Case Study 1: Evidence Based Practice Report

Theme: Interventions implemented by parents.

How Effective is Stepping Stones Triple P Programme for Reducing Behaviour Difficulties Among Children with Autism Spectrum Disorder?

Summary

This systematic literature review examined the effectiveness of a parenting intervention, Stepping Stones Triple P (SSTP), on reducing behavioural difficulties for children with Autism Spectrum Disorder (ASD). Through screening articles across three different databases (Web of Science, ERIC and PsycINFO), five studies were selected to be reviewed. Gough's (2007) Weight of Evidence Framework and Kratochwills' (2003) coding protocol were used to assess the quality of each study. All five studies found a significant reduction in behaviour problems for children with ASD, with small to large effect sizes, indicating the effectiveness of SSTP for reducing behaviour difficulties for this population. Two of the studies received high WoE D ratings, two received medium ratings and 1 received a low rating, highlighting variation in the quality of the studies reviewed. Limitations, future directions and implications are discussed.

Introduction

Stepping Stones Triple P

Triple P Positive Parenting Programme is a series of parenting and family interventions for parents of children who are at risk of developing emotional or behavioural problems (Sanders, 1999). SSTP is one of the Triple P Positive Parenting Programmes, designed specifically for children aged 0 to 12, with a disability. Standard Triple P strategies, such as specific praise and planned ignoring are included in SSTP, alongside additional strategies sourced from research on disabilities, such as social and communication

skills, independence skills, emotional self-regulation and problem solving skills (Early Intervention Foundation, 2023; Sanders et al., 2004). SSTP was developed because of the prevalence of children with a disability having behavioural difficulties. Study findings suggest that children with an intellectual disability are 2 to 3 times more likely to demonstrate behaviour problems (Sanders et al., 2004).

SSTP has five different levels of interventions; Level 1 Universal Triple P (general information to all parents interested in promoting their child's development); Level 2 Selected Triple P (two sessions of specific advice for parents with a particular concern about their child's development or behaviour); Level 3 Primary Care Triple P (four sessions for parents with specific concerns regarding their child's behaviour or development, involving parent skills training); Level 4 Standard Triple P (broad focus parenting skills training, over 8-10 sessions, usually for parents of children with more severe behaviour difficulties); Level 5 Enhanced Triple P (a behavioural family intervention, involving up to 11, 60-90 minute sessions, for parents of children with concurrent behaviour problems and family dysfunction) (Sanders, et al., 2004). There is a group SSTP training course, classified as a 'Level 4' programme, developed for parents who have concerns regarding their child's behaviour or who want to learn parenting skills to promote their child's development. This is a universal strategy for parents of children with a disability to prevent behavioural and emotional problems (Triple P, 2021a). The versions of SSTP that studies included in this review are based on are outlined in Table 1.

Table 1

Levels of SSTP Included in this Review

Level of Intervention	Intervention Methods	Sessions	Target Behaviours
Primary Care SSTP (Level 3)	4 sessions (15-30 minutes each) with families over a 4-6 week period Teaching parenting skills to apply to	**Session 1**: Assessment of presenting problem **Session 2:** Developing a parenting plan	Child behaviour problems e.g. tantrums, fighting with siblings; Developmental issues e.g.

	specific target behaviours, using active skills training, rehearsal and self-evaluation Face to face or telephone contact	**Session 3:** Review of Implementation **Session 4:** Follow up	independent self-care skills

Level of Intervention	Intervention Methods	Sessions	Target Behaviours
Standard SSTP (Level 4)	8-10 sessions (10 hours in total) Generalisation enhancement strategies and positive parenting skills. Parenting skills applied to a wide range of target behaviours Individual, group or self-directed options	**Session 1:** Initial interview **Session 2:** Observation and sharing of assessment findings **Session 3:** Promoting children's development **Session 4:** Managing misbehaviour **Session 5-7:** Practice sessions 1-3 **Session 8:** Planning ahead **Session 9:** Planning ahead practice session **Session 10:** Programme close	Multiple child behaviour difficulties; oppositional defiant disorder; learning difficulties; conduct disorder; aggressive behaviour; challenging behaviour; developmental issues e.g. sharing, toilet training, communication
Group SSTP (Level 4)	6 group sessions (2.5 hours per session), educating and actively training skills 3 individual (15-30 minute) telephone consultations helping parents use their parenting skills and develop problem solving skills	**Session 1:** Positive parenting **Session 2:** Promoting children's development **Session 3:** Teaching skills and behaviours **Session 4:** Managing misbehaviour and parenting routines	Mild to moderate challenging behaviours

	Session 5: Planning ahead **Session 6-8:** Implementing parenting routines **Session 9:** Program close

Note. This overview is based on information from Triple P (2021a); Triple P (2021b); Triple P (2021c) and Early Intervention Foundation (2023).

Psychological Basis

SSTP is a parenting intervention underpinned by social learning theory, coercion theory, behavioural family intervention and operant conditioning (Sanders, 1999).

Social Learning Theory

Social learning theory proposes that learning is a cognitive process, occurring in a social context, through observing, modelling and imitating behaviours and attitudes of others (Bandura et al., 1963, 1977). Observational learning involves observing individuals, formulating an idea of how new behaviours are performed and using this information as a guide for one's own actions. Vicarious reinforcement occurs through observation of rewards or punishments as a consequence of the behaviour (Bandura et al., 1963, 1977). SSTP is based on social learning principles, using techniques such as modelling, role play, direct feedback and home practice (Ruane & Carr, 2019).

Coercion Theory

Patterson's (1982) coercion theory, derived from behavioural research, outlines how aggressive and antisocial behaviours develop in children. Patterson (1982) proposes that ineffectual parental responses to a child's problem behaviour can unintentionally reinforce the undesired behaviour. These parent-child interactions can lead to a coercive cycle with problem behaviours increasing and escalating (Reid et al., 2002). Similar to Triple P, SSTP teaches positive parenting skills instead of coercive parenting practices (Sanders, 1999).

Operant Conditioning

Operant conditioning aims to reduce problem behaviours and encourage positive behaviours through positive reinforcement (Skinner, 1953). SSTP uses operant principles, using techniques, such as praise and planned ignoring, to reinforce positive behaviours and reduce problem behaviours (Kasperzack et al., 2020; Whittingham et al., 2009).

Behavioural Family Intervention

Through changing a child's environment which is reinforcing the child's problem behaviour, behavioural family interventions (BFI) aim to change the child's behaviour, using a therapeutic approach. SSTP is a BFI, aimed at reducing problem behaviours through altering parental behaviour and practices (Morawska & Sanders, 2006).

Public Health Perspective

Triple P adopts a public health approach to supporting families through increasing parenting support, ensuring parents participate to create meaningful change at a universal level rather than solely individual level. The approach ensures parents' rights and decision making are promoted, offering parents information and strategies to make more informed choices about parenting (Sanders, 2012).

Rationale for Review

A previous systematic review and meta-analysis evaluated the treatment effects of SSTP for parents of children with a disability. SSTP was found to be an effective intervention, improving outcomes for parents of children with a disability (Tellegen & Sanders, 2014). However, this review did not evaluate the specific effectiveness of SSTP for children with ASD on reducing behavioural difficulties. Furthermore, one of the authors, Professor Sanders, founded the SSTP programme, highlighting a conflict of interest.

