Case Study 1: An Evidence-Based Practice Review Report

Theme: School (setting) based interventions for children with special educational

ary

ttention involves two or more individuals coordinating their attention to
me thing (Baldwin, 1995), and has been found to be a deficit in children
utism Spectrum Disorder (ASD) (Meindl & Cannella-Malone, 2011).
abilities are strongly associated with language (Scaife & Bruner, 1975)
cial development (Kim & Mundy, 2012). Consequently, they are
nt to address in the preschool years .

This systematic literature review explored the effectiveness of joint attention interventions in preschool aged children (2-5 years old) with ASD. A literature search identified a total of ten studies which were critically appraised using Gough's (2007) Weight of Evidence Framework. Six studies were given medium weightings and four received low ratings. Findings revealed a mix of small, medium and large effect sizes.

Overall, findings indicate a positive effect of intervention on joint attention, however weaknesses were identified in study methodologies. As such, future research may wish to address the limitations discussed in this review. For

example, whether joint attention improvements are maintained and if interventions are also effective in a UK setting.

Introduction

Joint attention and impairments in ASD

Joint attention has been defined as “simultaneous engagement of two or more individuals in mental focus on one and the same external thing.” (Baldwin, 1995, p. 132). This coordinated attention can be initiated by verbal mea

the child or present in their natural home setting as this enables joint attention skills to be practised beyond the intervention sessions (White et al., 2011).

Rationale and Relevance

While a systematic review of joint attention interventions for children with ASD was conducted in 2016 (Murza et al.), this review included studies of children from eighteen months to eight years old. Therefore, this review aims to not only update this previous review but to look at a more specific age range of two-five year olds (preschool age) with the hope of ascertaining whether early intervention in joint attention is effective. Moreover, Muzra et al.'s (2016) review concluded that while joint attention interventions appear effective for this population it is unclear for whom such interventions are more or less effective for. Exploring the narrower age range of the preschool years enables this to be investigated.

This area is also of particular importance for EPs as early interventions can be cost effective by reducing the support and intervention those with ASD need later in life (Jacobson et al., 1998) and increasing "the likelihood of improved long-term outcomes" for children with ASD (Koegel et al., 2014, p. 52). Such early intervention has been associated with significant reductions in ASD symptoms and improved outcomes (Howard et al., 2005; McEachin, Smith & Lovaas, 1993; Sallows & Graupner, 2005). This may be due to the brain being primed to learn social skills in the early years, meaning early

intervention facilitates the further development of these skills later in a child's life (Franz & Dawson, 2019).

Furthermore, as there is an increasing number of children being diagnosed with ASD (Roman-Urrestarazu et al., 2021), the pressure on educational settings to use interventions which have been shown to be successful in this group has increased (Ali & Frederickson, 2006). Therefore, it is crucially relevant for EPs to know joint attention interventions for those with ASD are effective and evidence-based (Greenway, 2000) to appropriately and successfully support settings catering for these individuals.

Review Question:

How effective are joint attention interventions in children aged eighteen months to five years with a diagnosis of Autism Spectrum Disorder?

Critical Review of the Evidence

Literature Search and Screening

Literature searches were conducted on the 19th December 2021 using the online databases: PsycINFO (Ovid), Education Resource Information Centre (ERIC, EBSCO) and Web of Science (EBSCO). The search terms for this literature search are shown in table 1. The search term "Bucket time" was used as this is a named attention intervention for ASD which is part of the Attention Autism approach developed by Gina Davies (n.d.).

The search returned 119 text results (PsycINFO, 35; ERIC, 15; Web of Science, 69), of which 56

were then screened (titles and abstracts) based on the inclusion and exclusion criteria (Table 2) which led to 38 studies being excluded from the review. For the remaining 25, full text screening was conducted with 15 of these studies being excluded (Appendix A). Figure 1 shows a flow diagram to illustrate this process. Table 3 lists the final 10 studies ddeJ dppTbl -ns A

	Factor	Inclusion Criteria	Exclusion Criteria	Rationale
1	Participants	All participants aged 18 months to 5 years	Some or all participants aged 0-17 months, or older than 5 years	This review is looking at outcomes for children aged 18 months to 5 years.
2	Diagnosis	All participants have a diagnosis of Autism Spectrum Disorder	Some or all participants do not have a formal Autism Spectrum Disorder diagnosis or have a diagnosis other than Autism Spectrum Disorder e.g. Attention deficit hyperactivity disorder	

