Reviews and Cochrane Protocols [NOTE:FINAL LINE 2019]
#16#12 and #13 with Cochrane Library publication date from Mar 2019 to Jun 2020, in Cochrane Reviews and Cochrane Protocols [NOTE:FINAL LINE 2020]
#17#12 and #13 with Cochrane Library publication date from Jun 2020 to Feb 2022, in Cochrane Reviews and Cochrane Protocols [NOTE:FINAL LINE 2022]

Database of Abstracts of Reviews of Effects (DARE), part of the Cochrane Library

Issue 2 of 4, April 2015 Searched 17 May 2017 (62 records). No new issues after this date.

#1[mh ^"conduct disorder"]
#2[mh ^"Attention Deficit and Disruptive Behavior Disorders"]
#3[mh "disruptive, impulse control, and conduct disorders"]
#4[mh ^"child behavior disorders"]
#5(oppositional near/1 (defian* or disorder*)):ti
#6(disrupt* near/1 (behav* or disorder*)):ti

#7(defian* near/1 (behav* or disorder*)):ti
#8(impuls* near/1 (behav* or disorder*)):ti
#9(aggressiv* near/1 (behav* or disorder*)):ti
#10(behav* near/1 (agonistic or antisocial or anti-social or challeng* or disorder* or disturb* or externali* or problem*))
#11(conduct* near/1 (difficult* or disorder* or disturb* or problem*)):ti
#12{or #1-#11}
#13(boy* or child* or delinquen* or girl* or graders or junior* or juvenile* or kindergarten or minors or preadolescen* or prepubert* or prepubescen* or preschool* or preteen* or pubert* or pubescen* or school* or toddler*)
#15#12 and #13 in Other Reviews

Epistemonikos (www.epistemonikos.org)

Searched 17 May 2017 (13 records)
Searched 27 March 2019 (1 record added to database 17 May 2017 to 27 March 2019)
Searched 25 June 2020 (15 records added to database 27 March 2019 to 26 June 2020)
Searched 3 February 2022 (14 records added to database 25 June 2020 to 3 February 2022)

Title:(aggressive or aggression OR conduct OR disrupt* OR opposition* OR externali* OR defian* OR impuls*) AND title:(disorder* OR behav*) AND title:(child*)

Limits applied| PUBLICATION TYPE: SYSTEMATIC REVIEW

ClinicalTrials.gov (clinicaltrials.gov)

Searched 18 May (341 records)
Searched 27 March 2019 (76 records first posted from 18 May 2017 to 27 March 2019)
Searched 25 June 2020 (73 records first posted from 27 March /2019 to 25 June 2020)
Searched 3 February 2022 (13 records first posted from 25 June 2020 to 3 February 2022)

Condition/disease: "conduct disorder" OR oppositional OR externalising OR disruptive OR aggression OR aggressive
Intervention/Treatment: behavior OR behaviour OR other OR device
Applied filters: Interventional Child

International Clinical Trials Registry Platform (ICTRP)

Searched 18 May (330 records after eliminating duplicates)
Searched 27 March 2019 (116 records after eliminating duplicates)
Unable to search on 25 June 2020 because of heavy demand due to COVID-19
Searched 3 February 2022 (88 records first registered from 27 March 2019 to 3 February 2022 after eliminating duplicates)

Search 1
CONDITION: conduct disorder OR oppositional OR externalising OR disruptive OR aggression OR aggressive
CLINICAL TRIALS IN CHILDREN IS SELECTED
RECRUITMENT STATUS IS ALL

Search 2
TITLE: conduct disorder OR oppositional OR externalising OR disruptive OR aggression OR aggressive
CLINICAL TRIALS IN CHILDREN IS SELECTED
RECRUITMENT STATUS IS ALL

Appendix 2. Unused methods from the published protocol

Methods section	Unused method, as planned in the protocol (Kennedy 2017)
Measures of treatment effect	**Continuous data** In studies that used different measures to assess the impact of the intervention on the same outcome, we will use the SMD with 95% CIs.
Unit of analysis issues	**Cluster-randomised trials** In the identification of cluster-randomised trials, we will take into account the level at which randomisation occurred, to determine whether individuals were randomised individually or in groups.

