

Case study 1: An Evidence-based practice review report.

Theme: School/Setting Based Interventions for Social, Emotional and Mental Health.

How effective is the peer-mediated intervention, Stay, Play, Talk, in improving social communication skills for preschoolers with identified social communication difficulties?

Summary

Social communication is a fundamental life skill that enables an individual to understand and use appropriate conversational skills, to communicate effectively with others and develop meaningful relationships. Preschool settings provide a safe environment to practice and develop these skills during early childhood, however, children with social communication difficulties may not respond to opportunities for social learning in the same way as their typically developing peers. *Stay, Play, Talk* is a peer-mediated intervention which involves training peers to stay, play and talk with the target child, who presents with social communication difficulties. This review aims to assess the efficacy of Stay, Play, Talk on preschoolers with identified social communication needs. Five studies were selected using the pre-determined inclusion and exclusion criteria and their weight of evidence was assessed. Two studies displayed a significant difference in the target child's social communicative behaviour from baseline to treatment conditions and during a maintenance phase. These studies included target children with less severe social communication difficulties and the peer groupings had been carefully considered. Due to the inconsistency of the results and heterogeneity of the studies, it is difficult to directly compare these findings. Potential recommendations for future educational practice are discussed, alongside further research suggestions.

Introduction

Social Communication

Social communication, or pragmatic language, can be defined as the way in which one uses language within a social context (Norbury, 2014). The development of social communication skills is a primary focus in early years, as these skills provide the necessary foundations to access the curriculum (Katz & Girolametto, 2013). These skills are also strongly related to positive outcomes in future academic performance (Malecki & Elliott, 2002) and emotional wellbeing (Mashburn et al., 2008). Consequently, preschools have a responsibility to nurture the early development of social communication skills, to improve the educational outcomes and emotional wellbeing for all children.

However, for children with social communication difficulties, regular social-based activities within school can become a persistent barrier. A distinguishing feature of Autism Spectrum Disorders (ASD) is a deficit in social communication (Wing & Gould, 1979). This can include verbal communication, e.g. initiating and responding to interactions, and non-verbal communication, e.g. eye contact (Cashin et al., 2009). Social communication difficulties are also prevalent amongst other disabilities, including Down syndrome and developmental delay (Guralnick et al., 2011; Petersen et al., 1998).

Inclusive Education

In the UK there is a growing number of inclusive provisions, where educational settings cater for the needs of all children (UNESCO, 2009). It is essential for schools to understand the needs of all

their pupils, as they have a statutory responsibility to make reasonable adjustments to avoid students from being unfairly disadvantaged (DfE & DH; 2014; 2015).

While high quality inclusive environments provide opportunities for typically developing children and children with disabilities to interact (Strain & Bovey, 2011), simply integrating these children is not sufficient for improving those positive interactions (Odom et al., 2011). Social interactions of children with disabilities can be significantly lower than those of typically developing children (Honig & McCarron, 1988). Therefore, it is important to acknowledge that not all children respond to social learning in the same way, so many will require additional intervention to develop these fundamental skills.

Peer-mediated Intervention

Peer-mediated intervention is one strategy that can be used within an integrated setting to support the social engagement of target pupils who have social communication difficulties (Kalfus, 1984). During peer-mediated intervention, typically developing children become the intervention agent and are taught specific strategies to engage with a target child. Preschool settings provide a conducive environment for peer-mediators to support target pupils, through the naturally occurring structure of the day.

Stay, Play, Talk (SPT) is an example of a peer-mediated intervention designed to support the social interactions of preschool children (Ledford et al., 2016). This intervention is centred around three fundamental components: (i) stay close to your partner, (ii) play with your partner, and (iii) talk to your partner. Typically developing peer buddies, who have age-appropriate play and social skills, receive regular training on these three elements. The target child is selected because of their difficulties in socialising, in comparison to their peers. The peer groupings can consist of one peer buddy and a target child (dyad), two peer buddies per target child (triad), or alternatively, the whole class can be taught the fundamental principles of SPT. Once these groupings have been established and the training applied, the SPT activities can occur during the children's natural play. Visuals and verbal prompts to remind the peer buddies of their training can be used to reinforce the

intervention. Rewards are also recommended to acknowledge when a skill has been demonstrated successfully.

Psychological Theory

Peer-mediated intervention is grounded in the Social Learning Theory (Bandura, 1977), which emphasises the importance of observing, modelling and imitating the behaviours of others. Peer-mediated interventions encourage typically developing peers to behave as social models to improve the target child's social interactions (Chang & Locke, 2016). This theory posits that children can learn through attending to an individual performing the desired behaviour, rehearsing that skill, and replicating this within different contexts. Learning skills directly from peers also eliminates the need to generalise the skills from adult partners, which is often a challenge for children with ASD (Rogers, 2000).

Regular SPT mediations will encourage the target child to observe and imitate their peer buddies, improving the development of social skills and their confidence to interact with others (Gresham et al., 2010). Imitations are more likely to occur if the child perceives the model to be similar to themselves and if the behaviour is positively reinforced (McLeod, 2016). Therefore, peer buddies are well placed to act as social models for the target child, whilst the preschool educators can positively reinforce this behaviour, through the use of praise and rewards.

Relevance to Educational Psychologist Practice

SPT offers a simple framework to support the social communication between typically developing peers and children with social communication difficulties. It empowers peer buddies to support the target child, which requires minimal training and adult prompting. SPT strategies also take the pressure away from the target child, as they do not require responses or intense involvement (Goldstein et al., 1997). This approach not only supports the development of social communication skills for target children, but also promotes an inclusive ethos within the setting, through emphasising the need to engage with all children. Furthermore, it is particularly important to implement interventions to support the needs of children in the early years, as failure to do so can lead to low self-esteem, frustration in learning and subsequent behavioural difficulties (DfE, 2014).

If effective, SPT offers a feasible early-intervention to support the target child's social communication needs, whilst promoting an inclusive preschool ethos. This would provide an impactful recommendation within Educational Psychologist (EP) practice in the UK.

Current Review

This review will focus on the implementation of SPT in preschools, to determine the efficacy of using this intervention to improve the social communication difficulties for target children. Although SPT teaches peer buddies the valuable skills of being able to interact and engage with children who exhibit social difficulties, this review will focus specifically on the social communicative development of the target child, as they are most at risk of social isolation and academic failure (Barber et al., 2015).