The DSM-5 Autism Spectrum Disorder (ASD) diagnostic criteria, conceptualises ASD as persistent deficits in social communication and interaction and experiencing restricted, repetitive patterns of behaviour, interests or activities (American Psychiatric Association, 2013). ASD is a neurodevelopmental disorder and there is a spectrum of presentation, including less severe, more subtle behavioural difficulties to more severe behavioural difficulties. Challenging behaviours include oppositional defiant disorder, aggressiveness, temper tantrums, emotional lability and self-injurious behaviour (Yates & Le Couteur, 2016). Findings show that children with ASD display more maladaptive behaviours than peers with different intellectual disabilities and their typically developing peers (Hartley et al., 2008; McClintock et al., 2003). Similarly, a meta-analysis found a prevalence rate of 42% for self-injurious behaviour among children with ASD which is significantly higher than the prevalence rate among typically developing peers (Steenfeldt-Kristensen et al., 2020). Challenging behaviours are a concern to schools and families due to the negative effects these behaviours can have on the child and people around them (Nicholls et al., 2020). Challenging behaviours can increase social isolation and interfere with daily living skills, impacting a child's quality of life, alongside problem behaviours increasing parental mental health difficulties (Hartley et al., 2008; Weiss et al., 2012).

Therefore, with the prevalence of behavioural difficulties among children with ASD and the impact of these difficulties on the children and surrounding people. Educational psychologists, alongside other professionals, have a role in providing appropriate support for these families. Consequently, the purpose of this review was to examine the specific effectiveness of SSTP for children with ASD, with the review question outlined below:

"How Effective is Stepping Stones Triple P Programme for Reducing Behaviour Difficulties among Children with Autism Spectrum Disorder?"

Critical Review of the Evidence

Literature Search

A systematic literature review was conducted in December 2022, using the following databases: Web of Science, PsycINFO and ERIC (EBSCO). These databases were chosen to find articles relating to health, psychology and education. The same search terms were used across all three databases and are outlined in Table 2.

Table 2

Literature Search Terms

Search Term
"Stepping Stones Triple P" OR "SSTP" OR "Stepping Stones" OR "Triple P" AND "Autism Spectrum Disorder" OR ASD OR ASC OR "Autism Spectrum Condition" OR autis*

The initial search identified 114 studies. From this search, 22 duplicates were removed, leaving 92 studies. These studies were screened using the inclusion and exclusion criteria (see Table 3), resulting in 12 studies remaining. No date cut-off was used for screening the studies due to SSTP being a relatively new intervention and few studies have evaluated the effectiveness of SSTP. Full text screening led to 7 studies being excluded (see Appendix A), leaving 5 studies remaining (see Table 4). Figure 1 outlines the literature search process.

Table 3

Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion	Rationale
1. Participants	Must include parents/carers of	Sample includes participants who are not parents/carers	SSTP was designed for parents/carers of children aged 0-12

		children with ASD aged between 0-12	of children aged between 0-12	
2.	Diagnosis	Children must have a diagnosis of ASD or ASC	Children without a diagnosis of ASD or ASC	To evaluate the outcomes of this programme for children with ASD or ASC

	Criteria	Inclusion	Exclusion	Rationale
3.	Intervention	The study is based on the Stepping Stones Triple P Programme	The study does not include the Stepping Stones Triple P Programme	This review is evaluating the effectiveness of the Stepping Stones Triple P Programme
4.	Outcome	The study must include behavioural outcomes	The study does not include behavioural outcomes	This review is evaluating the effectiveness of SSTP for reducing children's behavioural difficulties
5.	Methodology	Quasi-experimental designs or randomised control trials	Studies without quasi-experimental designs or randomised control trials	Petticrew and Roberts' (2003) Hierarchy of Evidence proposes that randomised control trials, followed by quasi-experimental designs are most appropriate for evaluating effectiveness of interventions.
6.	Language	The study is published in English	The study is not published in English	Author's first language is English. Ensures the article can be fully understood and evaluated

