

Case Study 1: An Evidence-Based Practice Review Report

Theme: School Based Interventions for Children and Young People with Special Educational Needs (SEN).

How effective are imitative interventions in increasing the social communication skills of adolescents and young people with autism spectrum disorder?

Summary

There is a range of imitative interventions that aims to increase social skills, including reciprocal imitation training and video modelling. Both interventions involve viewing oneself or others using positive social interactions in order to teach the social function of imitation. Imitative interventions have been used in school settings to support adolescents and young people with difficulties in developing their social skills.

This systematic literature review aims to investigate the effectiveness of imitative interventions that could be delivered in a school-based setting. It will examine if they are effective in developing social skills in adolescents and young people with autism spectrum disorder. The main conclusion of this review is that, while there is promising initial evidence, there is inadequate research to support the effectiveness of these interventions with young people with autism spectrum disorder. This is due to the lack of replication of significant findings across studies. There is also limited literature pertaining to the effectiveness of such interventions with adolescents, and so it is recommended that additional research is required.

Introduction

a. The need for social communication interventions for adolescents and young people with autism spectrum disorder

Autism spectrum disorder is a developmental condition that is categorised by persistent deficits in social communication and social interaction (American Psychiatric Association, 2013). Tantam (2003) argued that these social difficulties become increasingly pronounced once individuals reach adolescence. This is reflected in research findings which show that adolescents with autism spectrum disorder report greater feelings of loneliness and limited friendships when compared with their peers (Bauminger & Kasari, 2000). Increased feelings of anxiety is another negative outcome associated with the social exclusion of adolescents with autism spectrum disorder (Sebastian, Blakemore & Charman, 2009). It is, therefore, important to support adolescents and young people in developing their social communication skills, in order to promote inclusion and increase positive outcomes for their future.

b. Imitative interventions-Psychological Basis

Imitation has widely been linked to our ability to learn other social communication skills, including joint attention and social reciprocity (McDuffie et al., 2007). Research has found that children with autism spectrum disorder have impaired imitative abilities of both actions with objects and facial expressions (Markodimitraki, Kypriotaki, Ampartzaki, & Manolitsis, 2013). However, studies have also established that children with autism spectrum

disorder develop their verbal interaction skills through contingent imitation (Ishizuka & Yamamoto, 2016). It has, therefore, been concluded that teaching and using imitation skills within interventions will increase the social communication abilities of children with autism (Ingersoll, 2008).

Recent neuroscientific research has identified five key component skills of imitation (Kana, Wadsworth & Travers, 2011):

- Attention
- Self-other matching
- Goals
- Familiarity
- Social relevance

Historically, behavioural approaches have been used to teach imitation.

Ingersoll (2008) concluded, however, that these interventions fail to produce generalised and maintained imitation skills in the absence of reinforcement.

Reciprocal imitation training is a naturalistic intervention that attempts to teach children and young people with autism spectrum disorder the social function of imitation (Ingersoll, 2008). Reciprocal imitation training aims to elicit intrinsic motivation by emphasising the social role of imitation using naturalistic behavioural and developmental strategies during ongoing play interactions. A typical reciprocal imitation training session includes exercises where the therapist imitates the child's actions in an attempt to increase their imitation rates during contingent imitation activities. The therapist will also verbally describe the actions of the child that they are imitating. A child will

usually have two to six, 20- minute sessions of reciprocal imitation training per week, over a 10-week period. Reciprocal imitation training was designed as an individual intervention for pre-schoolers and delivered by therapists, but more recently, it has been delivered by teaching staff in schools to adolescents with autism spectrum disorder (Ingersoll, Berger, Carlsen & Hamlin, 2017; Ingersoll, Walton, Carlsen & Hamlin, 2013).