(v)8 (ie)-6 (w)10 (is)-2 ( )]TJETQBT0.00

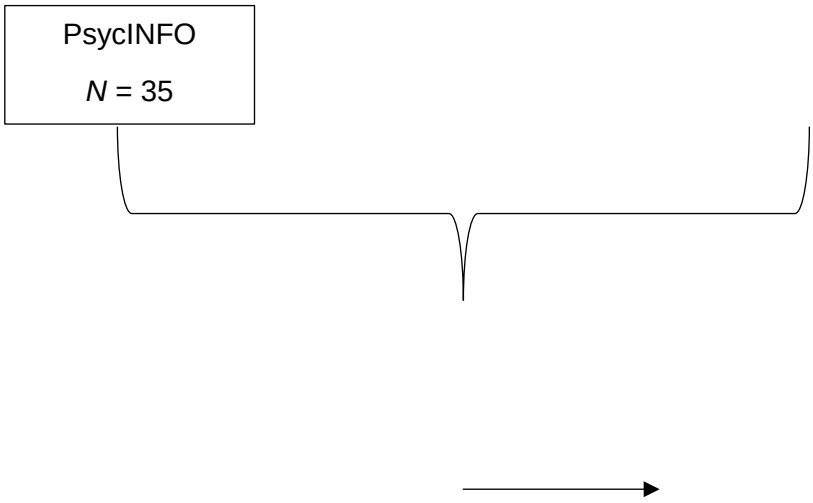
	Factor	Inclusion Criteria	Exclusion Criteria	Rationale
4	Methodology	Quantitative methodology	Qualitative methodology	This review is looking to explore the effectiveness of Attention interventions & quantitative methodology is most appropriate for this purpose.

mnP(at)2 w 124

	Factor	Inclusion Criteria	Exclusion Criteria	Rationale
		Development member countries	operation and Development	countries similar to the UK.
7	Outcome	Joint attention outcomes are reported	Joint attention outcomes are not reported (e.g. focus on spontaneous communication)	This review question is considering the effectiveness of Attention interventions on joint attention.
8	Language	Written in English	Not written in English	The author is monolingual and time/ cost restraints do not allow for translation.

Figure 1

Flow Chart of the Literature Search



Study	WoE A	WoE B	WoE C	WoE D
Gulsrud et al. (2007)	1.6	1	2.2	1.6 (Low)
Kasari et al. (2006)	2.3	2	2.2	2.2 (Medium)
Kasari et al. (2010)	1.7	3	2	2.2 (Medium)
Kasari et al. (2015)	2	1	1.8	1.6 (Low)
Lawton & Kasari (2012a)	1.9	3	2.2	2.4 (Medium)
Lawton & Kasari (2012b)	2.7	2	2.4	2.4 (Medium)
Rocha et al. (2007)	2.7	0	2	1.6 (Low)
Whalen et al. (2003)	2.6	0	2	1.5 (Low)
Zheng et al. (2020)	2.3	2	2.2	2.2 (Medium)

Note. WoE D ratings are described as 'Low' for scores 0-1.7, 'Medium' for scores 1.8-2.4, and 'High' for scores 2.5-3.

Participants

A total of 476 participants, aged two to five years, took part in the included studies. Eight studies reported the mean age of participants and the standard deviations (Boyd et al., 2018; Gulsrud et al., 2007; Kasari et al., 2006; Kasari et al., 2010; Kasari et al., 2015; Lawton & Kasari, 2012a; Lawton & Kasari, 2012b), resulting in a high rating for the WoE C criterion E 'Age of Participants'. The two single case design studies (Rocha et al., 2007; Whalen et al., 2003) reported the age of participants but did not include a mean age which resulted in a medium rating. However, participants in these two studies were within the age range of two-five years. Therefore, the generalisability of the findings from this review to children of preschool age is high, as all participants were also of preschool age.

All ten studies reviewed took place in the USA. They were therefore all deemed relatively applicable to the UK education system, resulting in a medium rating for the WoE C criterion B 'Location'.

Design

Two studies (Kasari et al., 2010; Lawton & Kasari, 2012a), used a randomised control trial (RCT) design which is considered the gold standard for providing evidence on the effectiveness of interventions (Higgins et al., 2011). Therefore, these studies received a 'high' WoE B rating. While four other studies randomised participants to groups and included a control group (Boyd et al., 2018; Kasari et al., 2006; Lawton & Kasari, 2012b; Zheng et al., 2020), they did not report follow-up measures resulting in 'medium' WoE B ratings.

Four studies included a second intervention group involving a symbolic play intervention (Gulsrud et al., 2007; Kasari et al., 2006; Lawton & Kasari, 2012a) and a parent only psychoeducational intervention (Kasari et al, 2015). In contrast, Kasari et al. (2010) and Wheng et al. (2020) implemented a wait list control design.