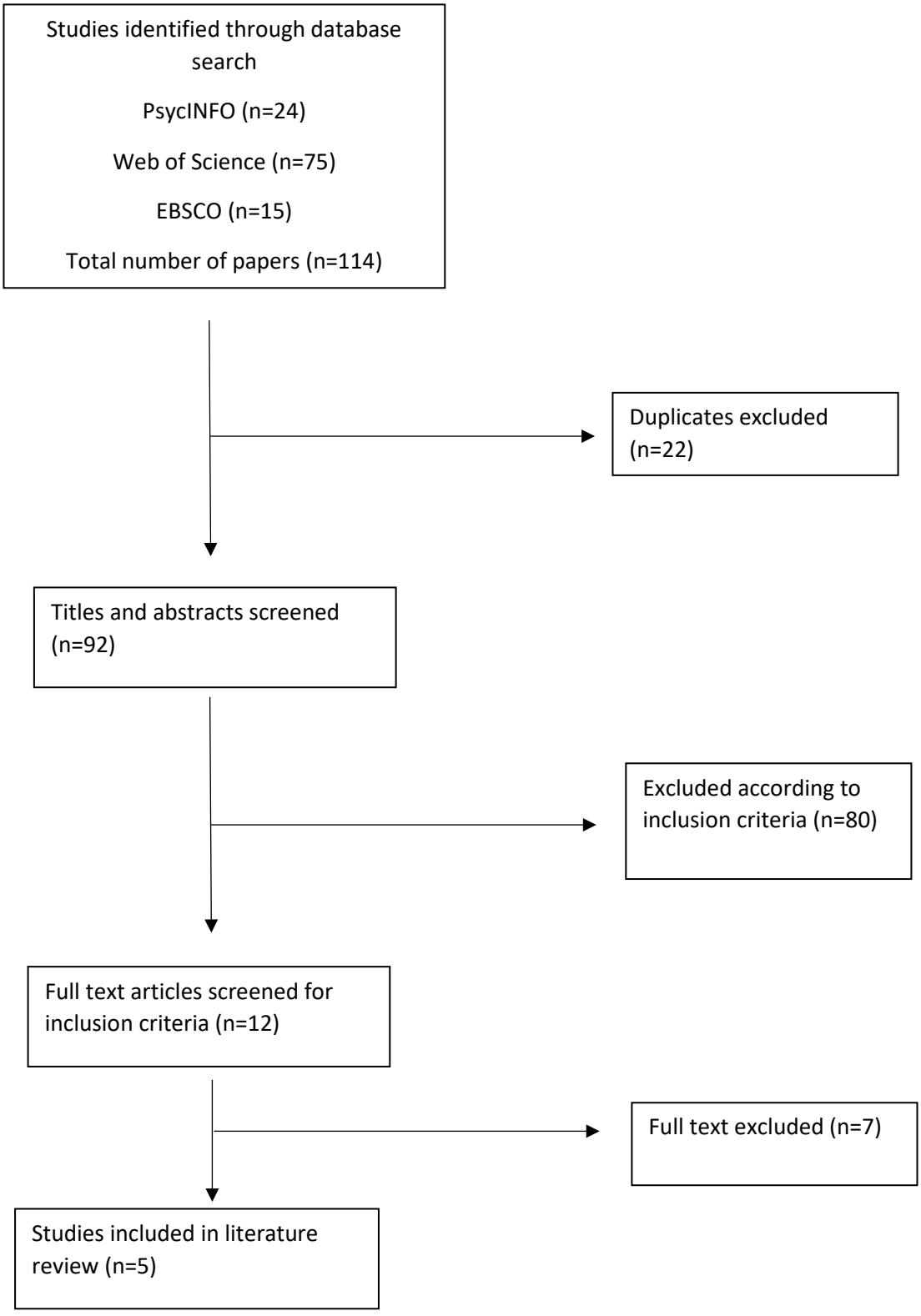
Table 4

References of Studies Included within this Review

References
Brian, J., Tint, A., Branson., J & Pilkington., M. (2021). Effectiveness of group stepping stones positive parenting program for children with autism spectrum disorder and disruptive behaviour: Program evaluation from a large community implementation. *Journal on Developmental Disabilities*, *26*(2), 1-20.
Kasperzack, D., Schrott, B., Mingebach, T., Becker, K., Burghardt, R., & Kamp-Becker, I. (2020). Effectiveness of the Stepping Stones Triple P group parenting program in reducing comorbid behavioral problems in children with autism. *Autism the international journal of research and practice*, *24*(2), 423–436. https://doi.org/10.1177/1362361319866063
Tellegen, C. L., & Sanders, M. R. (2014). A randomized controlled trial evaluating a brief parenting program with children with autism spectrum disorders. *Journal of Consulting and Clinical Psychology, 82*(6), 1193–1200. https://doi.org/10.1037/a0037246
Whittingham, K., Sofronoff, K., Sheffield, J., & Sanders, M. R. (2009). Stepping Stones Triple P: An RCT of a parenting program with parents of a child diagnosed with an autism spectrum disorder. *Journal of Abnormal Child Psychology*, 37, 469–480.
Zand, D.H., Bultas, M.W., McMillin, S.E., Halloran, D., White, T., McNamara, D. and Pierce, K.J. (2018). A Pilot of a Brief Positive Parenting Program on Children Newly Diagnosed with Autism Spectrum Disorder. *Family Process*, 57: 901-914. https://doi-org.libproxy.ucl.ac.uk/10.1111/famp.12334

Figure 1

Literature Search Process



Mapping the Field

The included studies in this review all examined the effectiveness of SSTP in reducing behavioural difficulties among children with ASD. However, they differed in the version of SSTP, research design, participants, outcome measures and implementation fidelity. In Appendix B, the key features of each study are outlined.

Weight of Evidence

Gough's (2007) Weight of Evidence (WoE) framework was used to critically appraise the five included studies, evaluating each study for quality and relevance. The WoE framework comprises of three main components; assessing methodological quality of a study (WoE A), the methodological relevance (WoE B) and the relevance of the topic to the review question (WoE C) (Gough, 2007).

For WoE A, the methodological quality of each study was assessed, using Kratochwills' (2003) coding protocol which was deemed appropriate for group-based designs. Kratochwills' (2003) coding protocol was chosen because it was developed to identify, review and code studies involving psychological and educational interventions for behavioural, emotional and academic problems for school aged children and their families, which is relevant to this review question. Modifications to Kratochwills' (2003) coding protocol, rationale for use and the criteria are outlined in Appendix C.

A coding protocol was designed by the researcher for WoE B, based on Petticrew and Roberts' (2003) Hierarchy of Evidence, to assess the methodological relevance of each study. Petticrew and Roberts (2003) proposed that when considering the effectiveness of an intervention, Randomised Control Trials (RCTs) are the highest quality of evidence, followed by quasi-experimental designs. A rating for the quality of evidence was given for each study, with higher quality of evidence receiving a higher rating. For WoE C, a coding protocol was developed by the researcher to assess how relevant the focus of the study is to answering the review

question. The rationale and detailed criteria are shown in Appendix D for WoE B and Appendix E for WoE C. To calculate WoE D, the weightings from WoE A, B and C were averaged to give an overall weighting, outlined in Table 5.

Table 5

Weight of Evidence Ratings

Study	WoE A: Methodological quality	WoE B: Methodological relevance	WoE C: Topic relevance	WoE D: Overall weight of evidence
Brian et al. (2021)	1	1	2.25	1.42 (low)
Kasperzack et al. (2020)	1.8	1	3	1.93 (medium)
Tellegen et al. (2014)	2.8	3	3	2.93 (high)
Whittingham et al. (2009)	2.6	3	2.25	2.62 (high)
Zand et al. (2018)	1.6	3	2.25	2.28 (medium)

Note. <1.7= low, 1.7-2.4=medium, >2.4= high

Participants

Across all five studies, there was a total of 299 participants (parents) recruited, ranging from 21 participants (Zand et al., 2018) in the smallest study to 131 participants in the largest study (Brian et al., 2021). All studies had low attrition rates, with less than 20% post-intervention and less than 30% at follow-up, reducing attrition bias and improving internal validity (Kratochwill, 2003), consequently receiving higher WoE A ratings.

In all five studies, the children of parents were between age 0-12 and therefore all received the highest weighting for WoE C ‘participant age’ criteria. This is because SSTP is designed for children aged 0-12 years. The raising of children in early years is argued to affect life outcomes, such as brain development, social skills, emotional regulation, mental and physical health, risk behaviours and self-control. Therefore, parenting programmes for

children are suggested to prevent behavioural problems and improve wellbeing (Sanders et al., 2014).

All children had a confirmed diagnosis of ASD from a paediatrician or other medical professional. Three of the studies (Kasperzack et al., 2020; Tellegen et al., 2014; Whittingham et al., 2009) conducted a pre-study screening for ASD, consequently receiving a higher WoE C 'participant diagnosis' criteria rating. Whittingham et al. (2009) and Tellegen et al. (2014) used a diagnostic interview based upon DSM-IV criteria and Kasperzack et al. (2020) used the Autism Diagnostic Observation Schedule-second edition (ADOS-2). Two studies (Brian et al., 2021; Zand et al., 2018) confirmed the children had a diagnosis of ASD but did not conduct a pre-study based screening and therefore received a 'low' rating on the WoE C 'participant diagnosis' criteria.

Study Design

Three of the five studies reviewed (Tellegen et al., 2014; Whittingham et al., 2009; Zand et al., 2018) were Randomised Control Trials (RCTs). RCTs are considered the 'gold standard' for evaluating the effectiveness of interventions (Petticrew & Roberts, 2003) and therefore received the highest WoE B rating. Randomisation is a rigorous tool for examining cause-effect relationships and reduces selection bias which is a threat to internal validity, allowing differences in outcomes to be attributed to the intervention (Barker et al., 2016; Hariton & Locascio, 2018).

Two of the studies (Brian et al., 2021; Kasperzack et al., 2020) used pre-post test group designs. They did not include randomisation to a treatment or control group in their design and therefore received a 'low' WoE B rating. This is because pre-post test designs without a control group are less able to determine causality due to not being able to eliminate confounding variables that might have affected outcomes rather than the intervention itself (Marsden & Carole, 2012).

Intervention

Four of the studies (Brian et al., 2021; Kasperzack et al., 2020; Tellegen et al., 2014; Zand et al., 2018) examined the effectiveness of SSTP as the primary intervention with two of the studies (Tellegen et al., 2014; Zand et al., 2018) using the Primary Care SSTP version and two of the studies (Brian et al., 2021; Kasperzack et al., 2020) utilising the group format of SSTP. Therefore due to using SSTP, like outlined in the manuals, as the primary intervention, they all received the highest WoE C rating for 'intervention criteria.'

Whittingham et al. (2009) used a partial group format of SSTP, with five group teaching sessions and four individual sessions for observation, practice and feedback. Comic Strip Conversations and Social Stories were added into one session (session seven), which are not currently used in SSTP, and therefore Whittingham et al. (2009) received a 'low' rating for 'intervention criteria'. This is because this review is examining the effectiveness of SSTP and additional interventions might have altered the outcomes.

Outcome Measures

All five studies included behavioural outcome measures, as outlined in the inclusion criteria, which were well-validated. Two studies (Kasperzack et al., 2020; Tellegen et al., 2014) used more than one assessment method for child behaviours. Kasperzack et al. (2020) used multiple assessment tools, including; Developmental Behaviour Checklist (DBC), Social Responsiveness Scale (SRS), Eyeberg Child Behaviour Inventory (ECBI) and Strengths and Difficulties Questionnaire (SDQ). Tellegen et al. (2014) utilised the ECBI and conducted 30-minute observations of parent-child behaviours. These two studies also used at least two sources for collecting data with Kasperzack et al. (2020) using teacher and parent data and Tellegen et al. (2014) using a parent questionnaire and observations conducted by the researchers. Reliable scores were reported for the majority of primary outcomes.