Review Question:

How effective is the peer-mediated intervention, Stay, Play, Talk, in improving social communication skills for preschoolers with identified social communication difficulties?

Critical Review of the Evidence

Systematic Literature Search

A systematic search of the literature was conducted using the online databases: PsycINFO (Ovid), Education Resource Information Centre (ERIC, EBSCO) and Web of Science (EBSCO). A scoping search on Google Scholar identified the commonly used language within the identified research topic. These concepts were combined to create a comprehensive literature search, presented in Table 1.

From conducting a literature search in all three databases and removing the duplicates, 115 studies were identified. These were screened from their titles and abstracts based on the inclusion and exclusion criteria, as specified in Table 2. Where studies did not meet the criteria, they were excluded from the review. From the ten studies that remained, each text was fully screened and assessed using the criteria in Table 2. Five of these studies were excluded, for further details refer to Appendix A. Figure 1 shows a flow diagram to illustrate this process.

Table 1

Literature search terms

Intervention	Participants	Outcome measure
“Stay play talk” OR “Stay, play, talk” OR “Stay-play-talk” OR “Stay, Play & talk” OR “Stay, play and talk” OR “SPT”	Child* OR “Primary school” OR “Pre-school setting*” OR “Preschool*” OR Infant* OR “Autism*” OR “Autism Spectrum Disorder*” OR “ASD” OR “Autism Spectrum Condition*” OR “ASC” OR “PWA” OR “Social delay”	“Social-emotional skill*” OR “Social skill*” OR “Social development” OR Social OR “Social* interaction*” OR play*

Note. Quotation marks used to ensure all words in the phrase are included. An Asterix allows for truncated words with alternative endings. Each column was combined with ‘AND’.

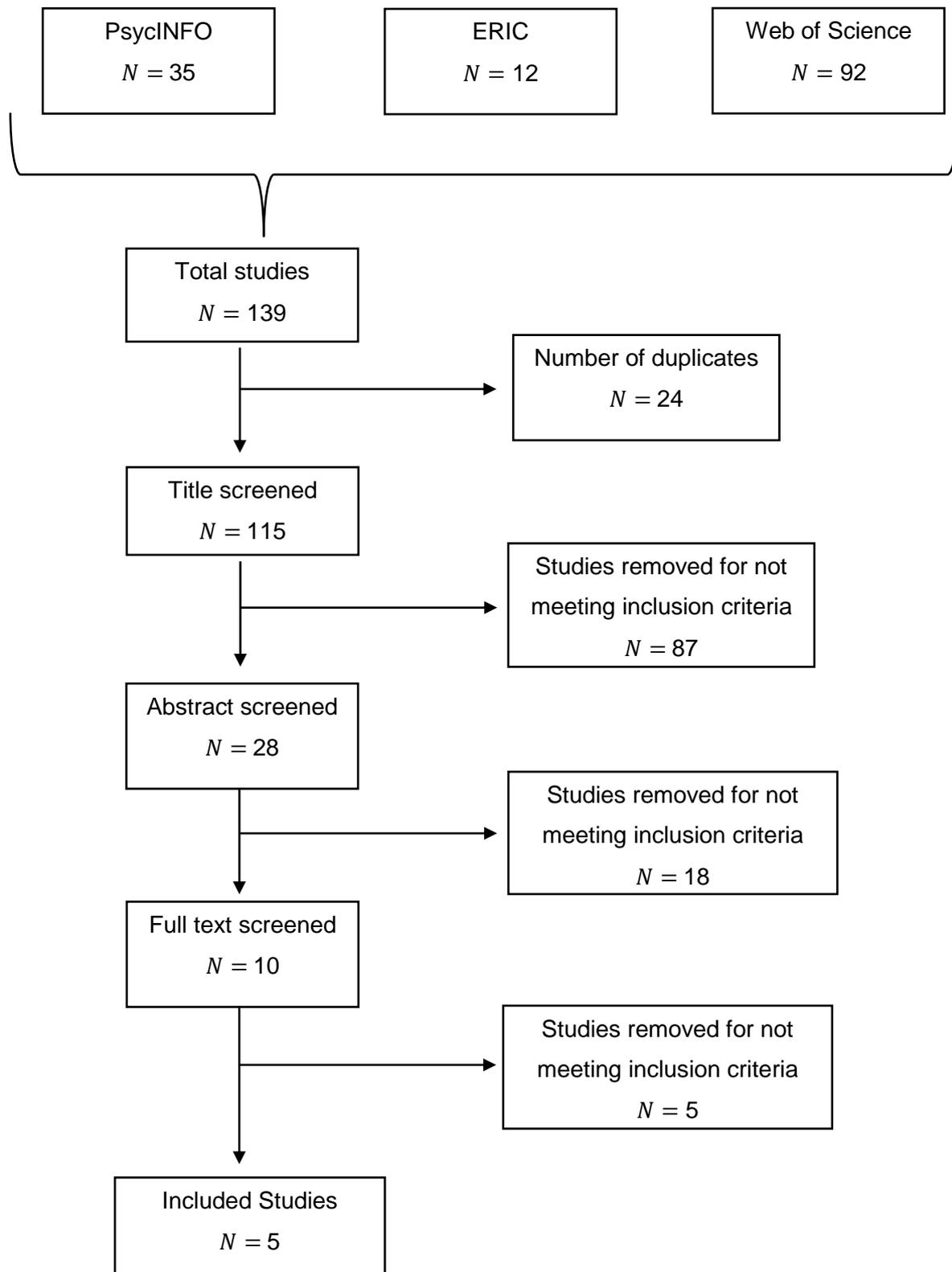
Table 2*Inclusion and Exclusion Criteria*

	Feature	Inclusion Criteria	Exclusion Criteria	Rationale
1	Target child and peer buddies' age	Preschoolers (aged between 1.5 – 6 yrs)	Children >6 yrs & <1.5 yrs	Stay, Play, Talk is an intervention designed for preschool children, who are between 1.5 – 6 yrs old.
2	Target child's criteria	Presenting social-communication difficulty or diagnosis, including Autism and Down's syndrome	Age-appropriate social skills	The intervention is targeted for children with social communication difficulties. This includes children with a variety of disabilities (Ledford et al., 2016).
3	Peer buddies' criteria	Age-appropriate social skills	Presenting social communication difficulty or diagnosis	The peers are trained on how to help their target child with their social skills, so need to have the appropriate skills and abilities to do so.
4	Setting	Preschool setting, including: University development centres; integrated preschools; child care centres; early childhood programs	The home environment; primary school; secondary school etc.	Review question focussed on implementing SPT in a preschool educational setting.