The other two studies (Rocha et al., 2007; Whalen et al., 2003) both used a single subject, multiple baseline design across participants which has been criticised for lacking external validity as findings can only be confidently related to the included participants and are not generalisable (Engel & Schutt, 2008). As a result, these two studies received a 'very low' WoE B rating. However, single case designs are suitable for use in heterogenous populations (Horner et al., 2005), of which the population of this review (those with ASD) is (Hassan & Mokhtar, 2019). Moreover, Plavnick and Ferreri (2013) argue these designs are particularly appropriate for use in educational research as they lead to greater understanding of who a particular intervention is and is not effective for and why. Furthermore, as these studies were rated 'High' for WoE A 'Methodological Quality' the findings of these studies are of significant relevance to the review in question.

All studies included pre- and post- intervention measures, however follow up data was only included for four studies (Kasari et al., 2010; Lawton & Kasari, 2012a; Rocha et al., 2007; Whalen et al., 2003). As the latter two studies (Kasari et al., 2010; Lawton & Kasari, 2012a) were also RCTs they were awarded the highest rating for WoE B.

Intervention

The majority of studies reviewed used a researcher- developed joint attention intervention (Gulsrud et al, 2007; Kasari et al, 2006; Kasari et al. 2010; Lawton & Kasari, 2012a; Rocha et al., 2007; Whalen et al., 2003). Two studies (Kasari et al., 2015; Lawton & Kasari, 2012b) used JASPER (joint attention symbolic play, engagement and regulation) intervention, while one study used ASAP (advancing social-communication and play) intervention (Watson et al., 2011). In contrast to the other nine studies reviewed, Zheng et al. (2020) used a robot mediated intervention. All ten studies included a joint attention component in the intervention, however in three studies (Boyd et al., 2018; Kasari et al., 2015; Lawton & Kasari, 2012b) this was not the sole focus of the intervention leading to lower WoE C ratings for criteria A 'Intervention'.

Interventions were carried out by a variety of individuals with three studies using trained external professionals only (Kasari et al. 2006; Lawton & Kasari, 2012a; Whalen et al., 2003), and two studies using them in conjunction with caregivers (Kasari et al., 2010; Kasari et al., 2015). Trained caregivers were solely used as interventionalists in Rocha et al. (2007) and two studies used trained professionals from within the setting (Boyd et al., 2018; Lawton & Kasari, 2012b). Finally, Zheng et al. (2020) used a humanoid robot to deliver the intervention.

The average duration of joint attention interventions reviewed was 6.8 weeks, however three studies were excluded from this calculation (Boyd et al., 2018; Rocha et al., 2007; Whalen et al., 2003). The former as the intervention lasted a minimum of six months which would have significantly skewed the average, and the latter two as due to their design the intervention duration

varied between participants. With regards to frequency of intervention, nine studies delivered the intervention more than once a week with four of these (Gulsrud et al., 2007; Kasari et al., 2006; Lawton & Kasari, 2012a; Lawton & Kasari, 2012b) delivering the intervention daily. One study did not provide information on exact intervention frequency (Zheng et al., 2020), only that four sessions were delivered over the course of three-nine weeks.

For further detail regarding the interventions used in the ten reviewed studies, see Appendix B.

Measures

All the studies reviewed had at least one measure of joint attention. Four of the studies (Kasari et al., 2006; Lawton & Kasari, 2012a; Lawton & Kasari, 2012b; Whalen et al. 2003) used the Early Social Communication Scales (ESCS) (Mundy et al., 1996), a structured observational measure. A researched-developed coding system to analyse observations was used by five studies (Boyd et al. 2018; Gulsrud et al., 2007; Kasari et al., 2010; Kasari et al., 2015; Whalen et al., 2003). An adaption of the Unstructured Joint Attention Assessment (UJAA) (Loveland & Landry, 1986) was used by the two single case studies (Rocha et al., 2007; Whalen et al., 2003) and the Screening Tool for Autism in Toddlers and Young Children (STAT) used by one study (Zheng et al., 2020).

Four studies used multiple methods to measure joint attention. Whalen et al.

classroom observation and Zheng et al. (2020) used STAT and a within system (computer) measurement. This is reflected in their higher WoE A ratings for 'Measurement/ Dependent Variable(s)'.

Reliability and validity of measures were also evaluated, with the extent to which these were discussed being reflected in WoE A 'Measurement' ratings.

Findings and Effect Sizes

Only two studied reported effect sizes (Boyd et al., 2018; Kasari et al., 2006). The effect sizes calculated for the other six RCTs was the standardised mean difference (Cohen's d) (see table 4 for descriptors of effect size values). These were calculated from reported F test data or means and standard deviations using the Campbell Collaboration online calculator (Wilson, n.d.). The effect sizes calculated for the two single case design studies (Rocha et al., 2007; Whalen et al., 2003) was Tau-U (baseline corrected) (see table 5 for descriptors of effect size values). For these studies, data for calculating effect sizes was not available in the paper directly. Therefore, WebPlotDigitizer (Rohatgi, 2020) was used to enable means and standard deviations to be calculated through extracting data points from the graphs. This

This suggests interventions had different effects on different measures of joint attention. One study (Kasari et al., 2006) had large effect sizes for all three outcome measures, and one study had large effect sizes at all three times the outcome was measured (Lawton & Kasari, 2012b). Whalen et al. (2003) also had some large effect sizes, however this study had an overall 'Low' rating for WoE due to its design and small sample size therefore these findings should be interpreted with caution.