Research suggests that reciprocal imitation training is successful in teaching young children gesture imitation and spontaneous object imitation (Ingersoll, Lewis & Kroman, 2007). Ingersoll and Schreibman (2006) also found that by targeting object imitation, reciprocal imitation training increases pretend play. Furthermore, these findings have shown that imitation skills were generalised by the children within the study and that the increase in imitation was maintained over time. Findings from the use of reciprocal imitation training with adolescents with autism spectrum disorder have also identified an increase in social emotional functioning (Ingersoll et al., 2017). There has been a recent review that found reciprocal imitation training to be an effective treatment for imitation deficits in young children with autism spectrum disorder (Krupicz, 2014). However, there has been no review of the evidence to assess how effective reciprocal imitation training is for adolescents and young people with autism spectrum disorder in developing their social skills.

Video modelling interventions have widely been used in developing social communication skills of children and adolescents, with empirical evidence supporting its effectiveness (Bellini & Akullian, 2007). Video modelling is based on Bandura's 'Social Learning Theory' (Bandura, 1977), whereby individuals can learn through observing others' behaviour. A typical video

modelling session involves viewing a vignette of oneself or a model, engaging in a specific behaviour. The child or young person is then given the chance to imitate the observed behaviour. Video modelling has been increasingly used in school settings (Hitchcock, Dowrick & Prater, 2003) due to its flexibility in format, for example, using a laptop, television or iPod. With regard to increasing social communication skills, research has previously found that video modelling is effective in developing social initiation skills (Nikopoulos & Keenan, 2003) and reciprocity (Nikopoulos & Keenan, 2004). Similar to the reciprocal imitation training literature, there is a range of evidence to support the use of video modelling interventions to increase social skills in young children (as previously cited), but limited data to validate its effectiveness with adolescents and young people with autism spectrum disorder.

c. Relevance to Educational Psychology Practice and Rationale for Review

The role of the educational psychologist includes supporting children and young people between the ages of 0 and 25, with varying special educational needs and disabilities. These include 'communication and interaction' and 'social, emotional and mental health' needs (DfE & DoH, 2015). Educational provisions have been focusing on trying to promote the inclusion of adolescents with autism spectrum disorder of late (Williams, Johnson & Sukhodolsky, 2005), and increasingly school-based social skills interventions have been adopted (Stanton-Chapman, Denning & Jamison, 2012). Arguably, there is currently limited evidence to support these interventions,

especially pertaining their effectiveness for children and young people with autism spectrum disorder (Howlin, Gordon, Pasco, Wade & Charman, 2007).

Educational psychologists are often involved in recommending, training and supervising the implementation of interventions within school-based settings. Therefore, it is important that educational psychologists apply and endorse evidence-based practice, which includes research that reports the effectiveness of interventions. Based on this information, it would be beneficial for educational psychology practice to explore the literature basis for social communication interventions for adolescents and young people with Autism.

Therefore, the question addressed in this review is:

How effective are imitative interventions in increasing the social communication skills of adolescents and young people with autism spectrum disorder?

Critical Review of the Evidence Base

a. Literature Search and Screening Strategy

Literature searches of published journals were conducted in February 2018, using electronic databases PsycInfo, ERIC and UCL online library. These databases allowed access to articles that were related to both education and psychology. The inability to retrieve a sufficient number of peer-reviewed journal articles, resulted in the inclusion of dissertations within the literature search. The search included the following terms:

“reciprocal imitation training OR video model*”

AND

“adolescen* OR teenag* OR young person”

AND

“autism OR asd OR autism spectrum disorder”

The database searches produced 162 results. Of these 162 studies, 152 were excluded from the review as they were found to be duplicates or they did not meet the inclusion criteria (see Table 1). The exclusion of these studies was carried out by screening the titles and abstracts. The remaining 10 articles were read in their entirety, and a further five were excluded with rationale (see Appendix A).

b. Weight of Evidence

Gough’s Weight of Evidence (WoE) Framework (Gough, 2007) was used to appraise each study within the review. They were weighted on their methodological quality (WoE A) which assesses the procedures used within each study, based on quality standards. The methodological relevance (WoE B) was also weighted to assess the appropriateness of the research design, and the significance of the topic to the review question (WoE C). To provide an overall Weight of Evidence score for each study (WoE D), the scores for each of these categories were then averaged (see Table 3). To assess the WoE A for the studies within the review, two different coding protocols were adapted and used. For group designs, the Kratochwill et al. (2003) coding protocol was used, and for single case experimental designs, the Horner et al. (2005).