5	Study design	Randomised control trials and single case experimental designs with multiple time points.	All other design types, or single case experimental designs with only a pre and post measure.	According to Petticrew and Roberts (2003), randomised control trials produce the highest methodological rigour when investigating efficacy of an intervention. Single case experimental designs are also necessary to draw out inferences from interventions for low incidence study populations (Tate et al., 2016).
6	Intervention	Stay, Play, Talk intervention. This involves training peers on these core features.	An intervention that is not specified as Stay, Play, Talk, or does not involve the peer training.	The research question is focussed on the effectiveness of the intervention Stay, Play, Talk. Training the peer confederates is an essential element to this programme.
7	Outcome measures	Social communicative outcomes for target child.	Measures other than social communication, or does not specify the outcomes for the target child.	The focus of the review question is on social communication skills for target pupils.
8	Language of report	Reported in English	Reported in a language other than English.	Author's first language is English.
9	Publication	Peer-reviewed journals	Research that has been published in sources other than peer-reviewed journals.	Peer-reviewed journals have been scrutinised by experts within the field, so encourages high quality research.

Figure 1

Flow Chart of the Literature Search



Included Studies

From the 139 studies that were screened using the inclusion and exclusion criteria, only single case experimental designs were found. Five studies remained and these were the focus of the current review. Refer to Table 3 for a list of their references. For further details on the included studies in the review, see Appendix B.

Table 3

Reference List of the Five Included Studies in the Review

Reference	
1	Barber, A. B., Saffo, R. W., Gilpin, A. T., Craft, L. D., & Goldstein, H. (2015). Peers as clinicians: Examining the impact of Stay Play Talk on social communication in young preschoolers with autism. <i>Journal of Communication Disorders</i> , 59, 1-15. https://doi.org/10.1016/j.jcomdis.2015.06.009
2	Goldstein, H., English, K., Shafer, K., & Kaczmarek, L. (1997). Interaction among preschoolers with and without disabilities: Effects of across-the-day peer intervention. <i>Journal of Speech, Language, and Hearing Research</i> , 40(1), 33-48. https://doi.org/10.1044/jslhr.4001.33
3	Maich, K., Hall, C. L., Van Rhijn, T. M., & Squires, K. (2018). Investigating Stay, Play, & Talk: A peer-mediated social skills intervention for young children with autism spectrum disorder and other social challenges. <i>Exceptionality Education International</i> , 28(2), 82–104. https://doi.org/10.5206/eei.v28i2.7766
4	Milam, M. E., Hemmeter, M. L., & Barton, E. E. (2020). The effects of systematic instruction on preschoolers' use of Stay-Play-Talk with their peers with social delays. <i>Journal of Early Intervention</i> , 43(1), 80-96. https://doi.org/10.1177/1053815119900253
5	Severini, K. E., Ledford, J. R., Barton, E. E., & Osborne, K. C. (2019). Implementing Stay-Play-Talk with children who use AAC. <i>Topics in Early Childhood Special Education</i> , 38(4), 220-233. https://doi.org/10.1177/0271121418776091

Note. Phase 2 of research by Maich et al. (2018) was excluded due to the research design only having pre and post measures, please see criteria 5 on Table 2.

Critical Appraisal

To appraise the selected studies, Gough's (2007) Weight of Evidence (WoE) framework was used. The three dimensions of weighting include: the methodological quality (WoE A); the methodological relevance (WoE B); and the topic relevance (WoE C). An average of these scores were calculated to provide an overall WoE D.

WoE A is the study's quality of execution in comparison to the quality standards for other studies of this type (Gough, 2007). As all five studies were single-subject research designs, the Horner et al. (2005) coding protocol was employed to calculate WoE A. Each item was given a rating out of three based on how much of the criteria was fulfilled, see Appendix C for completed protocols. Appendix D provides an overview of the WoE A and respective weightings for each study.

To calculate WoE B, Petticrew and Roberts (2003) typology was used to determine the appropriateness of the research design to address the research question. As all studies were single case experimental designs, they met the 'medium' weighting threshold. See Appendix E for further details.

WoE C is a review-specific judgement on the topic's relevance of the focus of the evidence to the review question (Gough, 2007). Studies were assessed according to criteria in Appendix F. An overview of these scores for each study is provided in Appendix G. See Table 4 for the calculated weightings for each study.

Table 4*Weight of Evidence (WoE) ratings for the studies evaluated in this review*

	WoE A	WoE B	WoE C	WoE D
Barber et al. (2016)	2.14	2	2.29	2.14 (Medium)
Goldstein et al. (1997)	2.21	2	2.57	2.26 (Medium)
Maich et al. (2018)	2.19	2	2.14	2.11 (Medium)
Milam et al. (2020)	2.27	2	2.57	2.26 (Medium)
Severini et al. (2019)	2.43	2	2.29	2.24 (Medium)

Note. <1.5 is low; 1.5 – 2.4 is medium; >2.4 is high

Participants and Setting

Nineteen preschoolers, between the ages of 1.5 - 5 years old, participated across the five studies within this literature review. As all studies were single case experimental designs, they required operational descriptions of the participants, setting, and the recruitment process of participants (Wolery & Ezell, 1993).

All five studies included the target child's age and a qualitative summary of their social behaviour and presenting difficulties. The target pupil's identified social communication difficulties included: seven children with Down's syndrome; six children with ASD; three children with social delay; and three children with delayed language development. However, not all studies replicated the same level of detail within their participant descriptions. Only two of the studies included details regarding the target child's ethnicity (Milam et al., 2020; Severini et al., 2019) and three studies clearly stated the instrument used to measure their disability (Barber et al., 2016; Maich et al., 2018; Milam et al., 2020). This was reflected in a higher weighting for question 1a in the WoE A protocol.

There was also variability between the reported recruitment process for participants. Barber et al. (2016) stated ethical approval from the University Review Board and used surveys to recruit

participants, whereas Maich et al. (2018) conducted recruitment via community-based consultants working with individuals with ASD. As not all studies reported the process of selecting their participants with replicable precision, they scored lower on question 1b in the WoE A protocol. Overall, all studies achieved mostly all the criteria within the WoE A 'Description of Participants and Settings' so were given the 'medium' weighting, with two studies meeting the threshold for a 'high' weighting (Maich et al., 2018; Severini et al., 2019).