Kasperzack et al. (2020) reported a strong internal consistency score for the DBC (α = 0.94), SRS (α = 0.91), ECBI (Intensity scale: α = 0.94; Problem scale: α = 0.95) and total SDQ problem score (α = 0.82). A test-retest reliability score for DBC subscales was reported to be sufficient (ranging from .76 to .96). Similarly, Tellegen et al. (2014) reported internal consistency scores for the ECBI (Intensity scale = .91; Problem scale =.89). Through using multiple, reliable assessment tools and sources for data collection, this increases reliability and validity and consequently generalisability of results (Barker et al., 2016). Therefore these two studies received a 'high' rating for WoE A 'measurement' criteria.

The other three studies (Brian et al., 2021; Whittingham et al., 2009; Zand et al., 2018), utilised one assessment method and one source for collecting outcome data. This reduces construct validity, making it difficult to generalise these findings to other settings, such as school (Barker et al., 2016). Therefore, these studies received lower ratings for the 'measurement' WoE A criteria.

Whittingham et al. (2009) reported strong internal consistency scores for the ECBI Intensity (α=0.92) and Problem scores (α=0.88). Zand et al. (2018) reported strong reliability estimates for the ECBI for the Problem (a = .79) and Intensity (a = .86) scales. Brian et al. (2021) was the only study not to report reliability estimates, consequently receiving the lowest score on the 'measurements criteria' for WoE A.

Follow-Up

Three of the studies (Kasperzack et al., 2020; Tellegen et al., 2014; Whittingham et al., 2009) used outcome measures at three time points (pre-intervention, post-intervention and 6-month follow-up). Outcome measures allow researchers and practitioners to monitor changes in a participant or client (Hatfield & Ogles, 2004). Through using outcome measures pre-post-intervention and at follow-up, researchers can measure the effectiveness of an intervention over a longer period of time. Therefore, these studies received 'high' WoE A 'follow-up' and WoE C 'outcome measure' criteria ratings. The other two studies (Brian et al., 2021; Zand et al., 2018) used

outcome measures at two time points (pre and post intervention) and therefore received a lower rating for WoE A and C.

Implementation Fidelity

Implementation fidelity was assessed across all five studies and scored, forming part of the WoE A weighting. Three of the studies (Tellegen et al., 2014; Whittingham et al., 2009; Zand et al., 2018) received a 'strong evidence' WoE A 'implementation fidelity' rating because the practitioners had received formal SSTP training and received supervision throughout the programme to ensure the programme was implemented correctly. Kasperzack et al. (2020) and Brian et al. (2021) received a 'weak evidence' rating of 1 due to having received training for SSTP but not receiving supervision. Implementation fidelity is important for ensuring the programme is implemented correctly to determine if the intervention is responsible for the outcomes reported (Kratochwill, 2003).

Findings

All studies reported effect sizes, with two reporting Cohen's *d* (Brian et al., 2021; Tellegen et al., 2014) and three Eta Squared (Kasperzack et al.,2020; Whittingham et al., 2009; Zand et al., 2018). To compare studies, all effect sizes were converted into Cohen's *d* using the Psychometrica website (Lenhard & Lenhard, 2016). All five studies utilised behavioural outcome measures and found, post SSTP intervention, a reduction in child problem behaviours with significant effect sizes ranging from small to large. Table 6 outlines the outcome measures, effect sizes, using cohen's *d*, and main findings from the studies. Effect sizes reported were within-group for studies without a control group, including Brian et al. (2021), Kasperzack et al. (2020), comparing pre-post mean differences within the group. For Tellegen et al. (2014), Whittingham et al. (2009) and Zand et al. (2018), effect sizes were based on between-group mean differences, comparing mean differences between treatment and control group pre-post intervention.

Table 6
Summary of Effect Sizes

Study	Outcome Measure	Effect Size (Cohen's *d*)	Descriptor	Significance (p value)	Main Findings	WoE D
Brian et al. (2021)	SDQ Total Difficulties	d=0.43	Small	<0.001*	Parent SDQ ratings indicated significant decreases in their child's behavioural difficulties. These improvements were modest and had a small effect size.	1.42 (low)
Kasperzack et al. (2020)	DBC (parent)	d=1.0931	Large	<0.001*	The study findings indicated the effectiveness of SSTP for reducing behaviour difficulties among children with ASD. Significant medium to large effect sizes for all parent behavioural measures were found and for the teacher SRS and SDQ measure but not the DBC measure.	1.93 (medium)
	DBC (teacher)	d=0.667	Medium	0.234		
	SRS (parent)					
	SRS (teacher)	d=0.7385	Medium	0.036*		
	ECBI Intensity	d=0.9686	Large	0.039*		
	ECBI Problem					
	SDQ (parent)	d=0.937	Large	0.003*		
	SDQ (teacher)	d=1.0312	Large	0.002*		
		d=0.7731	Medium	0.021*		
		d=0.9686	Large	0.044*		

Study	Outcome Measure	Effect Size	Descriptor	Significance (p value)	Main Findings	WoE D
Tellegen et al. (2014)	ECBI Intensity	d=0.40	Small	0.025*	Significant short-term reductions in behavioural difficulties were found from the ECBI outcome measures. No significant effect was found from observations. This effect was not maintained at follow-up.	2.93 (high)
	ECBI Problem	d=0.56	Medium	0.018*		
	Observed child and parent behaviours: Disruptive child	d=-0.05	Small	.848		
Whittingham et al. (2009)	ECBI Intensity	d=1.18855	Large	<0.001*	Significant reductions found in child behaviour difficulties and effects were maintained at 6-month follow-up.	2.7 (high)
	ECBI Problem	d=0.8729	Large	<0.001*		
Zand et al. (2018)	ECBI Intensity	d=1.1855	Large	<.02*	Reductions in child behaviour problems, with a significant reduction found for intensity of child disruptive behaviours but a non-significant effect for number of problem behaviours.	2.28 (medium)
	ECBI Problem	d=0.937	Large	<.06		

Note: *d* 0.2 = small effect size, *d* 0.5 = medium effect size, *d* 0.8 = large effect size

*Significant at p<0.05

Brian et al. (2021) found a significant improvement ($p<.001$), with a small effect size, in children's behavioural difficulties, measured through the SDQ. Kasperzack et al. (2020) found a significant reduction in behaviour problems across all four parental rated outcome measures (DBC, ECBI, SRS and SDQ) with medium to large effect sizes, at both post-intervention and follow-up, indicating the effectiveness of SSTP. Supporting this, teacher ratings on the SDQ and SRS showed a significant reduction ($p<.05$) in behaviour difficulties with large effect sizes. However, the teacher rated DBC did not reach statistical significance. Nevertheless, this accentuates the effect of SSTP in reducing behavioural difficulties across different settings.

Although Kasperzack et al. (2020) and Brian et al. (2021) reported significant reductions in behavior problems, indicating the effectiveness of the SSTP intervention, they did not have a control group to make a comparison against. This makes it difficult to determine the effectiveness of the intervention with potential confounding variables not being controlled for. This resulted in a lower WoE A and B score and consequently WoE D score.

Tellegen et al. (2014) found significant short-term intervention effects ($p<0.05$) on reducing behaviour difficulties through using the ECBI Intensity and Problem scales with a small effect size found for ECBI Intensity and medium effect size for ECBI Problem scale. However, this effect was not maintained at 6-month follow-up and significant effects were not found for observed child behaviours.

A significant decrease ($p<0.001$) in behaviour problems, with a large effect size, was found in Whittingham et al.'s (2009) study, through using the ECBI Intensity and Problem scales. This was found at both post-intervention and 6-month follow-up.