Two of the included studies (Boyd et al., 2018; Zheng et al., 2020) did not find significant improvement in joint attention following intervention. Boyd et al. (2018) cite intervention implementation issues as a possible reason for this, with Zheng et al. (2020) arguing significant changes in joint attention for subgroups of participants, suggesting it is difficult to establish a clear pattern of response to interventions in this heterogenous group (ASD). The other eight studies reviewed all found improvements in joint attention, with Whalen et al. (2003) and Kasari et al. (2006) finding these improvements generalized from intervention sessions to the natural environment for the former, and to play interactions with a caregiver for the latter. In addition, two studies reported sustained improvements in joint attention at follow up (Kasari et al., 2010; Kasari et al. 2012a; Kasari et al., 2015).

There is a mixed picture with regards to the effects found on different measures of joint attention. For example, Kasari et al. (2015) found larger improvements for duration of joint engagement compared to joint attention initiations. Whereas, Gulsrud et al. (2007) found larger effects for quantity of joint attention than duration of joint attention. Such differences could again be attributable to the heterogeneity of the population being studied. Moreover,

there appear no clear difference in study findings between different settings such as preschool classrooms or Early Intervention Program centres. However, the two studies (Boyd et al., 2018; Lawton & Kasari, 2012b) conducted in educational settings did receive higher WoE ratings for criterion L ‘Site of Implementation’ due to their higher ecological validity.

Study	Sample Size	Research Design	Outcome Measure	Effect Size	*p*	Descriptor	WoE D
			Coordinated joint looks (Early Social Communication Scales)	*d*= 1.32	>0.05	Large	
			Child-initiated Joint Engagement (Early Social Communication Scales)	*d*= 1.38	>0.05	Large	
Kasari et al. (2010)	38	Randomized wait list control study	Joint Engagement (Early Social Communication Scales)	*d*= 0.58	>0.05	Medium	2.2
			Responsiveness to Joint Attention	*d*= 0.66	>0.05	Medium	
			Joint Engagement at 1y follow up	*d*= 0.56	>0.05	Medium	
			Responsiveness to Joint Attention at 1y follow up	*d*= 0.11	>0.05	Small	
	86	Randomized comparative					

Study	Sample Size	Research Design	Outcome Measure	Effect Size	*p*

Study	Sample Size	Research Design	Outcome Measure	Effect Size	*p*	Descriptor	WoE D
				Tau-U= -0.706			
Whalen et al. (2003)	10	Single subject, multiple baseline design across participants	Joint Attention responses (Early Social Communication Scales & Unstructured Joint Attention Assessment)	Carrie: Tau-U= 0.319			

Study	Sample Size	Research Design

Study

Conclusions & Recommendations

Summary

This review updated and refined a previous systematic review and aimed to investigate the effectiveness of joint attention interventions in preschool aged children with ASD. Of the ten studies reviewed, six received a medium WoE D rating, and four received a low rating (Gulsrud et al., 2007; Kasari et al., 2015; Rocha et al., 2007; Whalen et al., 2003;). All studies measured joint attention using observational measures, with one study also using a computerised measure (Zheng et al., 2020).

Overall, evidence for the effectiveness of joint attention interventions in this population appears moderate and mixed with two studies finding no effects (Boyd et al., 2018; Zheng et al., 2020) and the other eight studies finding small to large effects. This demonstrates the significant variability in the effects of joint attention interventions in the ten studies reviewed. Furthermore, the studies reviewed varied in the type of joint attention evaluated, the setting in which the intervention was delivered and who by and the joint attention outcome/s measured. Despite this variation, all ten studies reviewed had a common aim of increasing joint attention in young children with ASD.

Limitations

The ten studies reviewed varied significantly, making it difficult to determine the overall effectiveness of joint attention interventions evaluated in this review. Furthermore, four studies did not include a control group (Gulsrud et al., 2007; Kasari et al., 2006; Kasari et al, 2015; Lawton & Kasari, 2012a) but

instead included a second intervention group meaning changes cannot be solely attributed to joint attention intervention and may be due to other factors (Gopalan et al., 2020). In addition, two studies adopted a single case design (Rocha et al., 2007; Whalen et al., 2003) which has been criticised for producing findings which cannot be generalised to wider populations (Engel & Schutt, 2008). However, these were included in the review as they have been described as suitable for use in heterogenous populations (Horner et al., 2005) so are appropriate when exploring the ASD population as this is a heterogenous group (Hassan & Mokhtar, 2019). Furthermore, such designs are of relevance for educational research (Plavnick & Ferreri, 2013).