A description of the intervention setting was also included. Interventions that were conducted in naturalistic preschool settings were awarded a higher WoE C, as the purpose of this review is to support school-based interventions. Three of the studies were conducted in an integrated preschool setting (Goldstein et al., 1997; Maich et al., 2018; Milam et al., 2020), whereas two studies were conducted in researcher contrived environments (Barber et al., 2016; Severini et al., 2019). This is reflected in the variability of the WoE C for 'Intervention Setting', as the three studies conducted in a naturalistic environment received a 'high' weighting, compared to the 'low' and 'medium' weightings for Barber et al. (2016) and Severini et al. (2019) respectively.

As all studies were conducted in OECD countries (U.S. and Canada), they were provided a 'medium' rating for the 'Location of Intervention' in the WoE C.

Research Design

All research studies implemented a single case experimental design, as no randomised control trials were identified to be relevant to the review question from the database search. Small-N designs are integral when investigating evidence-based practice in special education (Horner et al., 2005). They enable researchers to focus on the individual characteristics and study low-incidence population, for example, children with identifiable social communication difficulties. However, Petticrew and Roberts (2003) suggest that randomised control trials produce the highest methodological rigour when investigating the efficacy of an intervention. As all studies used a single case experimental design, they were consistently awarded a 'medium' rating for the WoE B. To determine a functional relationship between manipulation of the intervention and change in target pupil's social communication skills, the design must document three demonstrations of the experimental effect (Horner et al., 2005). This can include three different points in time with a

single participant or across three different participants. All studies included at least three target participants, apart from Severini et al. (2019) where only two target pupils were investigated, so an A-B-A-B withdrawal design was used. As all studies included at least three demonstrations of experimental effect, this is reflected in the 'high' rating for question 5a in the WoE A protocol.

Intervention

The focus of the review question was on the peer-mediated intervention, SPT. The fidelity of SPT is dependent on the peer training, resources, and relationship between target child and peer confederate. These three categories were included in the WoE C criteria.

SPT involves training peer confederates to stay with their peer, play with their peer and talk with their peer. Ledford et al. (2016) recommends regular training that includes peer buddies and target child, so they are both involved in the process. Three of the selected studies (Maich et al., 2018; Milam et al., 2020; Severini et al., 2019) included detailed training with peer buddies and target children so received 'high' weightings, compared to two studies (Barber et al., 2016; Goldstein et al., 1997) where training was only conducted with the peer buddies, so a 'medium' rating was awarded.

Although variations in the training content was not reflected in the WoE, it is important to note that some studies included additional training topics. Sensitivity training was provided for all participants in Barber et al. (2016) and Goldstein et al. (1997) studies. This involves sensitising participants to the different attention-getting and requesting behaviours that target children may use. Similarly, Maich et al. (2018) conducted diversity awareness training for all participants, which is an adult-led activity that focuses on similarities and differences. Where there was an extended school break for participants in the study by Milam et al. (2020), booster training sessions were applied to recap the core themes. Whilst reviewing the effectiveness of SPT, it is important to consider the variation of training that was applied.

For the resources for SPT, Ledford et al. (2016) recommends using play materials in preschool settings; visual reminders of the taught strategies; adult led reinforcements; and data collection tools. As all studies included an exhaustive list of the resources, they scored a 'high' rating in the 'Intervention Resources' section of the WoE C.

SPT is also centred around the relationship between the target child and peer buddy interacting and playing together. Therefore, ensuring the peer groupings are compatible is vital. Goldstein et al. (1997) utilised the results of non-standardised assessments, including play preferences and sociometric data, to establish peer dyads. However, for three of the studies (Maich et al., 2018; Milam et al., 2020; Severini et al., 2019) target child and peer buddies were randomly allocated, irrespective of social history and shared interests. This is reflected in the difference of the WoE C ratings for 'Relationships between Target Child and Peer Confederate'.

Outcome Measures

Although this review question is focussed on social communication skills, a variety of outcome measures were included. All five studies used multiple time points to assess the efficacy of implementing SPT, with two studies (Barber et al., 2016; Maich et al., 2018) including pre and post questionnaires. As the questionnaires were not well-known standardised assessments, there was not the adequate data to calculate the Reliable Change Index (Jacobson & Traux, 1991). The present review, therefore, will focus on the results generated from multiple data points.

Four studies measured social communication skills through the frequency of interactions (Barber et al., 2016; Goldstein et al., 1997; Maich et al., 2018; Severini et al., 2019). However, Milam et al. (2020) operationalised social communication skills as the duration of a target child's social play. See Table 5 for the specific outcome measures of each study.

Alongside the presenting social communicative behaviour in the treatment condition, all participants partook in a generalisation phase. This was to investigate whether any positive outcomes could be generalised following the intervention input or with a different peer buddy. One study conducted a maintenance phase three months after the intervention (Barber et al., 2016), compared to two studies that included a maintenance phase relatively soon after the intervention (Maich et al., 2018; Milam et al., 2020). Studies also varied in how the maintenance phase was initiated and conducted. Maich et al. (2018) ensured there was no adult prompting after the 12-session implementation, whereas Milam et al. (2020) systematically faded the intervention after the peer buddies' strategies were stable. Contrastingly, Goldstein et al. (1997) included a generalisation probe where the target children were reassigned to different peer buddies.

Three studies (Barber et al., 2016; Goldstein et al., 1997; Milam et al., 2020) used at least three different social outcome measures, including a generalisation phase, providing a comprehensive overview for the target's child social communicative behaviour. This was reflected in a 'high' WoE C for the 'Scope of Outcome Measures'.

Results

To investigate the relationship between baseline social-communicative behaviour and treatment social-communicative behaviour, Tau-U effect sizes were calculated using a single case research calculator (Vannest et al., 2016). Two studies (Goldstein et al., 1997; Milam et al., 2020) found a significant difference in social communication skills from baseline to treatment condition.

To assess the efficacy of social communicative outcomes from baseline to the generalisation phase, the single case research calculator was also used to generate these effect sizes. As Severini et al. (2019) implemented an A-B-A-B experimental design, a maintenance phase was calculated between the initial baseline condition and final SPT intervention phase.

For the majority of target children in two studies (Goldstein et al., 1997; Milam et al., 2020), significant effect sizes were calculated from baseline to the generalisation condition. See Table 5 for full details regarding the effect sizes.