Each study was assigned a WoE description of 'low' where scores are 1.4 or below, 'medium' if between 1.5 and 2.4, or 'high' if the score lies between 2.5 and 3.0. Overall, findings from studies with high WoE D ratings (Ingersoll et al., 2013; Ingersoll et al., 2017; Williamson, Casey, Robertson & Buggey, 2013; Whittington-Barnish, 2012) should be given more weight in drawing conclusions from the review, than studies that receive medium or low ratings (Mason, Rispoli, Ganz, Boles & Orr, 2012).

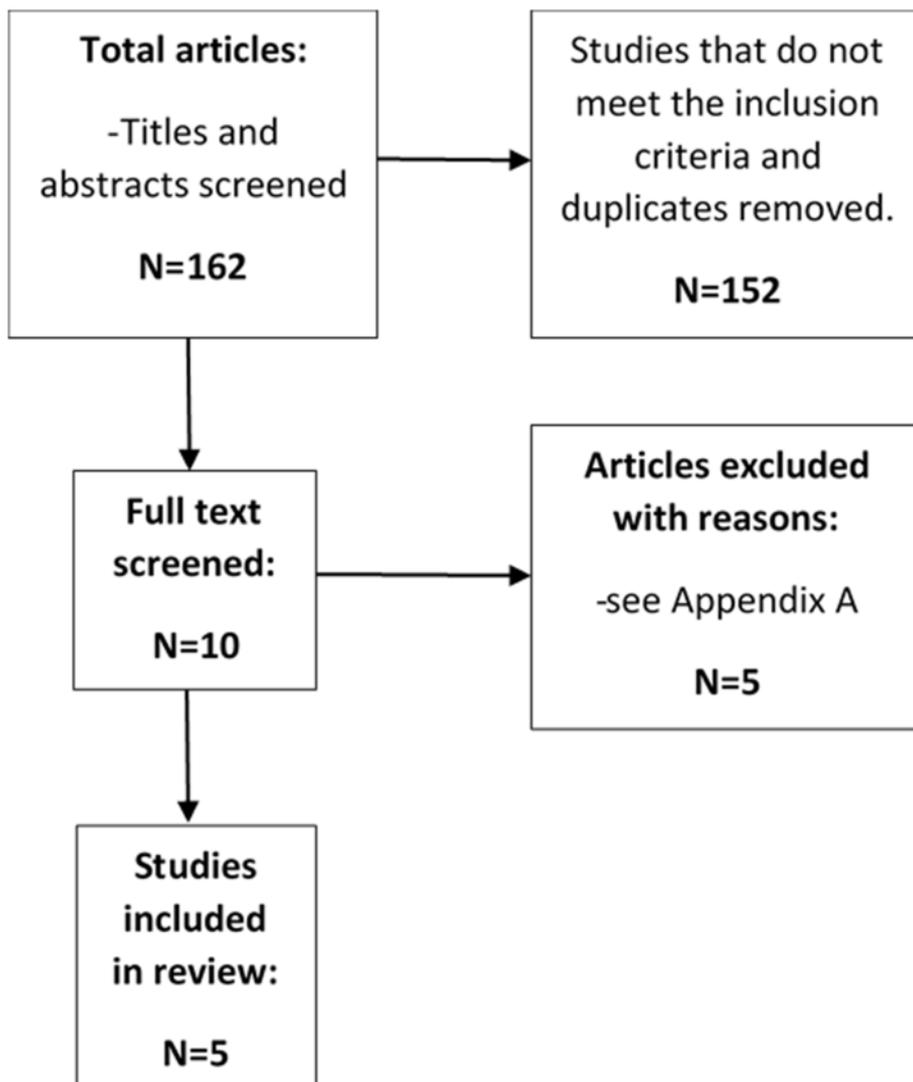


Figure 1. Literature search strategy

Table 1

Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion Criteria	Rationale
1. Participants	Study participants who are adolescents or young adults aged 11-25 and have a diagnosis of ASD.	Participants do not have a diagnosis of ASD and are younger than 11.	The focus of the study was on adolescents and young adults with a diagnosis of ASD. EPs work with young people up to the age of 25 (DfE, 2015) and predominantly in school settings. No review has yet looked at the effectiveness of imitative interventions with adolescents and young adults.
2. Setting	Study is set in an academic setting.	Study is in a home setting, clinic or early years.	No review has yet looked at whether imitative interventions are effective for adolescents and young adults with ASD when delivered in academic settings.
3. Intervention	Intervention is based on imitation.	Intervention is not based on imitation.	Review aims to investigate effectiveness of imitative interventions.
4. Type of study	Study uses an experimental design.	Study does not use an experimental design.	The aim of the review is to compare findings from experimental studies in order to determine the effectiveness of

Criteria	Inclusion	Exclusion Criteria	Rationale
			the interventions.
5. Outcome variables	Study has outcome variables measuring social communication skills.	Study does not have outcome variables measuring social communication skills.	Social communication skills were chosen as the outcome variable to investigate within the review, due to persistent difficulties for young people with ASD.
6. Language	Report is in the English language.	Report is not in the English language.	Unable to translate reports.

Table 2

Full References of Included Studies

Ingersoll, B., Berger, N., Carlsen, D. & Hamlin, T. (2017). Improving social functioning and challenging behaviors in adolescents with ASD and significant ID: A randomized pilot feasibility trial of reciprocal imitation training in a residential setting. *Developmental Neurorehabilitation*, 20(4), 236-246.

Ingersoll, B., Walton, K., Carlsen, D. & Hamlin, T. (2013). Social Intervention for Adolescents with Autism and Significant Intellectual Disability: Initial Efficacy of Reciprocal Imitation Training. *American Journal on Intellectual and Developmental Disabilities*, 118(4), 247-261.