Two studies reviewed (Kasari et al., 2015; Lawton & Kasari, 2012b) used the JASPER intervention developed by Kasari and colleagues. Therefore, their findings should be viewed with caution as the researchers may have been biased to produce positive results that support the intervention they developed being effective.

All studies reviewed used observational methods, however many lacked triangulation through the use of multiple measures which led to WoE A penalties by lowering reliability and validity (Annan et al,. 2013; Moon, 2019). Moreover, only five of the ten studies reviewed included follow up data (Kasari et al., 2010; Kasari et al., 2015; Lawton & Kasari, 2012a; Rocha et al., 2007; Whalen et al., 2003). As interventions have been associated with later improvements in social communication (Howard et al., 2005; McEachin, Smith & Lovaas, 1993; Sallows & Graupner, 2005), ascertaining whether improvements are maintained over time is important. Thus, future research

could explore the effects of joint attention interventions over time and use multiple methods to monitor impact on outcomes.

With regards to specificity, five studies targeted responses to joint attention bids only (Gulsrud et al., 2007; Kasari et al., 2010; Lawton & Kasari, 2012a; Rocha et al., 2007, Zheng et al., 2020), while the other five targeted both responses to and initiations of joint attention (Boyd et al., 2018; Kasari et al., 2006; Kasari et al., 2015; Lawton & Kasari, 2012b; Whalen et al., 2003). Therefore, as interventions focused on different aspects of joint attention and found differing effects, further research is needed to explore if these behaviours can be effectively targeted simultaneously or if specific separate intervention is needed.

Finally, all ten studies reviewed were conducted in the USA, so research is needed in the UK setting to assess whether similar effects are seen within the UK population and educational settings.

Implications for EP practice

EPs should work in collaboration with not only children and young people but the adults who care for them (Gutkin & Curtis, 2009) as such collaboration increases the likelihood of interventions being implemented and sustained (Reynolds et al., 2017). Three of the studies reviewed involved parents implementing the intervention (Kasari et al., 2010; Kasari et al., 2015; Rocha et al., 2007), with small to large effects being found. This suggests, parents could be effective agents to improving joint attention skills in young children with ASD. This is of importance for EPs when considering interventions which are collaborative and feasible. Moreover, parental involvement has

Higgins, J., Altman, D., Gotzsche, P., Juni, P., Moher, D., Oxman, A., Savovic, J., Schulz, K., Weeks, L., Sterne, J. , Cochrane Bias Methods Group & Cochrane Statistical Methods Group (2011). *BMJ, 343*, 1-9.

Horner, R., Carr, E., Halle, J., McGee, G., Odom, S. & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional children, 71*(2), 165-179.

Howard, J., Sparkman, C., Cohen, H., Green, G. & Stanislaw, H. (2005). Comparisons of intensive behavior analytic and eclectic treatments for young children with autism.

Loveland, K. & Landry, S. (1986). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders, 16*, 335-349.

McEachin, J., Smith, T. & Lovaas, O. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment.

Sigman, M. & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, Down syndrome and developmental delays. *Monographs of the Society for Research in Child Development, 64,* 1-114.

Tarlow, K. (2016). Baseline Corrected Tau Calculator. Retrieved January 28, 2021, from http://ktarlow.com/stats/tau/

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Watson, L., Boyd, B., Baranek, G. & Crais, E. (2011). *Advancing social-communication and play: An intervention program for pre-schoolers with autism manuals.* The University of North Carolina.

Wh is/TT2 O’yBre0e (e0e (e0)10 (/TTM 0 Tw 9.34 0 Td( )Tj)10.1(s)4 (,)2 (0. is)4-5CID 6 >

Appendices

Appendix A - Excluded Studies

List of Excluded studies at full review

Reference	Criteria number	Rationale
Alotaibi, A. (2020). The effect of teacher implemented Joint Attention intervention on		

Reference	Criteria number	Rationale
nt Attention Training With Autism Spectrum Disorders. *Autism, 20*(2), 172-182.	1	Organisation for Economic Co-operation and Development: Taiwan.
Eissa, M. (2015). The Effectiveness Of A Joint Attention Training Program On Improving Communication Skills Of Children With Autism Spectrum Disorder. *International Journal of Psycho-Educational Sciences, 4*(3), 3-12.		