Overall, the research revealed inconsistent findings, where participants in only two studies (Goldstein et al., 1997; Milam et al., 2020) displayed significant improvements in their social communication skills following the SPT intervention. All the research was judged to be of a 'medium' quality, as the methodological quality, methodological relevance and topic relevance, met the appropriate criteria. Whilst none of the studies used randomised control trials, they all had the necessary resources to implement SPT, were conducted in OECD countries and reported at least two measures of social outcomes. Therefore, sufficient weight should be given to the effect sizes of all five studies.

Table 5

Effect Sizes to show the difference between the type of social communication on target children from baseline to treatment and baseline to maintenance.

Study	Outcome	Target Child	Baseline V Intervention			Baseline V Maintenance			WoE D
			Tau - U	<i>p</i>	CI (95%)	Tau - U	<i>p</i>	CI (95%)	
Barber et al. (2016)	Combined initiation and response ("social interaction")	A1	0.313	0.269	-0.152<>0.777	-0.29	0.739	-1<>0.671-	2.14
		A2	0.04	0.867	0.469<>0.531	-0.286	0.558	1<>0.563	
		A3	- 0.67	0.005	-0.395<>0.484	0.80	0.206	-1<>0.517	
Goldstein et al. (1997)	Number of interactions per 10-minute sample with class peers	B1	0.84**	<0.001	0.345<>1	1	0.127	-0.283<>1	2.26
		B2	1***	0.002	0.356<>1	1	0.064	-0.059<>1	
		B3	0.99***	0.003	0.347<>1	1	0.157	-0.386<>1	
		B4	0.95***	0.005	0.281<>1	1***	0.025	0.123<>1	
		C1	0.63**	0.018	0.108<>1	1***	0.046	0.020<>1	
		C2	0.89***	0.003	0.308<>1	1***	0.020	0.157<>1	
		C3	0.91***	0.001	0.361<>1	0.96***	0.019	0.158<>1	
		C4	0.80**	0.023	0.110<>1	1***	0.025	0.123<>1	

Maich et al. (2018)	Mean social interactions across 10-minute data collection intervals	D1	0.029	0.935	-0.661<>0.718	0.87	0.053	-0.010<>1	2.11
		D2	0.200	0.570	-0.490<>0.890	1	0.053	-0.012<>1	
		D3	-0.200	0.558	-0.869<>0.469	0.11	0.847	-0.912<>1	
Milam et al. (2020)	Duration of social play (mean number of seconds)	E1	1***	0.007	0.278<>1	1***	0.017	0.181<>1	2.26
		E2	0.978***	<0.001	0.444<>1	1***	0.006	0.294<>1	
		E3	0.924***	<0.001	0.512<>1	0.96***	<0.001	0.452<>1	
Severini et al. (2019)	Number of social interactions between PB and TC	F1	0.475	0.164	-0.194<>1	0.58	0.071	-0.049<>1	2.24
		F2	-1	0.317	-1<>0.960	-0.83	0.211	-1<>0.473	

Note. Starred values are considered significant with a p -value < 0.05 . *small effect size 0 – 0.31; **medium effect size 0.32 – 0.84; ***large effect size 0.85 – 1

Conclusions and Recommendations

Conclusions

This review aimed to investigate the effectiveness of using the peer-mediated intervention, SPT, in improving the social communication skills of preschoolers with identified social communication difficulties. The results across the five included studies for preschoolers were inconsistent, where participants in only two of the studies, with either moderate developmental delay (Goldstein et al., 1997) or social delay (Milam et al., 2020), had significant improvements in their social communication skills. These positive outcomes for target children were also sustained when different peer buddies were used (Goldstein et al., 1997) and when the intervention strategies were systematically phased out (Milam et al., 2020).

For studies that included participants with diagnosable disabilities, for example ASD (Barber et al., 2016; Maich et al., 2018) and Down's syndrome (Severini et al., 2019), results were insignificant. Maich et al. (2018) discovered an overall increase in the number of social interactions, but only during structured activities. Opposing play interests between peer buddy and target child may also impact the level of interaction. Goldstein et al. (1997) used non-standardised assessment tools to consider the dynamic relationship between peers, which may be a contributing factor to the stronger effect size.

Limitations

Although all five studies received a 'medium' weighting according to the WoE framework (Gough, 2007), the heterogeneous nature of single case experimental designs causes difficulties when generalising the results. The target population sample, outcome measures and implementation of the intervention, all varied between studies. Therefore, it is important to draw inferences from the effectiveness of the intervention on a participant-by-participant basis.

In addition, it is important to note that none of the studies were conducted in the UK and the largest target-child sample size consisted of four participants (Goldstein et al., 1997). Although all studies occurred in OECD countries, the results cannot be directly extrapolated to the UK population, decreasing their external validity.

Despite SPT being a naturally occurring intervention in an integrated preschool, two studies were conducted in researcher contrived environments (Barber et al., 2016; Severini et al., 2019).

Furthermore, the studies that were conducted in a naturalistic setting also involved researcher training and input (Goldstein et al., 1997; Milam et al., 2020). Therefore, it is difficult to assume that the positive social communicative outcomes could be generalised within an educational setting, as school staff would be required to deliver the intervention.

Recommendations

For integrated preschools, SPT is a feasible intervention that holds promising implications when characteristics of the target child and peer buddy are carefully considered. EPs can encourage settings to consider the individual characteristics of the target child's disability when deciding the most appropriate social skills-based intervention. To access the benefits of SPT, target children need a basic or limited ability to engage in functional play prior to receiving additional support, as this intervention provides an opportunity to develop these skills further. Where children display no social communication skills, a more comprehensive intervention package that involves one-to-one adult input may be more suitable.

Whilst it is important to consider the needs of the target child, it is also beneficial to consider the appropriateness of the peer buddies. Milam et al. (2020) created a comprehensive inclusion criterion for peer buddies, which included: chronological age of 36 months; consistent attendance; frequent interaction with peers; and ability to attend during teacher lead activities. Settings should consider these factors when selecting peer buddies, as they may impact the effectiveness of SPT outcomes.

Variations in the peer buddy training programmes were also noted. Significant effects occurred in research where sensitivity training was included (Goldstein et al., 1997) and additional booster session to refresh the content following extended school breaks occurred (Milam et al., 2020). A bespoke training package for peer buddies and target children that cater specifically for the presenting needs may be most appropriate.