Mason, R. A., Rispoli, M., Ganz, J. B., Boles, M. B. & Orr, K. (2012). Effects of video modeling on communicative social skills of college students with asperger syndrome Intervention Name (description of study): Video Modeling. *Developmental Neurorehabilitation*, 15(6), 425-434.

Whittington-Barnish, A. K. (2012). *Research to practice: Evaluation of conversation skills video modeling intervention for adolescents with autism*. ProQuest Dissertations and Theses: ProQuest Dissertations Publishing.

Williamson, R. L., Casey, L. B., Robertson, J. S. & Bugghey, T. (2013). Video Self-Modeling in Children with Autism: A Pilot Study Validating Prerequisite Skills and Extending the Utilization of VSM across Skill Sets. *Assistive technology*, 25(2), 63-71.

See Appendix B for a summary of the included studies.

Table 3

Weight of Evidence of Included Studies

Authors	Methodological Quality WOE A	Methodological Relevance WOE B	Relevance to Review Question WOE C	Overall Weight of Evidence WOE D
Ingersoll, Berger, Carlsen & Hamlin, 2017	2.6 (High)	3 (High)	3 (High)	2.9 (High)
Ingersoll, Walton, Carlsen & Hamlin, 2013	2.9 (High)	3 (High)	2.8 (High)	2.9 (High)
Mason, Rispoli, Ganz, Boles & Orr, 2012	2.1 (Medium)	3 (High)	2 (Medium)	2.4 (Medium)
Whittington-Barnish, 2012	2.7 (High)	3 (High)	2.5 (High)	2.7 (High)
Williamson, Casey, Robertson & Buggiey, 2013	2.6 (High)	3 (High)	2.3 (Medium)	2.6 (High)

c. Participants

In total, there were 39 participants included within the studies in this review, aged between 11 and 25. The participants had variable needs across the studies, but all participants had a diagnosis of autism spectrum disorder. One study did not have any clear rationale for participant selection and did not screen for need (Williamson, Casey, Robertson & Bugghey, 2013). Two studies used referrals from academic staff of students with existing diagnoses of autism spectrum disorder (Ingersoll et al., 2013; Mason, Rispoli, Ganz, Boles & Orr, 2012). Two studies (Ingersoll et al., 2017; Whittington-Barnish, 2012) conducted separate analyses of data of participants with autism spectrum disorder and social skills deficits. This will likely impact the outcomes of the studies, and so this was reflected in the WoE C ratings. Higher ratings were given to studies that clearly identified participants starting the study with social skills deficits through explicit analyses (Ingersoll et al., 2017; Whittington-Barnish, 2012). Studies that identified participants with social skills deficits through staff observations (Ingersoll et al., 2013; Mason, Rispoli, Ganz, Boles & Orr, 2012) or without any clear rationale (Williamson et al., 2013) were given lower WoE C ratings.