(Continued)

	We will analyse cluster-randomised trials using the mean cluster size and an estimate of the ICC to adjust sample sizes to the 'effective sample size'. This process will follow the methods described in the *Cochrane Handbook of Systematic Reviews of Interventions* (Higgins 2021). Where an estimate in the ICC is not available, we will contact trial authors to obtain this information. Where that is unavailable, we will use an estimate from a similar trial or a trial with a similar population. We will combine single RCTs with cluster-RCTs if we consider the designs and interventions as sufficiently similar and the effect of the intervention is unlikely to be influenced by the method of randomisation. We will conduct sensitivity analyses if RCTs have not statistically accounted for clustering.
Assessment of reporting bias	Where an individual meta-analysis contains ≥ 10 studies, we will construct and visually assess funnel plots for skewness of data. A relationship between effect size and standard error could be due to publication or related biases, or differences between small and large studies. We will assess funnel plot asymmetry using Egger's test (Egger 1997).
Subgroup analysis and investigation of heterogeneity	We may conduct subgroup analyses to further investigate causes of heterogeneity. Possible subgroups include: 1. age of participating children across groups (aged < 5 years; 4–8 years; 9–12 years); 2. gender (boys versus girls); and 3. initial severity of conduct problems.
Sensitivity analysis	We will perform sensitivity analyses to evaluate the impact of study quality on the robustness of the conclusions drawn. Specifically, we will explore potential heterogeneity between studies according to: 1. missing data (we will conduct a sensitivity analysis when we cannot assume that data are missing at random, attrition is > 20%, or where an appropriate ITT analysis has not been undertaken); and 2. cluster effects (we will conduct a sensitivity analysis if cluster-RCTs have not been adjusted for clustering). If necessary, we will conduct additional analyses for any further potential issues identified that may impact the robustness of review findings.

CI: confidence interval; ICC: intraclass correlation coefficient; ITT: intention to treat; RCT: randomised controlled trial; SMD: standardised mean difference.

HISTORY

Protocol first published: Issue 8, 2017

CONTRIBUTIONS OF AUTHORS

EK, EH, LF, SS, RS, and LK were involved in the conception, design, and co-ordination of the review.

EH drafted the review.

EH, LF, LK, LJ, CL, GS, and VR screened records to determine eligibility and selected studies for inclusion in the review, with support from EK throughout the screening process.

LK, GS, and EH extracted data.

EH, GS, CL, and LF assessed the risk of bias with support from EK who resolved any disagreements.

EH, CL, GS, and LF rated the certainty of the evidence using GRADE.

EH, CL, and GS analysed the results.

GS, EH, LF, LK, CL, SS, and EK interpreted the data.

EK is the guarantor for the review and provided feedback on drafts of the review.

SS and RS provided feedback on drafts of the review.

CR provided statistical support.

DECLARATIONS OF INTEREST

SS, RS, and EK are lead/co-applicants on a National Institute for Health Research (NIHR) grant, the long-term objectives of which are to investigate outcomes of parent training and develop personalised interventions for children with conduct problems.

CL: none.

EH: none.

LK: none.

LF: none.

VR: none.

LJ: none.

CR: none.

SS has declared that he is Chief Investigator on a grant relevant to this review funded by the National Institute for Health Research, UK. No studies from this grant were included in our review. He declares no conflicts of interest.

RS: none.

GS: none.

EK: none.

SOURCES OF SUPPORT

Internal sources

- The Tavistock and Portman NHS Foundation Trust, UK

  Provided base and employment for all review authors except Christopher Roberts and Stephen Scott.

  The views and opinions do not necessarily reflect those of the NHS or Department of Health. The authors alone are responsible for the views expressed in this publication.
- Division of Population Health, University of Manchester, Manchester, UK, UK

  Provided a base and employment for Christopher Roberts.

  The views and opinions do not necessarily reflect those of the University of Manchester. The author alone are responsible for the views expressed in this publication.
- Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK

  Provided a base and employment for Stephen Scott.

  The views and opinions do not necessarily reflect those of King's College London. The author alone are responsible for the views expressed in this publication.

External sources

- National Institute for Health Research (NIHR), UK

  This review is being completed as part of a NIHR Programme Grant for Applied Research (PGfAR). Grant number: LTC-RP-PG-0814-20001.

  The views and opinions do not necessarily reflect those of the NIHR. The authors alone are responsible for the views expressed in this publication.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. Types of outcome measures

a. We did not define short-term, medium-term, and long-term effects. We decided to define short-term effects as one-month postintervention or less, medium-term as one month to less than 12 months, and long-term as 12 months or greater, in accordance with the definition in previous Cochrane Reviews and on consultation with the review group.