Further Research

Future research should consider investigating the long-term impact of SPT on the attitudes towards children with identified social-communication difficulties. Case studies have revealed a qualitative difference between friendships with two typically developing peers and friendships between a typically developing peer and child with SEN, which often mirrors a 'helper-helpee' dynamic (Van der Klift & Kunc, 2002). Although SPT aims to equip typically developing children with the relevant strategies to engage with their target peers, this may also discourage the authenticity of a reciprocated friendship. Therefore, it would be worthwhile to investigate whether SPT has a long-term impact on children's understanding and social schemas towards children with disabilities. It is also recommended for future research to be conducted in the UK and to utilise their teaching staff to deliver and implement SPT, to provide more externally valid conclusions.

References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Barber, A. B., Saffo, R. W., Gilpin, A. T., Craft, L. D., & Goldstein, H. (2015). Peers as clinicians: Examining the impact of Stay Play Talk on social communication in young preschoolers with autism. *Journal of Communication Disorders*, 59, 1-15. <https://doi.org/10.1016/j.jcomdis.2015.06.009>
- Cashin, A., & Barker, P. (2009). The triad of impairment in autism revisited. *Journal of Child and Adolescent Psychiatric Nursing*, 22(4), 189-193. <https://doi.org/10.1111/j.1744-6171.2009.00198.x>
- Chang, Y., & Locke, J. (2016). A systematic review of peer-mediated interventions for children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 27, 1-10. <https://doi.org/10.1016/j.rasd.2016.03.010>
- Department for Education and Department for Health (DfE & DH). (2014). *Special educational needs and disability code of practice: 0 to 25 years*. London: The Stationery Office. <https://www.gov.uk/government/publications/send-code-of-practice-0-to-25>
- Department for Education and Department for Health (DfE & DH). (2015). *Special educational needs and disability code of practice: 0 to 25 years*. London: The Stationery Office. (Revised January 2015). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/398815/SEND_Code_of_Practice_January_2015.pdf
- Department for Education (2014). *Early years: Guide to the 0 to 25 SEND code of practice. Advice for early years providers that are funded by the local authority*. London: Department for Education. <https://www.gov.uk/government/publications/send-guide-for-early-years-settings>
- Goldstein, H., English, K., Shafer, K., & Kaczmarek, L. (1997). Interaction among preschoolers with and without disabilities: Effects of across-the-day peer intervention. *Journal of Speech, Language, and Hearing Research*, 40(1), 33-48. <https://doi.org/10.1044/jslhr.4001.33>

- Gough, D. (2007). Weight of Evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213–228.
<https://doi.org/10.1080/02671520701296189>
- Gresham, F. M., Elliott, S. N., & Kettler, R. J. (2010). Base rates of social skills acquisition/performance deficits, strengths, and problem behaviors: An analysis of the social skills improvement system – rating scales. *Psychological Assessment* 22(4), 809–815.
- Guralnick, M. J., Connor, R. T., & Johnson, L. C. (2011). The peer social networks of young children with down syndrome in classroom programmes. *Journal of Applied Research in Intellectual Disabilities*, 24(4), 310-321. <https://doi.org/10.1111/j.1468-3148.2010.00619.x>
- Honig, A., & McCarron, P. A. (1988). Prosocial behaviors of handicapped and typical peers in an integrated preschool. *Early Child Development and Care*, 33(1-4), 113-125.
<https://doi.org/10.1080/0300443880330109>
- Horner, R. H., Carr, E. G., 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. <https://doi.org/10.1177/001440290507100203>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12–19. <https://doi.org/10.1037/10109-042>
- Kalfus, G. R. (1984). Peer Mediated Intervention. *Child & Family Behavior Therapy*, 6(1), 17-43.
https://doi.org/10.1300/J019v06n01_02
- Katz, E., & Girolametto, L. (2013). Peer-mediated intervention for preschoolers with ASD implemented in early childhood education settings. *Topics in Early Childhood Special Education*, 33(3), 133-143. <https://doi.org/10.1177/0271121413484972>
- Ledford, J. R., Osborne, K., & Chazin, K. T. (2016). Stay, play, talk procedures. *Evidence-based Instructional Practices for Young Children with Autism and Other Disabilities*.
<http://ebip.vkcsites.org/stay-play-talk-procedures>

- Maich, K., Hall, C. L., Van Rhijn, T. M., & Squires, K. (2018). Investigating Stay, Play, & Talk: A peer-mediated social skills intervention for young children with autism spectrum disorder and other social challenges. *Exceptionality Education International*, 28(2), 82–104.
<https://doi.org/10.5206/eei.v28i2.7766>
- Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly*, 17(1), 1-23.
<https://doi.org/10.1521/scpq.17.1.1.19902>
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732-749. <https://doi.org/10.1111/j.1467-8624.2008.01154.x>
- McLeod, S. A. (2016, February 5). *Bandura - social learning theory*. Simply Psychology.
<https://www.simplypsychology.org/bandura.html>
- Milam, M. E., Hemmeter, M. L., & Barton, E. E. (2020). The effects of systematic instruction on preschoolers' use of Stay-Play-Talk with their peers with social delays. *Journal of Early Intervention*, 43(1), 80-96. <https://doi.org/10.1177/1053815119900253>
- Norbury, C. F. (2014). Practitioner Review: Social (pragmatic) communication disorder conceptualization, evidence and clinical implications. *Journal of Child Psychology and Psychiatry*, 55(3), 204-216. <https://doi.org/10.1111/jcpp.12154>
- Odom, S. L., Buysse, V., & Soukakou, E. (2011). Inclusion for young children with disabilities. *Journal of Early Intervention*, 33(4), 344-356. <https://doi.org/10.1177/1053815111430094>
- Petersen, M. C., Kube, D. A., & Palmer, F. B. (1998). Classification of developmental delays. *Seminars Paediatric Neurology*, 5(1), 2-14. [https://doi.org/10.1016/S1071-9091\(98\)80012-0](https://doi.org/10.1016/S1071-9091(98)80012-0)
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health*, 57(7), 527–529.
<http://dx.doi.org/10.1136/jech.57.7.527>
- Rogers, S. J. (2000). Interventions That facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders*, 30(5), 399-409.