There is a gender disparity amongst all of the studies within the review, with double or more males than females (Ingersoll et al., 2017; Ingersoll et al., 2013; Mason et al., 2012; Whittington-Barnish, 2012; Williamson et al., 2013). This indicates a gender bias in the literature towards males with autism spectrum disorder. Research, however, has highlighted that there are higher rates of autism spectrum disorder diagnoses in males than females (Halladay et al., 2015). This could explain the disparity in autism spectrum

disorder literature of males and females, although there is very limited research that has focused on identifying the possible contributing factors.

One study used a comparison group design, with a total of 20 participants (Ingersoll et al., 2017). This study is underpowered based on Cohen's (1992) suggested group sizes for detecting effect sizes with a significance level of 0.5 and a power level of 0.8 (large effect sizes are significant with 23 participants; medium with 64 participants; small with 393 participants). As a result, when considering the effects of the group study, it is important to do so with caution. This is accounted for in the WoE A criteria. A single subject design was employed by four of the studies, whereby the sample size can be one (Ingersoll et al., 2013; Mason et al., 2012; Whittington-Barnish, 2012; Williamson et al., 2013). This design methodology allows the participant to act as their own control. As a result, separate criteria for WoE A and WoE B were used to judge methodological quality of single case experimental designs (whereby Horner et al.'s, 2005 criteria was used) and group designs (using Kratochwill et al.'s 2003 criteria) to ensure equitable weighting.

Various settings were used for the studies, in the United States of America (USA). Two studies took place in mainstream schools (Whittington-Barnish, 2012; Williamson et al., 2013) and the other two studies took place in special schools (Ingersoll et al., 2017; Ingersoll et al., 2013). One study was conducted in a college (Mason et al., 2012), which limits the generalisability of findings to school settings. This is reflected in the WoE C criteria. The USA is an Organisation for Economic Co-operation and Development (OECD) country, with a similar education system to the United Kingdom. As all of the studies were carried out in the same country, and based on the similarity of

the education system of the USA to the UK, country was not deemed to have a significant impact on WoE when conducting this review.

d. Intervention Content and Fidelity

All interventions were based on imitation but with varying content and fidelity. Two of the studies used the reciprocal imitation training intervention (Ingersoll et al., 2017; Ingersoll et al., 2013), whilst the other three studies used a video modelling intervention (Mason et al., 2012; Whittington-Barnish, 2012; Williamson et al., 2013).

Interventions were delivered by school staff (Ingersoll et al., 2017; Ingersoll et al., 2013; Williamson et al., 2013), a lead researcher (Mason et al., 2012) or research assistants (Whittington-Barnish, 2012). This was reflected in the WoE C ratings, whereby studies that used school staff only to deliver the intervention were rated higher than those which were delivered by staff and researchers, or researchers only. All of the studies which used school staff or research assistants to implement the intervention included varying degrees of training, either 2 days, 2 weeks or weekly. Two studies also explicitly stated the use of facilitator supervision for school staff and research assistants (Ingersoll et al., 2017; Whittington-Barnish, 2012). Four of the studies used treatment fidelity checklists, either self-check (Williamson et al., 2013) or via video recordings (Ingersoll et al., 2017; Ingersoll et al., 2013; Whittington-Barnish, 2012). Experimenter effects may have also affected the outcomes of all five studies, as they did not counterbalance the change agents used to administer the interventions. This is accounted for in WoE A.