Reference	Criteria number	Rationale
Intervention in the Preschool Classroom. *The Journal of Special Education, 53*(2), 96-107.		Autism Spectrum Disorder.
Jones, E., Carr, E. & Feeley, K. (2006). Multiple Effects of Joint Attention Intervention for Children with Autism. *Behaviour Modification, 30*(6), 782-834.	2	Some participants did not have formal diagnosis of Autism Spectrum Disorder.
Jones, E., Feeley, K. (2007). Parent Implemented Joint Attention Intervention for Preschoolers with Autism. *The Journal of Speech and Language Pathology – Applied Behavior Analysis, 2*(3), 253-268.	5	Article published in non-peer reviewed journal.
Kaale, A., Smith, L. & Sponheim, E. (2012). A randomized controlled trial of preschool-based joint attention intervention for children with autism. *Journal of Child Psychology and Psychiatry, 53*(1), 97-105.	6	Study conducted country that is not member of Organisation for Economic Co-operation and Development: Norway.

Reference	Criteria number	Rationale
		measured & reported for one of two studies described.
Whalen, C., Schreibman, L. & Ingersoll, B. (2006). The Collateral Effects of Joint Attention Training on Social Initiations, Positive Affect, Imitation, and Spontaneous Speech for Young Children with Autism. *Journal of Autism and Developmental Disorders, 36,* 655-664.	7	Joint attention outcomes not reported.
Wong, C. (2013). A play and joint attention intervention for teachers of young children with autism: A randomized controlled pilot study. *Autism, 17*(3), 340-357.	1	Some participants aged 6 years and above.

	Author & Location	Design	N	Participants	Intervention	Interventionalist & Setting	Duration	Outcome & Measures
2	Gulsrud et al. (2007) USA	Randomized controlled intervention study	35	2-4-				

Author & Location	Design	N	Participants	Intervention

Author & Location	Design	N	Participants	Intervention	Intervention Setting	Intervention Delivery	Outcome Measures

	Author & Location	Design	N	Participants	Intervention	Interventionalist & Setting	Duration	Outcome & Measures
					maintaining periods of joint engagement. Involved 2 30-minute sessions a week for 10 weeks.			pre, post & 6m follow-up
6	Lawton & Kasari (2012a) USA	Randomized controlled intervention study	52	3–4-year-olds with clinical diagnosis of ASD	Joint attention using discrete trial training (with a hierarchy of different prompts and positive reinforcement). Intervention was based on applied behaviour	Trained Educational Psychology graduate students (experienced with children with ASD) in Early	5-6 weeks	Joint attention and shared positive affect at pre, post & 6m follow-up

	Author & Location	Design	N	Participants	Intervention	Interventionalist & Setting	Duration	Outcome & Measures
					analysis and developmental approach of responsive and facilitative interaction.	intervention program centre		
7	Lawton & Kasari (2012b) USA	Randomized controlled trial	16	3–5-28.08 12 riaam	10			
				12 -f3-2.1TEM (er)7 ((r)7 pons)4a and vapprovre				

	Author & Location	Design	N	Participants	Intervention	Interventionalist & Setting	Duration	Outcome & Measures
					joint engagement. Involved 2 30-minute sessions a week for 5 weeks.			engagement at pre & post
8	Rocha et al. (2007) USA	Single subject, multiple baseline design across participant pairs	3	2-4 year olds with clinical diagnosis of ASD	Joint attention training for parents using behaviour analytic techniques to increase parents joint attention initiations and subsequently responses in children.	Trained parents in playroom in clinic & child's home	N/A	Responses to joint attention at pre, post & follow-up

Author

&

Location

Appendix C - Weight of Evidence

Weight of Evidence A- Methodological Quality

WoE A assessed the methodological quality of studies. The group design coding protocol from Kratochwill (2003) that has been used in this review for the eight studies which adopted a group design was amended. The amendments and rationale are detailed in Table 1. The single case design coding protocol from Horner et al. (2005) was used to review the other two studies which adopt this form of research design. The WoE A ratings for the eight group design studies are shown in Table 2 and for the two single case designs in Table 3.

Table 1

Amendments to the group design coding protocol

Section heading		Section removed	Rationale
I.	General Study Characteristics	A: General Study Characteristics	This is discussed in detail in the review.
		B: General Design Characteristics	This is discussed in detail in the review.
		C: Data Analysis	This is not relevant for the current review.
		D: Type of Program	All studies included in this review are intervention programmes.

Section heading	Section removed	Rationale
	E: Stage of Program	This is not relevant for the current review.
	F: Concurrent or Historical Intervention Exposure	This is not relevant for the current review.
II. Key Features for Coding Studies and Rating Level of		

Section heading	Section removed	Rationale
	J4.9 Intervention Style or Orientation	This is discussed in detail in the review.
	J4.10 Cost Analysis Data	This is not relevant for the current review.
	J4.11 Training and Support Resources	This is discussed in detail in the review.
	J4.12 Feasibility	This is discussed in detail in the review.
	K: Replication	This is not relevant for the current review.