Zand et al. (2018) found a significant reduction with a large effect size in the intensity of children's disruptive behaviour, measured through the ECBI Intensity scale. Although parent's reported a decline in the number of problem behaviours, measured through the ECBI Problem scale, this result was not statistically significant.

Discrepancy was found between studies regarding the statistical analysis. Two of the studies (Brian et al., 2021; Whittingham et al., 2009) received a 'high' weighting on the WoE A 'statistical analysis' criteria, having used appropriate statistical analysis, controlling for familywise error rate through applying Bonferroni correction and having a sufficiently large sample size. Tellegen et al. (2014) used appropriate unit of analysis for all three measures and had a sufficiently large sample size, indicating promising evidence. However, familywise error rate was not shown to be controlled for in the study. Kazperzack et al. (2020) used appropriate unit of analysis for all measures and controlled for familywise error rate through Bonferroni correction but did not have a sufficiently large sample size. Zand et al. (2018) received a lower 'analysis' rating due to not controlling for familywise error rate or having a sufficiently large sample size, indicating 'weak evidence', demonstrated in the WoE A and D scores. A small sample size can affect the reliability of the results and reduce generalisability (Barker et al., 2016). Therefore, the significant effect sizes for these studies (Kazperzack et al., 2020; Zand et al., 2018) should be interpreted with caution.

Conclusions and Recommendations

The aim of this study was to evaluate the effectiveness of SSTP on reducing behavioural difficulties among children with ASD. Five studies which met the inclusion criteria were reviewed. Gough's (2007) weight of evidence framework was used to assess the quality of the findings. Four of the studies received 'medium' to 'high' WoE D scores (Kasperzack et al., 2020; Tellegen et al., 2014; Whittingham et al., 2009; Zand et al., 2018) and one (Brian et al., 2018) a 'low' WoE D rating. All five studies found the SSTP intervention significantly reduced behavioural difficulties among children with ASD with effect sizes ranging from small to large. Overall, these findings indicate that SSTP is an effective intervention for parents of children with ASD for reducing behavioural difficulties.

However, most of the findings were based on parent reported data with only two studies utilising two sources (Kasperzack et al.,2020; Tellegen et al.,

2014). This reduces the ecological validity of the findings and makes it difficult to generalise these findings to other settings, such as school. Future research should utilise more information sources, such as teachers, and more assessment tools, to examine the impact of SSTP on behavioural difficulties in other settings.

Furthermore, two of the studies did not include control groups, making it difficult to conclude whether the impact was due to the intervention or confounding variables that the study did not control for which could overestimate the effectiveness of the intervention (Marsden & Carole, 2012). Therefore, more randomised control trials would strengthen the argument for the effectiveness of SSTP for reducing behaviour difficulties among children with ASD.

Three of the studies (Kasperzack et al., 2020; Tellegen et al., 2014; Whittingham et al., 2009) examined longer term effects of SSTP through a 6-month follow up. Significant reductions in behavioural problems were found at 6-month follow up by two of the studies (Kasperzack et al., 2020; Whittingham et al., 2009) but not by Tellegen et al. (2014). Two of the studies (Brian et al., 2021; Zand et al., 2018) did not examine follow-up effects. Therefore, the long term effectiveness of SSTP is unclear and more research is needed in this area.

Implications

The current review supports the effectiveness of SSTP for reducing behavioural difficulties among children with ASD. With the prevalence of behavioural difficulties among children with ASD, parenting interventions, such as SSTP, can help reduce behaviour problems and improve outcomes for children with ASD and families. In the current climate with an increase in referrals for ASD and long waiting times (NHS, 2022), low intensity parenting interventions, such as SSTP, can provide a cost-effective and accessible option for parents of children with ASD. Educational psychologists and other professionals have a role in supporting schools and families to access programmes, such as SSTP, to manage behavioural difficulties and improve outcomes for children with ASD.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders.* 5th ed. Arlington, VA: American Psychiatric Association.

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *The Journal of Abnormal and Social Psychology, 66*(1), 3-11. doi:http://dx.doi.org/10.1037/h0048687

Bandura, A (1977) *Social Learning Theory.* New York: General Learning Press.

Barker, C., Pistrang, N., & Elliott, R. (2016). *Research methods in clinical psychology : An introduction for students and practitioners / Chris Barker and Nancy Pistrang, and Robert Elliott.* (Third ed.).

Brian, J., Tint, A., Branson., J & Pilkington., M. (2021). Effectiveness of group stepping stones positive parenting program for children with autism spectrum disorder and disruptive behaviour: Program evaluation from a large community implementation. *Journal on Developmental Disabilities*, *26*(2) ,1-20

Early Intervention Foundation. (2023, January). Group Stepping Stones Triple P. PDF. London; Early Intervention Foundation.

Gough D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Applied and Practice-based Research. Special Edition of Research Papers in Education, 22* (2), 213-228. https://doi.org/10.1080/02671520701296189

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG : an international journal of obstetrics and gynaecology*, *125*(13), 1716. https://doi.org/10.1111/1471-0528.15199

Hartley, S. L., Sikora, D. M., & McCoy, R. (2008). Prevalence and risk factors of maladaptive behaviour in young children with Autistic Disorder. *Journal of Intellectual Disability Research*, 52(10), 819–829. https://doi.org/10.1111/j.1365-2788.2008.01065.x

Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice*, *35*(5), 485–491. https://doi.org/10.1037/0735-7028.35.5.485

Kasperzack, D., Schrott, B., Mingebach, T., Becker, K., Burghardt, R., & Kamp-Becker, I. (2020). Effectiveness of the Stepping Stones Triple P group parenting program in reducing comorbid behavioral problems in children with autism. *Autism : the international journal of research and practice*, *24*(2), 423–436. https://doi.org/10.1177/1362361319866063

Kratochwill, T. R. (2003). Task Force on Evidence Based Interventions in School Psychology. American Psychological Association.

Lenhard, W. & Lenhard, A. (2016) *Computation of effect size.* Retrieved from: h ttps://www.psychometrica.de/effect_size.html. Psychoemtrica.

Marsden, E & Carole J. T. (2012) Single group, pre- and post-test research designs: Some methodological concerns, *Oxford Review of Education*, 38(5), 583-616, DOI: 10.1080/03054985.2012.731208

McClintock, K., Hall, S., & Oliver, C. (2003). Risk markers associated with challenging behaviours in people with intellectual disabilities: a meta-analytic study. *Journal of intellectual disability research: JIDR*, *47*(Pt 6), 405–416. https://doi.org/10.1046/j.1365-2788.2003.00517.x

Morawska, A., & Sanders, M. R. (2006). Self-directed behavioural family intervention.*International Journal of Behavioral Consultation and Therapy, 2*(3), 332–340. https://doi.org/10.1037/h0100787

NHS. (2022, March). *Autism Statistics, April 2019 to December 2021*. NHS choices. Retrieved February 10, 2023, from https://digital.nhs.uk/data-and-information/publications/statistical/autism-statistics/april-2019-to-december-2021

Nicholls, G., Hastings, R. P., & Grindle, C. (2020). Prevalence and correlates of challenging behaviour in children and young people in a special school

setting. *European Journal of Special Needs Education*, *35*(1), 40–54. https://doi.org/10.1080/08856257.2019.1607659

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health*, *57*, 527–529. doi:10.1136/jech.57.7.527

Reid, J. B., Patterson, G. R., & Snyder, J. J. (2002). *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. American Psychological Association. https:// doi.org/10.1037/10468-000

Ruane, A. and Carr, A. (2019), Systematic Review and Meta-analysis of Stepping Stones Triple P for Parents of Children with Disabilities. Family Process., 58: 232-246. https://doi.org/10.1111/famp.12352

Sanders M. R. (1999). Triple P-Positive Parenting Program: towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical child and family psychology review*, *2*(2), 71–90. https://doi.org/10.1023/a:1021843613840

Sanders, M. R., Mazzucchelli, T. G., & Studman, L. J. (2004). Stepping Stones Triple P: The theoretical basis and development of an evidence-based positive parenting program for families with a child who has a disability. *Journal of Intellectual and Developmental Disability, 29*(3), 265–283. https://doi.org/10.1080/13668250412331285127

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

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

Skinner, B. F. (1953). *Science and human behavior.* Macmillan.