e. Measures

Two of the studies used multiple sources and methods of data collection including standardised tests to measure the dependent variable 'social skills' (Ingersoll et al., 2017; Whittington-Barnish, 2012). The use of multiple sources increases the measure of outcomes for improvements in social skills. To measure the dependent variable, Ingersoll et al., (2017) used the Matson Evaluation of Social Skills for Individuals with Severe Retardation (MESSIER) and the Aberrant Behaviour Checklist-Residential (ABC-R) which were completed by school staff. Whittington-Barnish (2012) used the Autism Social Skills Profile (ASSP) and the Social Responsiveness Scale (SRS) which was completed by both parents and school staff. Using standardised tests also provides statistical evidence to support the measurement of outcomes. This is reflected in their high WoE A scores.

Three studies used observation as a single method of measurement for the dependent variable 'social skills' (Ingersoll et al., 2013; Mason et al., 2012; Williamson et al., 2013). This could decrease the reliability of the outcomes reported, due to observation bias. Inter-rater reliability was reported by Ingersoll et al., (2013) and Williamson et al., (2013) which increased their methodological quality, therefore obtaining a higher WoE A score.

f. Outcomes

There is inconclusive information regarding the effectiveness of imitative interventions based on the studies in the review. Effect sizes for the group study were provided by the researcher (Ingersoll et al., 2017) and are reported in Table 4. Effect sizes for single case experimental studies were

provided by one researcher (Mason et al., 2012), based on percentage of non-overlapping data points with the results reported in Table 5. Effect sizes could only be calculated for a further one of the remaining three studies (Whittington-Barnish, 2012) using the reported t test scores, which are reported in Table 6. The remaining two studies did not provide effect sizes within their reported data, or include enough information for one to be calculated (Ingersoll et al., 2013; Williamson et al., 2013). Effect sizes were given descriptors based on Cohen's (1988) criteria, whereby 0.2 constitutes a small effect; 0.5 is a medium effect and 0.8 is a large effect.

Positive effects were found for reciprocal imitation training across both studies (Ingersoll et al., 2017; Ingersoll et al., 2013), with regards to an increase in social skills. Medium effect sizes are reported by Ingersoll et al., (2017), showing increased social functioning of adolescents with autism spectrum disorder following the reciprocal imitation training intervention. Similarly, Ingersoll et al., (2013) reported findings that rates of joint interaction increased in half of their participants, further suggesting that reciprocal imitation training may improve social responsiveness in adolescents with autism spectrum disorder.

The data for video modelling effects on social skills were also variable. Mason et al. (2012) reported a significant large effect size for turn taking and moderate effect size for eye contact, but not for facial expression. Whittington-Barnish (2012) also reported a large significant effect of video modelling on increasing social skills, based on the 'teachers' reported Autism Social Skills Profile ratings. However, a further three measure effects were

not significant. Williamson et al. (2013) only reported increasing social skills in one of three participants.

g. Follow-up

Two of the studies (Ingersoll et al., 2013; Whittington-Barnish, 2012) collected follow-up data. Whittington-Barnish (2012) collected data monthly, for 6 months after the intervention was complete. Mean scores were presented at times, 1, 2, 3, 4, 5 and 6. Data shows the effects of the intervention on social skills as maintained and generalised, for between 4 and 6 months upon follow-up. The studies did not include enough information to calculate effect sizes for follow-up data, therefore it was not cited as criteria within the WoE ratings.

Table 4

Effect Sizes for Social Skills (Group Study)

Author	Measures	Intervention			Control			Cohen's d Pretest- posttest Effect Size (Intervention vs Control)	p	Effect Size Descriptor	Total WoE
		N	Pre M (SD)	Post M (SD)	N	Pre M (SD)	Post M (SD)				
Ingersoll, Berger, Carlsen & Hamlin, 2017	Matson Evaluation of Social Skills for Individuals with Severe Retardation (MESSIER)	10	148.80 (21.11)	157.80 (19.01)	9	136.11 (13.77)	138.11 (19.21)	0.76	.04	Medium	2.9 High
	Aberrant Behaviour Checklist- Residential (ABC-R)		48.50 (23.01)	37.50 (17.17)		48.89 (18.49)	56.11 (21.94)	0.70	.05	Medium	