Study	A: Research Methodology	B: Measurement	G1: Sampling Procedures	G3-6: External Validity	J1-3: Implementation Fidelity	J4: Identifiable Intervention Components	L: Site of Implementation	Overall WoE A rating (average)
Kasari et al. (2010)	2	2	2	2	2	1	1	1.7
Kasari et al. (2015)	3	2	1	3	2	2	1	2
Lawton & Kasari (2012a)	2	2	2	2	2	2	1	1.9
Lawton &	3	3	2	3	3	2	3	2.7

Table 3

Summary of WoE A Ratings for Single-participant design studies

Study	Description of participants & settings	Dependent variable(s)

Weight of Evidence B: Methodological Relevance

WoE B assesses how appropriate the type of study is in relation to the review question being addressed. Petticrew and Roberts (2003) propose a Hierarchy of Evidence which assigns Randomised Control Trials (RCT) as the 'gold standard' of designs measuring the effectiveness of interventions. This is followed by quasi-experimental and cohort studies being of lower quality in addressing such questions. Criteria to evaluate WoE B were developed based on Petticrew and Roberts (2003) recommendations as to the research most suitable for answering these types of questions, with each study being assigned a rating of 0-3 as shown in Table 4 and Table 5.

Table 4

WoE B Criteria

Rating	Criteria	Example
3 (High)	• Included a control group • Random assignment to treatment or control groups • Pre- and post- intervention and follow up measures reported	RCT
2 (Medium)	• Included a control group • Pre- and post- intervention measures reported	Quasi-experimental designs with a control group
1 (Low)	• Did not include a control group	Quasi-experimental designs without a

Rating	Criteria	Example
0 (Very Low)	• Pre- and post- intervention measures reported • Did not include a control group • No random assignment • Pre- and post- intervention measures may or may not be reported	control group & cohort studies Qualitative research, case-control studies, surveys & non-experimental evaluations

Weight of Evidence C: Topic Relevance

WoE C is a review-specific judgement about the relevance of the focus of the study to the review question being explored. The criteria in Table 6 were developed and each study was assigned a rating of 0-3 for each of the five criteria identified. These ratings were then averaged to produce an overall WoE C rating as shown in Table 7.

Table 6

WoE C Criteria and Rationale

Criteria	Rating	Descriptor	Rationale
A: Intervention	3	Joint Attention is the only focus of the intervention.	This review is looking at the effectiveness of interventions targeting joint attention. Therefore, interventions focusing on other areas are not suitable.
	2	Joint Attention is one of the foci of the intervention.	
	1	Joint Attention is not the focus of intervention.	
B: Location	3	Study is conducted in the UK.	To increase the generalisability of the findings to the UK, it is important the study has taken place in a location

Criteria	Rating	Descriptor	Rationale
	1	Study is conducted in a location which is not economically similar to the UK.	with a comparable education system.
C: Intervention description	3	The intervention is clearly described, with accompanying materials provided.	To allow for the intervention to be replicated, information regarding the content and implementation of the intervention should be given.
	2	The intervention is clearly described, but no accompanying materials are provided.	
	1	The intervention is not clearly described and no accompanying materials are provided.	
D: Intervention delivery	3	The intervention is delivered by existing staff working in the educational setting.	Findings are higher in external validity in studies where existing staff working in the educational setting e.g. school, have delivered the intervention.
	2	The intervention is delivered by researchers in an educational setting.	

Criteria	Rating	Descriptor	Rationale
	1	The intervention is delivered by researchers outside an educational setti6p39g æt.0g	

Table 7

Summary of WoE C Ratings

Study	Criteria A	Criteria B	Criteria C	Criteria D	Criteria E	Overall WoE C Rating
Boyd et al. (2018)	2	2	2	3	3	2.4 (Medium)
Gulsrud et al. (2007)	3	2	2	1	3	2.2 (Medium)
Kasari et al. (2006)	3	2	2	1	3	2.2 (Medium)
Kasari et al. (2010)	3	2	1	1	3	2 (Medium)
Kasari et al. (2015)	2	2	1	1	3	1.8 (Medium)
Lawton & Kasari (2012a)	3	2	2	1	3	2.2 (Medium)

Lawton & Kasari (2012b)	2	2	2	3	3	2.4 (Medium)

G. External Validity Indicators

G1. Sampling Procedures

G1.1 Sampling procedures described in detail

☒ 1 Yes

☐ 0 No

G1.2 Rationale for sample selection specified

☐ 1 Yes
Specify:

☒ 0 No

G1.3 Rationale for sample size specified

☐ 1 Yes
Specify:

☒ 0 No

G1.4 Evidence provided that sample represents target population

☒ 1 Yes

☐ 0 No

G1.5 Recruitment procedures congruent with target cultural group. Researcher used culturally appropriate ways/methods to contact, recruit, inform, and maintain participation.