Steenfeldt-Kristensen, C., Jones, C. A., & Richards, C. (2020). The prevalence of self-injurious behaviour in autism: A meta-analytic study. *Journal of Autism and Developmental Disorders*, *50*(11), 3857–3873. https://doi.org/10.1007/s10803-020-04443-1

Tellegen, C. L., & Sanders, M. R. (2014). A randomized controlled trial evaluating a brief parenting program with children with autism spectrum disorders. *Journal of Consulting and Clinical Psychology, 82*(6), 1193–1200. https://doi.org/10.1037/a0037246

Triple P (2021a). *Group Stepping Stones Triple P*. Triple P International Pty Ltd.

Triple P. (2021b). *Primary Care Stepping Stones Triple P*. Triple P International Pty Ltd.

Triple P (2021c). *Standard Stepping Stones Triple P*. Triple P International Pty Ltd.

Weiss, J. A., Cappadocia, M. C., MacMullin, J. A., Viecili, M., & Lunsky, Y. (2012). The impact of child problem behaviors of children with ASD on parent mental health: The mediating role of acceptance and empowerment. *Autism,* 16(3), 261-274. https://doi.org/10.1177%2F1362361311422708

Whittingham, K., Sofronoff, K., Sheffield, J., & Sanders, M. R. (2009). Stepping Stones Triple P: an RCT of a parenting program with parents of a child diagnosed with an autism spectrum disorder. *Journal of abnormal child psychology*, *37*(4), 469–480. https://doi.org/10.1007/s10802-008-9285-x

Yates, K., & Le Couteur, A. (2016). Diagnosing autism/autism spectrum disorders. *Paediatrics and Child Health*, *26*(12), 513–518. https://doi.org/10.1016/j.paed.2016.08.004

Zand, D.H., Bultas, M.W., McMillin, S.E., Halloran, D., White, T., McNamara, D. and Pierce, K.J. (2018), A Pilot of a Brief Positive Parenting Program on Children Newly Diagnosed with Autism Spectrum Disorder. *Family Process*., 57: 901-914. https://doi-org.libproxy.ucl.ac.uk/10.1111/famp.12334

Appendices

Appendix A: Excluded Studies at Full Paper Review

Table 7

Excluded Studies

Reference	Reason for Exclusion	Inclusion/Exclusion Criteria Number
Farris, O., Royston, R., Absoud, M., Ambler, G., Barnes, J., Hunter, R., Kyriakopoulos, M., Oulton, K., Paliokosta, E., Panca, M., Paulauskaite, L., Poppe, M., Ricciardi, F., Sharma, A., Slonims, V., Summerson, U., Sutcliffe, A., Thomas, M., & Hassiotis, A. (2020). Clinical and cost effectiveness of a parent mediated intervention to reduce challenging behaviour in pre-schoolers with moderate to severe intellectual disability (EPICC-ID) study protocol: A multi-centre, parallel-group Randomised Controlled Trial. *BMC Psychiatry*, *20*(1). https://doi.org/10.1186/s12888-020-2451-6	2	Did not measure child behavioural outcomes
Hodgetts, S., Savage, A., & McConnell, D. (2013). Experience and outcomes of stepping stones triple P for families of children with autism. *Research in Developmental Disabilities*, *34*(9), 2572–2585. https://doi.org/10.1016/j.ridd.2013.05.005	4	Did not measure child behavioural outcomes
Roux, G., Sofronoff, K., & Sanders, M. (2013). A randomized controlled trial of group Stepping Stones Triple P: A mixed-disability trial. *Family Process*, *52*(3), 411–424. https://doi.org/10.1111/famp.12016	2	Included participants without ASD

Reference	Reason for Exclusion	Inclusion/Exclusion Criteria Number
Schrott, B., Kasperzack, D., Weber, L., Becker, K., Burghardt, R., & Kamp-Becker, I. (2018). Effectiveness of the stepping stones triple P group parenting program as an additional intervention in the treatment of autism spectrum disorders: Effects on parenting variables. *Journal of Autism and Developmental Disorders*, *49*(3), 913–923. https://doi.org/10.1007/s10803-018-3764-x	4	Did not measure child behavioural outcomes
Sofronoff, K., Jahnel, D., & Sanders, M. (2011). Stepping stones triple P seminars for parents of a child with a disability: A randomized controlled trial. *Research in Developmental Disabilities*, *32*(6), 2253–2262. https://doi.org/10.1016/j.ridd.2011.07.046	2	Included participants without ASD
Whittngham, K., Sofronoff, K., & Sheffield, J. (2006). Stepping stones triple P: A pilot study to evaluate acceptability of the program by parents of a child diagnosed with an autism spectrum disorder. *Research in Developmental Disabilities*, *27*(4), 364–380. https://doi.org/10.1016/j.ridd.2005.05.003	4	Did not measure child behavioural outcomes
Whittingham, K., Sofronoff, K., Sheffield, J., & Sanders, M. R. (2009). Behavioural family intervention with parents of children with ASD: What do they find useful in the parenting program stepping stones triple P? *Research in Autism Spectrum Disorders*, *3*(3), 702–713. https://doi.org/10.1016/j.rasd.2009.01.009	4	Did not measure child behavioural outcomes

Appendix B: Mapping the Field

Table 8

Mapping the Field

Author	Participants	Demographics	Study Design	Intervention	Outcome Measures	Outcome	Follow up
Brian et al. (2021)	131 parents of children with ASD aged 2-12 years	Relationship to child: 80 mothers, 48 fathers and 1 aunt No data on ethnicity or gender of child	Single group pre-post design	Group Stepping Stones Triple P (SSTP) 9 weeks	Strengths and Difficulties Questionnaire (SDQ)	Significant improvements found through the SDQ, for children's total difficulties, behavioural (conduct) problems, hyperactivity/inattention, and prosocial behavior (all p's < .001) but not for emotional symptoms or peer problems. These results indicate a reduction in behaviour problems.	No follow-up

Author	Participants	Demographics	Study Design	Intervention	Outcome Measures	Outcome	Follow up
Kasperzack et al. (2020)	Parents of 24 children with ASD aged between 3.6 and 12 years	Gender of children: 21 male, 3 female Relationship to child: 20 mothers, 3 fathers No ethnicity data	Single group repeated measures design	Group SSTP 6 weeks	Developmental Behavior Checklist (DBC) (filled in by parents and teachers) Social Responsiveness Scale (SRS) (filled in by parents and teachers) Eyeberg Child Behavior Inventory (ECBI) (filled in by parent only) SDQ (filled in by teachers and parents)	Comorbid behavioural problems, measured with the parent and teacher ECBI, SDQ and SRS showed a significant decline at post-treatment DBC indicated a significant reduction in behaviour problems, measured by parent ratings but not by teacher ratings	Post-hoc analysis revealed significant effects were maintained at follow up

Author	Participants	Demographics	Study Design	Intervention	Outcome Measures	Outcome	Follow up
Tellegen & Sanders (2014)	64 parents/carers of children aged 2-9 years with an ASD diagnosis Intervention group=35 Control group=29	Gender of children: 55 male 9 female Ethnicity of children: 57 White Australian 7 'other' Parents relationship to child: 61 mothers 3 fathers	Randomized two-arm clinical trial with two conditions (intervention, CAU)	Primary Care SSTP Four individual sessions (between 15-105 minutes long)	Eyeberg Child Behaviour Inventory (ECBI) Observations of 30-minute parent-child interactions	Significant short-term improvements were found in the intervention group, measured through the ECBI No significant effects found from observations of child behaviour	Follow up univariate analysis found no significant interaction effects, indicating improvements were not maintained