Table 5

Effect Sizes of non-overlapping data for Social Skills (Single Case Experimental Design)

Author	Measures	Social Skill	Percent non overlap	IRD	p	Effect Size Descriptor	Total WoE
Mason, Rispoli, Ganz, Boles & Orr, 2012	Direct observation	Eye contact	0.81	0.56	0.01	Medium	2.4 Medium
		Facial expression	0.74	0.38	0.09	Small	
		Turn taking	1.00	1.00	0.00	Large	

IRD effect sizes below 0.50 are described as small, whereas IRD effect sizes from 0.50–0.70 are described as moderate, and those above 0.70 are described as large (Parker, Vannest & Brown, 2009).

Table 6

Effect Sizes of t test for Social Skills (Single Case Experimental Design)

Author	Measures	N	t score	Effect Size	p	Effect Size Descriptor	Total WoE
Whittington-Barnish, 2012	Teacher Autism Social Skills Profile (ASSP)	11	3.89	2.3	.00	Large	2.7 High
	Parent Autism Social Skills Profile (ASSP)	9	0.768	0.5	.77	Medium	
	Teacher Social Responsiveness Scale (SRS)	11	1.44	0.9	.09	Large	
	Parent Social Responsiveness Scale (SRS)	10	-1.10	-0.7	.85	Medium	

Effect sizes for this study was calculated by looking at within-participant pre-post change (Becker, 1988).

Conclusion and recommendations

In conclusion, evidence of the effectiveness of imitative interventions for increasing social skills in adolescents and young people with autism spectrum disorder is limited. The studies show some consensus on the effectiveness of imitative interventions; however, there were variable effects found across the five studies including insignificant outcomes. Given the insufficient data provided by two of the studies (Ingersoll et al., 2013; Williamson et al., 2013), and the absence of an effect size, it is difficult to identify any fixed conclusions.

It is difficult to identify the reason for the variability of findings across the studies, due to the variety of content and setting. Across the studies, there were several methodological weaknesses. This limits the generalisability of their findings for the effectiveness of imitative interventions. Ingersoll et al., (2017) had insufficient power, based on their small sample size. Thus, although they reported medium effects, these may not accurately reflect the strength of the intervention. All of the studies included within the review were based upon data that was skewed towards males, which also limits the generalisability of findings across genders. The effectiveness of imitative interventions on the development of social skills may vary significantly for females compared to males. Additionally, the lack of quantitative data reported for outcome measures was a limitation within the review, with particular reference to the two studies for which an effect size could not be calculated (Ingersoll et al., 2013; Williamson et al., 2013). Another key point to raise is that the researcher who conducted the studies to measure the effectiveness of reciprocal imitation training, was the originator of the

intervention (Ingersoll et al., 2017; Ingersoll et al., 2013). Therefore, the researcher's aims may have had an influential effect on the research findings.

Imitative interventions are based upon a naturalistic approach, which allows it to be easily applied to school-based teaching. In addition, both of the studies for reciprocal imitation training (Ingersoll et al., 2017; Ingersoll et al., 2013) and one study for video modelling (Whittington-Barnish, 2012) reported strong social validity responses about the interventions from teachers. While the evidence base for this approach is reasonably strong for children with autism spectrum disorder (Krupicz, 2014; Bellini & Akullian, 2007), the evidence base for adolescents and young people needs further development. This review should help guide educational psychologists in making an informed decision about the effectiveness and appropriateness of recommending the social skills interventions, reciprocal imitation training and video modelling. The literature review shows that some evidence suggests imitative interventions are effective in developing the social communication skills of adolescents and young people with autism spectrum disorder.

Furthermore, the significant findings have been reported when the intervention was delivered by researchers or school-based practitioners. This allows schools to deliver imitative interventions with limited external costs (for example, staff training). Research with young children has also shown that reciprocal imitative training (Ingersoll & Gergans, 2007) can also be effectively delivered by parents. If these findings were replicated with adolescents in future research, this would allow educational psychologists to

recommend one type of intervention for both home and school use, as part of their systemic approach to working.