☒ 1 Yes

☐ 0 No

G1.6 Inclusion/exclusion criteria specified

☒ 1 Yes

☐ 0 No

G1.7 Inclusion/exclusion criteria similar to school practice

☐ 1 Yes

☐ 0 No

G

☐ Unknown

Rating for Implementation fidelity 2

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

J4. Implementation Context (Conditions of Implementation)

J4.2 Adaptations in Implementation

☐ 3 Detailed account of the implementation and adaptations to fit the context or target population

☐ 2 Detailed account of the implementation but not of the adaptations to fit the context or target population

☐ 1 Partial description of the implementation and/or the adaptations to fit the context or target population

☐ 0 Vague or no account of the implementation

J4.3 Relationship of Researcher to Intervention

☐ 3 Detailed description of the researcher's level of involvement and safeguards used to minimize the bias of the researcher.

☐ 2 Detailed description of the researcher's level of involvement, but minimal description of safeguards to minimize the bias of the researcher

☐ 1 Minimal description of the researcher's level of involvement and of safeguards to minimize the bias of the researcher.

☐ 1 No information provided

J4.4 Relationship of Implementer/to Participants

☐ 3 Detailed description regarding the interpersonal processes used to establish and maintain the relationship between implementer and participants.

☐ 2 Detailed description of relationship development procedures, but lacking detail on some aspects of the relationship processes.

☐ 1 Provides overview of relationship development procedures and processes, but lack details

☐ 0 No description of relationship processes provided.

J. Overall Rating for Identifiable Intervention Components 2

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

L. Site of Implementation

L1. School (if school is the site, select one of the following options)

L1.1 ☐ Public

L1.2 ☐ Private

L1.3 ☐ Charter

L1.4 ☐ University Affiliated

L1.5 ☐ Alternative

L1.6 ☐ Not specified/ unknown

L2. Non School Site (if it is a non school site, select one of the following options)

L2.1 ☐ Home

L2.2 ☐ University Clinic

L2.3 ☐ Summer Program

L2.4 ☐ Outpatient Hospital

L2.5 ☐ Partial inpatient/ day Intervention Program

L2.6 ☐ Inpatient Hospital

L2.7 ☐ Private Practice

L2.8 ☐ Mental Health Center

L2.9 ☐ Residential Treatment Facility

L2.10 ☐ Other (specify): _____

L2.11 ☐ Unknown/ insufficient information provided

J. OVERALL Rating for Site of Implementation (select 0, 1, 2 or 3): 3

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

Summary of Evidence for Group- Based Design Studies

Indicator	Overall Evidence Rating (0 -3)	Description of Evidence (Strong, Promising, Weak or No/limited evidence)
Key Features		
Research Methodology	3	Strong
Measurement	2	Promising
Sampling	2	Promising
External Validity	2	Promising
Implementation Fidelity	2	Promising
Identifiable Intervention Components	2	Promising
Site of Implementation	3	Strong
Average	2.3	Promising

Appendix E- Example of a completed Weight of Evidence A coding protocol f or one single -case design stud y

Reference: Rocha, M., Schreibman, L. & Stahmer, A. (2007). Effectiveness of Training Parents to Teach Joint Attention in Children With Autism. *Journal of Early Intervention, 29*(2), 154-172.

Horner et al. (2005): Quality Indicators Within Single -Subject Research

Scoring criteria:

- All criteria fulfilled = 3
- Majority of criteria fulfilled= 2
- Half or less of criteria fulfilled= 1

Description of Participants and Settings	Participants are described with sufficient detail to allow others to select individuals with similar characteristics e.g. age, gender, disability, diagnosis	X
	The process for selecting participants is described with replicable precision	
	Critical features of the physical setting are described with sufficient precision to allow replications	X
Total		2
Dependent Variable	Dependent variables are described with operational precision	X
	Each dependent variable is measured with a procedure that generates a quantifiable index	X
	Measurement of the dependent variable is valid and described with replicable precision	X
	Dependent variables are measured repeatedly over time	X
	Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards e.g. IOA= 80%; Kappa= 60%	X
Total		3

Independent Variable	Independent variable is described with replicable precision	X
	Independent variable is systematically manipulated and under the control of the experimenter	X
	Overt measurement of the fidelity of implementation for the independent variable is highly desirable	X
Total		3
Baseline	Baseline phase provides repeated measurement of the dependent variable	X
	The baseline establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur	X
	Baseline conditions are described with replicable precision	X
Total		3
Experimental control/ Internal validity	The design provides at least 3 demonstrations of experimental effect at 3 different points in time	X
	The design controls for common threats to internal validity e.g. permits elimination of rival hypotheses	X
	The results document a pattern that demonstrates experimental control	X
Total		3
External validity	Experimental effects are replicated across participants	X
	Experimental effects are replicated across settings	X
	Experimental effects are replicated across materials	
Total		2
Social validity	The dependent variable is socially important	X
	The magnitude of change in the dependent variable resulting from the intervention is socially important	X