- Severini, K. E., Ledford, J. R., Barton, E. E., & Osborne, K. C. (2019). Implementing Stay-Play-Talk with children who use AAC. *Topics in Early Childhood Special Education, 38*(4), 220-233. <https://doi.org/10.1177/0271121418776091>
- Strain, P. S., & Bovey, E. H. (2011). Randomized, controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 31*(3), 133-15. <https://doi.org/10.1177/0271121411408740>
- Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., Wilson, B., Horner, R., Kazdin, A., Kratochwill, T., McDonald, S., Sampson, M., Shamseer, L., Togher, L., Albin, R., Backman, C., Douglas, J., Evans, J. J., Gast, D., ... Wilson, B. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 Statement. *Journal of Clinical Epidemiology, 73*, 142-152. <https://doi.org/10.2522/ptj.2016.96.7.e1>
- UNESCO (2009). *Policy guidelines on inclusion in education*. Paris: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000177849>
- Van der Klift, E., & Kunc, N. (2002) Beyond benevolence. In J.S. Thousand., R.A. Villa and A.I. Nevin (eds) *Creativity and Collaborative Learning: The Practical Guide to Empowering Student, Teachers, and Families* (pp. 21-8). Baltimore, MD: Paul H. Brook.
- Vannest, K.J., Parker, R.I., Gonen, O., & Adiguzel, T. (2016). Single Case Research: Web based calculators for SCR analysis. (Version 2.0) [Web-based application]. College Station, TX: Texas A&M University. Retrieved Friday 5th February 2021. Available from singlecaseresearch.org
- Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9*(1), 11-29.
- Wolery, M., & Ezell, H. K. (1993). Subject descriptions and single-subject research. *Journal of Learning Disabilities, 26*(10), 642-64. <https://doi.org/10.1177/002221949302601001>

Appendix A: Details of Excluded Studies

Table A1

List of excluded studies at full review

Reference	Criteria Number	Rationale
Hughett, K., Kohler, F. W., & Raschke, D. (2013). The effects of a buddy skills package on preschool children’s social interactions and play. <i>Topics in Early Childhood Special Education</i> , 32(4), 246-254. https://doi.org/10.1177/0271121411424927	6	Buddy skill Package, not the Stay, Play, Talk intervention.
Osborne, K., Ledford, J. R., Martin, J., & Thorne, K. (2019). Component Analysis of Stay, Play, Talk Interventions With and Without Self-Monitored Group Contingencies and Recorded Reminders. <i>Topics in Early Childhood Special Education</i> , 39(1), 5-18. https://doi.org/10.1177/0271121418815236	3	Peer confederates used a whole class of children with and without identifiable disabilities.
Thiemann-Bourque, K. S., McGuff, S., & Goldstein, H. (2017). Training peer partners to use a speech-generating device With classmates with autism spectrum disorder: Exploring communication outcomes across preschool contexts. <i>Journal of Speech, Language, and Hearing Research</i> , 60(9), 2648-2662. https://doi.org/10.1044/2017_JSLHR-L-17-0049	6	Peer-mediated intervention that does not involve Stay, Play, Talk

<p>Van Rhijn, T., Osborne, C., Ranby, S., Maich, K., Hall, C., Rzepecki, L., & Hemmerich, A. (2019). Peer play in inclusive child care settings: Assessing the impact of Stay, Play, & Talk, a peer-mediated social skills program. <i>Child Care in Practice: Northern Ireland Journal of Multi-disciplinary Child Care Practice</i>, 27(3),1-15. https://doi.org/10.1080/13575279.2019.1588707</p>	2	<p>Stay, Play, Talk was implemented on a whole-class level with typically developing children.</p>
<p>Whittaker, J. E. V., Williford, A. P., Carter, L. M., Vitiello, V. E., & Hatfield, B. E. (2018). Using a standardized task to assess the quality of teacher-child dyadic interactions in preschool. <i>Early Education and Development</i>, 29(2), 266-287. https://doi.org/10.1080/10409289.2017.1387960</p>	6	<p>Intervention involved Teacher-Child Structures Play Task (TC-SPT).</p>

Appendix B: Details of Included Studies

Table B1

Overview of the Included Studies: Mapping the Field

	Study Design	Sample Size	Peer Groupings	Setting	Presenting Difficulty for TC	Intervention	Country	Outcome Variable for TC
1. Barber et al. (2015)	Multiple baseline across participants design	$N = 6$	Dyads	University Child Development Centre	ASD	SPT over 16, 20-min intervention sessions (twice a week for 8 weeks)	U.S.	Combined initiation and response (“social interaction”) Early Communicative Index
2. Goldstein et al. (1997)	Multiple baseline across participants design	<i>Cohort 1:</i> $N = 8$ <i>Cohort 2:</i> $N = 8$	Dyads	Developmentally integrated preschool	Moderate Developmental Disabilities	SPT occurred during three time points: free play, snack time, activity time	U.S.	Number of social interactions per 10-min sample with class peers
3. Maich et al. (2018)	Single-subject AB design	$N = 3$	Whole class	Child care centre	ASD	Four step training procedure	Canada	Mean social interactions across 10-min data collection Researcher-created Social Skills Questionnaire for TC

4.	Milam et al. (2020)	Multiple probe design	$N = 9$	Triads	Community child care centre	Social delay	Four step training procedure	U.S.	Duration of social play
5.	Severini et al. (2019)	A-B-A-B withdrawal design	$N = 6$	Triads	University-based inclusive early childhood program	Down's syndrome	SPT with high-tech AAC training	U.S.	Number of social interactions between PB and TC

Note. TC refers to target child; PB refers to peer buddy; AAC refers to Augmentative and Alternative Communication

Appendix C: WoE A Coding Protocols

Coding protocol: Barber et al. (2015)

Coding Protocol: Single – Subject Research Design

Adapted from Horner et al. (2005), The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education

Date: 25.01.2021

Full Study Reference:

Barber, A. B., Saffo, R. W., Gilpin, A. T., Craft, L. D., & Goldstein, H. (2015). Peers as clinicians: Examining the impact of Stay Play Talk on social communication in young preschoolers with autism. *Journal of Communication Disorders*, 59, 1-15.
<https://doi.org/10.1016/j.jcomdis.2015.06.009>

1. Description of Participants and Settings

(a) Participants are described with sufficient detail to allow others to select individuals with similar characteristics (e.g., age, gender, disability, diagnosis).

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(b) The process for selecting participants is described with replicable precision.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(c) Critical features of the physical setting are described with sufficient precision to allow replication.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

2. *Dependent Variable*

(a) Dependent variable is described with operational precision.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(b) Each dependent variable is measured with a procedure that generates a quantifiable index.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(c) Measurement of the dependent variable is valid and described with replicable precision.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(d) Dependent variables are measured repeatedly over time.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(e) Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet the minimal standards (e.g., IOA = 80%; Kappa = 60%).