Author	Participants	Demographics	Study Design	Intervention	Outcome Measures	Outcome	Follow up
Whittingham et al. (2009)	59 families with a child with an ASD diagnosis aged between 2-9 Intervention group = 29 Control group =30 (wait list)	Gender of children: 47 male 12 female Relationship to child: 54 mothers 4 fathers 1 grandmother No ethnicity data	Randomised control trial	SSTP in a partial group format Comic Strip Conversations and Social Stories added to intervention Fortnight of practice sessions and 1-3 individual sessions	ECBI	A significant reduction in parent reported child behavioural problems for the intervention group compared to the wait-list group on the ECBI Intensity Scale and Problem Scale One third of the treatment group experienced significant change in child behaviour as measured by the ECBI Intensity Scale (34.5%) and the ECBI Problem scale (37.9%)	Follow-up data, using MANOVAS, indicated the change was maintained 6 months after the intervention was completed

Author	Participants	Demographics	Study Design	Intervention	Outcome Measures	Outcome	Follow up
Zand et al. (2018)	21 parents/carers with children with newly diagnosed ASD, ages 2-12 Intervention group=12 Control group=9	Gender of child: 18 male 3 female Race of children: Caucasian= 14 African-American=5 Bi-racial=2 Gender of parent/carer: 21 female	Pretest/posttest two group design with random assignment to the PC SSTP intervention versus the Wait List Control	Primary Care SSTP (4 session manualised positive parenting programme)	ECBI	Clinically and statistically significant reductions in; intensity of child disruptive behaviours but not the number of behaviours	No follow-up

Appendix C: WoE A Methodological Quality

An adapted version of Kratochwills' (2003) coding protocol was used to assess the methodological quality of all five studies and weighted, giving a WoE A rating. Kratochwills' (2003) coding protocol was chosen because it was formed to identify, review and code studies involving educational and psychological interventions for emotional, behavioural and academic problems for school students and their families. Kratochwills' (2003) manual is appropriate for group-based designs. Therefore, it was deemed appropriate for the five studies reviewed in this study.

Kratochwills' coding manual (2003) outlines how to code each section, resulting in weighting scores from 0-3. A score of three is considered 'strong evidence', two is 'promising evidence', 1 is 'weak evidence' and 0 is 'no evidence.' All studies scores were averaged across the five areas to give an overall WoE A rating. Items assessed and ratings can be found in table 10 and Appendix F. Details of the criteria for each feature can be found in Kratochwills' (2003) coding manual. Adaptions to the coding protocol and the rationale for these changes are outlined in Table 9.

Table 9

Adapted Sections to Kratochwills' Coding Protocol

Removed	Item	Rationale
I.	General Study Characteristics	Within the review, the study characteristics are discussed.
II.	C Data Analysis C7 Coding (removed) C8 Interactive process (removed) C9 Rival Interpretations (removed)	Not relevant for this review due to studies being quantitative not qualitative.

	Removed Item	Rationale
III.	Key Features for Coding Studies and Rating Level of Evidence/Support. A: Research Methodology (removed)	Relevant research methodology components discussed in review in WoE B section.
IV.	B Measurement. Removed: B4. Extent of engagement B5. Cultural appropriateness of the measures B7. Measures of key outcomes are linked to the conceptual model	Not relevant to the scope of this review.
V.	Primary/Secondary Outcomes that is Statistically Significant. Removed: D2 Percentage of primary outcomes that is statistically significant D3. Evidence of appropriate statistical analysis for secondary outcomes. D4. Percentage of secondary outcomes that is statistically significant D5. Overall Summary of Questions Investigated D6. Cultural Moderator Variables	Not relevant to this review. This review does not focus on secondary outcomes.

	Removed Item	Rationale
VI.	E. Cultural Significance	Not relevant to the scope of this review.
VII.	F. Educational/Clinical Significance	This is discussed elsewhere in this review.
VIII.	G. External Validity Indicators	This is discussed elsewhere in this review.
IX.	H. Durability of Intervention and Outcomes. Removed: H2 Durability/Generalization over time H3 Durability/Generalization across settings H4 Durability/Generalization across persons	This is discussed elsewhere in this review.
X.	I Identifiable Intervention Components (all sections)	Not relevant to this review.
XI.	J Implementation Fidelity: J4 Implementation Context (removed)	Not relevant to this review.
XII.	K. Replication (all sections)	Not relevant to this review.
XIII.	L. Site of Implementation	Not relevant to this review.

Table 10

Weight of Evidence (WoE) A ratings

Study	Measures	Comparison Group	Analysis	Follow up	Implementation Fidelity	Overall WoE A	Descriptor
Brian et al. (2021)	1	0	3	0	1	1	Low
Kasperzack et al., (2020)	3	0	2	3	1	1.8	Medium
Tellegen et al., (2014)	3	3	2	3	3	2.8	High
Whittingham et al., (2009)	2	2	3	3	3	2.6	High
Zand et al., (2018)	2	2	1	0	3	1.6	Low

Note. <1.7= low, 1.7-2.4=medium, >2.4= high

Appendix D WoE B Methodological Relevance

The methodological relevance of each study was assessed and weighted for WoE B, using Petticrew and Roberts' (2003) Hierarchy of Evidence. Petticrew and Roberts (2003) propose that Randomised Control Trial designs are the most appropriate design for answering 'effectiveness' questions, followed by quasi-experimental designs. Studies with a control group are argued to be of higher quality compared to studies without a control group. The criteria for WoE B ratings can be found in Table 11. A rating for the quality of evidence was given for each study, with higher quality of evidence receiving a higher rating (see Table 12).

Table 11

Criteria for WoE B Ratings

Criteria	WoE Rating	Descriptor	Rationale
A: Study Type	3 (high)	Randomised Control Trial	It is important to consider the study design when assessing the quality of each study
	2 (medium)	Quasi-experimental designs with a control group	
	1 (low)	Quasi-experimental designs without a control group, cohort studies	

Table 12

WoE B Ratings for Studies in this Review

Study	WoE B	Descriptor
Brian et al. (2021)	1	Low
Kasperzack et al., (2020)	1	Low
Tellegen et al., (2014)	3	High
Whittingham et al., (2009)	3	High
Zand et al., (2018)	3	High

Note. <1.7= low, 1.7-2.4=medium, >2.4= high

Appendix E: WoE C Topic Relevance

WoE C assessed how appropriate the design of the study was to the review question. The criteria in table 13 was developed by the reviewer, considering the intervention, participant diagnosis, participant age and outcome measures. The rationale for each criteria is described in Table 13 and WoE C ratings for each study can be found in Table 14.