2. Search methods for identification of studies
   a. We searched CENTRAL via the Cochrane Register of Studies Online (CRSO), rather than via the Cochrane library.
   b. We broadened our search for trials by searching the International Clinical Trials Registry Platform (ICTRP), which is populated by trials from several primary registries, including the contents of the ISRCTN registry listed in the protocol. We were unable to access ICTRP in June 2020 as it was not available outside the World Health Organization because of heavy web traffic generated by the COVID-19 pandemic.
3. Selection of studies
   a. We had planned for LF and EK to independently assess the full-text reports for inclusion against the selection criteria. When conducting the review, LF, EH, CL, GS and VR assessed the full-text reports for inclusion against the selection criteria (Criteria for considering studies for this review), with EK reviewing when any uncertainties arose.
4. Data collection and analysis
   a. We were unable to use the following methods in the review, which we have archived for use in future updates.
      i. We planned to include cluster-randomised trials (see Kennedy 2017); however, we did not identify any eligible cluster-randomised trials.
      ii. We planned to assess reporting bias (see Kennedy 2017); however, due to the small number of trials in each meta-analysis, we were unable to assess reporting bias.
      iii. We planned to conduct subgroup analysis and investigation of heterogeneity (see Kennedy 2017), but due to the small number of trials that could be included in each meta-analysis, this was not possible.
      iv. We were unable to perform the planned sensitivity analysis (see Kennedy 2017), due to the small number of trials in each meta-analysis.
5. Assessment of risk of bias in included studies
   a. We amended the 'other' risk of bias domain to include group differences, small sample size, and to implement objective criteria for defining a lack of adherence to treatment manual, in order to assist with replicability of assessing the risk of bias in this domain in future updates of the review. We considered a small sample size (fewer than 15 in the treatment group) to be at risk of bias. Group differences were defined as instances were differences between groups would directly impact upon the results of an intervention, or where group differences were not reported by a study. We defined a lack of adherence to treatment manual as adherence of 80% or less (Borrelli 2005).
6. Unit of analysis issues > Multiple treatment groups
   a. We added the following information to provide further details about the analysis of multiple intervention arms: "Where a study included multiple arms, we first assessed whether all arms met the inclusion criteria. Where two or more arms involved personalised interventions of interest and shared an eligible comparator (non-personalised intervention, TAU, or waitlist), we split the shared group into two or more groups with small sample sizes to avoid double counting the participants. These comparisons were separated into different forest plots, such that for three-arms trials with both a comparison intervention and no treatment control group, we made the following separate comparisons: personalised intervention compared to the comparison intervention and personalised intervention compared to the no treatment control. When combining groups in multiple-arm trials, we summed the sample sizes and events across groups for dichotomous data; and we combined sample sizes, means and SD for continuous data, in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021)."
7. Summary of findings and assessment of the certainty of the evidence
   a. We added details of the GRADE approach used by review authors to assess the certainty of the body of evidence and provide an overall rating for each outcome.
   b. We did not state the follow-up length that would be included in the summary of findings table in the protocol (Kennedy 2017). We chose to report short-term data because a diagnosis of a conduct disorder is strongly associated with poor educational performance, social isolation, and, in adolescence, substance misuse and increased contact with the criminal justice system. This association continues into adult life with poorer educational and occupational outcomes, involvement with the criminal justice system, and a high level of mental health problems (NICE 2017). With this in mind, a focus on preventing or reducing the escalation of existing conduct problems or disruptive behaviour as early as possible is key (NICE 2017).
   c. We only report mean differences in the summary of findings table, and not risk differences and risk ratios as described in the protocol. We made this decision because we considered that reporting the absolute difference between the mean value in our included interventions, rather than the estimated difference in the probability or the risk of an outcome increasing or decreasing, would better inform clinical decision-making.
   d. Given that the assessments were completed independently by two review authors, where discrepancies were noted, we resolved disagreements to reach consensus through discussion. This detail was not reported in the protocol (Kennedy 2017).

INDEX TERMS

Medical Subject Headings (MeSH)

Child Behavior; Child Rearing; Emotions; Parents [psychology]; *Problem Behavior; United States

MeSH check words

Adolescent; Child; Humans