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

3. *Independent Variable*

(a) Independent variable is described with replicable precision.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(b) Independent variable is systematically manipulated and under the control of the experimenter.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(c) Overt measurement of the fidelity of implementation for the independent variable is highly desirable.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

4. *Baseline*

(a) The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and 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.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(b) Baseline conditions are described with replicable precision.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

5. Experimental control/internal validity

(a) The design provides at least three demonstrations of experimental effect at three different points in time.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(b) The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(c) The results document a pattern that demonstrates experimental control.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

6. External Validity

(a) Experimental effects are replicated across participants, settings, or materials to establish external validity.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

7. Social validity

(a) The dependent variable is socially important.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(b) The magnitude of change in the dependent variable resulting from the intervention is socially important.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(c) Implementation of the independent variable is practical and cost effective.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

(d) Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.

- All of the criteria are fulfilled = 3
- Mostly all of the criteria are fulfilled = 2
- Limited criteria are fulfilled = 1
- None of the criteria are fulfilled = 0

Table C1

Weight of Evidence (WoE) Calculations

	Overall Evidence Rating (0 – 3)	Evidence Descriptors
Description of Participants and Settings	2.3	Medium
Dependent Variable	2.6	High
Independent Variable	2.6	High
Baseline	2.5	High
Experimental control/internal validity	2	Medium
External Validity	1	Low
Social Validity	2	Medium

Note. <1.5 is low; 1.5 – 2.4 is medium; >2.4 is high

Average Quality of Evidence across the Key Judgement Areas
$\frac{\sum x}{N}$
x = Individual quality of evidence for each judgement area
N = Number of judgement areas
Overall Rating for Weight of Evidence A: 2.14 (Medium)

Appendix D: Overview of the WoE A Calculations

Table D1

An overview of the calculated WoE A scores and descriptors for each category identified in the Horner et al. (2005) protocol

Study	Description of Participants and Settings	Dependent Variable	Category from the WoE A Protocol			External Validity	Social Validity
			Independent Variable	Baseline	Experimental control/internal validity		
Barber et al. (2016)	2.3	2.6	2.6	2.5	2	1	2
Goldstein et al. (1997)	1.7	2.2	2	3	2.3	2	2.25
Maich et al. (2018)	2.6	2.4	1.6	2	2	2	2.75
Milam et al. (2020)	2	2.6	3	1.5	2.3	2	2.5
Severini et al. (2019)	2.6	2.8	3	2.3	2.3	1	2.25

Note. <1.4 is low; 1.5 – 2.4 is medium; >2.5 is high

Appendix E: WoE B Coding Protocol

Table E1

Weight of Evidence (WoE B): Methodological Relevance

WoE B Rating (Qualitative Descriptor)	Criteria	Rational
3 (High)	Randomised control trials	
2 (Medium)	Cohort studies, quasi-experimental studies, single case experimental designs	Petticrew & Roberts (2003) researched the appropriateness of different study designs to investigate the effectiveness of an intervention.
1 (Low)	Qualitative research, survey, case control, non-experimental evaluation	

Appendix F: WoE C Coding Protocol

Table F1

Weight of Evidence (WoE C): Topic Relevance

		Weightings	Rational
Relationship between target child and peer confederate	3	Peer buddies have been carefully selected for target children based upon selective criteria and taken children's social history and shared interests into account	Peer-mediated intervention involves the participation of staying, playing and talking between peers. Ledford et al. (2016) outlines what to consider when grouping children.
	2	Peer buddies have been carefully selected for target children based upon selective criteria	
	1	Peer buddies have been randomly allocated to target children	
Location of intervention	3	Research conducted in the UK	The review question is directed to support school-based interventions in the UK.
	2	Research conducted in OECD countries	
	1	Research conducted in countries outside the OECD	
Intervention setting	3	Intervention occurs in naturalistic preschool settings and routines across the day (e.g. free play, break time)	To improve external validity, it is preferred for the study to be conducted in a naturalistic setting.
	2	Intervention occurs in naturalistic setting, but contrived by the researcher	
	1	Intervention occurs in an artificial/clinical environment, manipulated by the researcher	

Intervention fidelity	3	Overt measures of fidelity are documented for accurate replication, including a clear outline of the intervention environment, the type and number of adult prompts, an adequate description of the training procedures (i.e. amount of sessions, an overview of what was taught), the level of adult feedback and reward systems used.	Overt measures of the intervention will ensure the quality of the intervention is maintained.
	2	Overt measures of fidelity are documented, but not to the same level of detail for accurate replication.	
	1	Little/no consideration for fidelity of the intervention	
Training on Stay, Play, Talk	3	Detailed training on each component of the intervention (stay, play, talk) have been conducted with target children and their peer confederates	Ledford et al. (2016) recommends conducting the training with peer buddies to practice with the child they will also be working with.
	2	Detailed training on each component of the intervention (stay, play, talk) have been conducted with only peer confederates	
	1	Limited detail on the different components of the training	
Intervention resources (play materials, visual reminders, reinforcers, data collection tools)	3	All materials stated have been included	The recommended materials needed for the intervention
	2	Three out of the four materials have been included	
	1	One or less materials have been included	
Scope of outcome measures	3	At least 3 measures of social outcomes are reported (e.g. number of social interactions, duration of social play, questionnaires related to social skills), including at least one generalisation phase.	To calculate the effectiveness of the intervention on social communicative skills for target children, multiple outcomes measures should be assessed, including a generalisation phase.
	2	At least 2 measures of social outcomes are reported	
	1	Only one measure of social outcome is reported	

Appendix G: Overview of the WoE C Calculations

Table G1

An overview of the WoE C scores across the different categories on the study's relevance

Study	Relationship between target child and peer confederate	Location of intervention	Category from the WoE A Protocol			Intervention resources	Scope of outcome measures
			Intervention setting	Intervention fidelity	Training on Stay, Play, Talk		
Barber et al. (2016)	2	2	1	3	2	3	3
Goldstein et al. (1997)	3	2	3	2	2	3	3
Maich et al. (2018)	1	2	3	1	3	3	2
Milam et al. (2020)	1	2	3	3	3	3	3
Severini et al. (2019)	1	2	2	3	3	3	2