Table 13

Criteria for WoE C Ratings

Criteria	WoE Rating	Descriptor	Rationale
A. Intervention	3	Stepping Stones Triple P (SSTP) is the primary intervention	This review is evaluating the effectiveness of SSTP and therefore the studies using the original SSTP version are most appropriate.
	2	An alternative version of SSTP is used	
	1	SSTP is combined with another intervention	
	0	SSTP is not used within the study	
B. Participant Diagnosis	3	Children have a diagnosis of ASD from a medical professional, confirmed by a study based screening. Scores reported.	This review is evaluating the effectiveness of reducing behavioural difficulties among children with ASD.
	2	Children have a diagnosis of ASD from a medical professional, confirmed by a study based screening. Scores not reported.	
	1	Children have a diagnosis of ASD from a medical professional with no study based screening.	
	0	No information provided about the child's diagnosis	

	Criteria	WoE Rating	Descriptor	Rationale
C.	Participant Age	3	Age 0-12	SSTP is designed for children aged 0-12
		2	Age 12-18	
		1	Age 18-25	
		0	Age 25 and above	
D.	Outcome Measures	3	Behavioural outcome measures used at 3 time points or more	This study is measuring the effectiveness of the SSTP on reducing behavioural difficulties. Through measuring at multiple time points, this highlights the durability of the effect of SSTP.
		2	Behavioural outcome measures used at 2 time points or more	
		1	Behavioural outcome measures used at 1 time point	
		0	No behavioural outcome measures used	

Table 14

WoE C Ratings

Study	Criteria A	Criteria B	Criteria C	Criteria D	Overall WoE C	Descriptor
Brian et al. (2021)	3	1	3	2	2.25	Medium
Kasperzack et al., (2020)	3	3	3	3	3	High
Tellegen et al., (2014)	3	3	3	3	3	High
Whittingham et al., (2009)	1	2	3	3	2.25	Medium
Zand et al., (2018)	3	1	3	2	2.25	Medium

Note. <1.7= low, 1.7-2.4=medium, >2.4= high

Appendix F: Coding Protocols

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

Coding Protocol

Full Study Reference in proper format: Kasperzack, D., Schrott, B., Mingebach, T., Becker, K., Burghardt, R., & Kamp-Becker, I. (2020). Effectiveness of the Stepping Stones Triple P group parenting program in reducing comorbid behavioral problems in children with autism. *Autism : the international journal of research and practice*, *24*(2), 423–436. https://doi.org/10.1177/1362361319866063

Intervention Name (description of study): Stepping Stones Triple P

☐ Type of Publication:

☐ Book/Monograph

☒ Journal Article

☐ Book Chapter

☐ Other (specify):

1. General Characteristics

A. General Design Characteristics

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

☐ Completely randomized design

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

☐ Randomized block design (within participants)

☐ Randomized hierarchical design (nested treatments)

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

☐ Nonrandomized design

☐ Nonrandomized block design (between participants)

☐ Nonrandomized block design (within participants)

☐ Nonrandomized hierarchical design

☒ Optional coding for Quasi-experimental designs

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

☐ Very low (little basis)

☐ Low (guess)

☐ Moderate (weak inference)

☐ High (strong inference)

☐ Very high (explicitly stated)

☒ N/A

☐ Unknown/unable to code

B. Participants

Total size of sample (start of study): 23 families with 24 children with ASD

Intervention group sample size: 23 families with 24 children with ASD

Control group sample size: No control group

C. Type of Program

☐ Universal prevention program

☐ Selective prevention program

☐ Targeted prevention program

☒ Intervention/Treatment

☐ Unknown

C. Stage of Program

☐ Model/demonstration programs

☐ Early stage programs

☒ Established/institutionalized programs

☐ Unknown

D. Concurrent or Historical Intervention Exposure

☒ Current exposure

☐ Prior exposure

☐ Unknown

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

(Rating Scale: **3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)**

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

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

☒Yes

☐No

☐Unknown/unable to code

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

☒ Yes

☐ No

☐ N/A

☐ Unknown/unable to code

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

☒ Yes – teacher and parent

☐ No

☐ N/A

☐ Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

☒ Yes validated with specific target group

☐ In part, validated for general population only

☐ No

☐ Unknown/unable to code

Overall Rating for measurement=3

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

B. Comparison Group

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

☐ Typical intervention (typical intervention for that setting, without additions that make up the intervention being evaluated)

☐ Attention placebo

☐ Intervention element placebo

☐ Alternative intervention

☐ Pharmacotherapy

☐ No intervention

☐ Wait list/delayed intervention

☐ Minimal contact

☒ Unable to identify type of comparison

B2 Overall confidence of judgment on type of comparison group

☐ Very low (little basis)

☐ Low (guess)

☐ Moderate (weak inference)

☐ High (strong inference)

☐ Very high (explicitly stated)

☒ Unable to identify comparison group

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

☐ By change agent

☐ Statistical (analyse includes a test for intervention)

☐ Other

☒ Not reported/None

B4 Group equivalence established (select one of the following) – not reported

☐ Random assignment

☐ Posthoc matched set

☐ Statistical matching

☐ Post hoc test for group equivalence

B5 Equivalent mortality

☒ Low attrition (less than 20 % for post)

☒ Low attrition (less than 30% for follow-up)

☒ Intent to intervene analysis carried out?

Overall rating for Comparison group = 0

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

C. Appropriate Statistical Analysis

Analysis 1: DBC

☒ Appropriate unit of analysis

☒ Familywise/experimenter wise error rate controlled when applicable – Bonferri correction applied to avoid alpha error inflation

☐ Sufficiently large N - small sample size noted in limitations (no power anlaysis/calculation stated)

Analysis 2: ECBI

☒ Appropriate unit of analysis

☒ Familywise/experimenter wise error rate controlled when applicable - Bonferri correction applied to avoid alpha error inflation

☐ Sufficiently large N – small sample size noted in limitations (no power anlaysis/calculation stated)

Analysis 3: SDQ

☒ Appropriate unit of analysis

☒ Familywise/experimenter wise error rate controlled when applicable - Bonferri correction applied to avoid alpha error inflation

☐ Sufficiently large N - small sample size noted in limitations (no power anlaysis/calculation stated)

Analysis 4: SRS

☒ Appropriate unit of analysis

☒ Familywise/experimenter wise error rate controlled when applicable - Bonferri correction applied to avoid alpha error inflation

☐ Sufficiently large N - small sample size noted in limitations (no power anlaysis/calculation stated)

Overall rating for Statistical Analysis = 2

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

H. Durability/Generalization of Intervention and Outcomes

H1. Follow-up assessment

H1.1 Timing of follow up assessment: ☒yes ☐ no

Specify – 6-month follow up

H1.2. Number of participants included in the follow up assessment: ☒yes ☐ no

Specify 22 families

H1.3, Consistency of assessment method used: ☒yes ☐ no

Specify DBC, SRS, ECBI and SDQ

Overall Rating for Follow-up Assessment (select 0, 1, 2, or 3): ☒3 ☐2 ☐1 ☐0

J. Implementation Fidelity

J1. Evidence of Acceptable Adherence (answer J1.1 through J1.3) – N/A – not stated

J1.1 ☐Ongoing supervision/consultation

J1.2 ☐Coding intervention sessions/lessons or procedures

J1.3 ☐Audio/video tape implementation (select J1.3.1 or J1.3.2):

J1.3.1 Entire intervention

J1.3.2 Part of intervention

J2. Manualization (select all that apply) – official SSTP manual used

J2.1 ☐Written material involving a detailed account of the exact procedures and the sequence in which they are to be used

J2.2 ☒Formal training session that includes a detailed account of the exact procedures and the sequence in which they are to be used – 2 day workshop and accreditation in SSTP

J2.3 ☐Written material involving an overview of broad principles and a description of the intervention phases

J2.4 ☐Formal or informal training session involving an overview of broad principles and a description of the intervention phases

Rating for Implementation Fidelity (select 0, 1, 2, or 3): ☐3 ☐2 ☒1 ☐0

Summary of Evidence

Indicator	Overall evidence rating 0-3 NNR= No numerical rating	Description of evidence Strong Promising Weak No/limited evidence Or Descriptive ratings
General Characteristics		
Design	NNR	
Type of programme	NNR	
Stage of programme	NNR	
Concurrent/ historical intervention exposure	NNR	
Key features		

Measurement	3	Strong
Comparison group	0	No evidence
Appropriate Statistical Analysis	2	Promising evidence
Follow Up Assessment	3	Strong evidence
Implementation Fidelity	1	Weak evidence