Based on the findings discussed within this review, educational psychologists should note that while the evidence base for adolescents and young people is currently limited, it may still be beneficial to recommend imitative interventions for individuals with social communication difficulties. Imitative interventions are relatively new and emerging, such as reciprocal imitation training and video modelling, and have produced promising initial evidence. Future research should ensure the use of reliable and valid quantitative outcome measures, to increase methodological quality and to allow firm conclusions to be drawn from the data. Additionally, there should be sufficient power to calculate effect sizes.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: American Psychiatric Association text revision.
- Bandura, A. (1977). *Social Learning Theory*. New York: General Learning Press.
- Bauminger, N., & Kasari, C. (2000). Loneliness and friendship in high-functioning children with autism. *Child Development, 71*, 447–456.
- Becker, B. J. (1988). Synthesizing standardized mean change measures. *British Journal of Mathematical and Statistical Psychology, 41*, 257–278.
- Bellini, S., & Akullian, J. (2007). *Exceptional Children, 73*, 261-284.
- Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.
- Department for Education and Department of Health. (2015). *Special Educational needs and disability code of practice: 0-25 years*. Available at: <https://www.gov.uk/government/publications/send-code-of-practice-0-to-25> (Accessed: 18 February 2018).
- Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 37–41.

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., Messinger, D., Pelphrey, K., Sanders, S. J., Singer, A. T., Taylor, J. L. & Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. *Molecular autism*, 6, 36.

Hitchcock, C. H., Dowrick, P. W., & Prater, M. A. (2003). Video Self-Modeling Intervention in School-Based Settings. *Remedial and Special Education*, 24(1), 36-45.

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.

Howlin, P. R., Gordon, K., Pasco, G., Wade, A., & Charman, T. (2007). The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of Child Psychology and Psychiatry*, 48(5), 473–48.

Ingersoll, B. (2008). The social role of imitation in autism: Implications for the treatment of imitation deficits. *Infants & Young Children*, 21, 107-119.

Ingersoll, B., Berger, N., Carlsen, D. & Hamlin, T. (2017). Improving social functioning and challenging behaviors in adolescents with ASD and significant ID: A randomized pilot feasibility trial of reciprocal imitation training in a residential setting. *Developmental Neurorehabilitation*, 20(4), 236-246.

Ingersoll, B., & Gergans, S. (2007). The effect of a parent-implemented imitation intervention on spontaneous skills in young children with autism. *Research in developmental disabilities*, 28(2), 163-175.

Ingersoll, B., Lewis, E., & Kroman, E. (2007). Teaching the Imitation and Spontaneous Use of Descriptive Gestures in Young Children with Autism Using a Naturalistic Behavioral Intervention. *Journal of Autism and Developmental Disorders*, 37(8), 1446-1456.

Ingersoll, B. & Schreibman, L. (2006). Teaching Reciprocal Imitation Skills to Young Children with Autism Using a Naturalistic Behavioral Approach: Effects on Language, Pretend Play, and Joint Attention. *Journal of Autism and Developmental Disorders*, 36(4), 487-505.

Ingersoll, B., Walton, K., Carlsen, D. & Hamlin, T. (2013). Social Intervention for Adolescents with Autism and Significant Intellectual Disability: Initial Efficacy of Reciprocal Imitation Training. *American Journal on Intellectual and Developmental Disabilities*, 118(4), 247-261.

Ishizuka, Y., & Yamamoto J. (2016). Contingent Imitation Increases Verbal Interaction in Children with Autism Spectrum Disorders. *Autism: The International Journal of Research and Practice*, 20(8), 1011-1020.

Kana, R. K., Wadsworth, H. M. & Travers, B. G. (2011). A systems level analysis of the mirror neuron hypothesis and imitation impairments in autism spectrum disorders. *Neuroscience and Biobehavioral Reviews*, 35(3), 894-902.

Kratochwill, T., Fuchs, L., Abidin, D., Hughes, J., Batch, G., Keith, T., & Strein, W. (2003). *Task Force on Evidence-Based Interventions in School Psychology*. Available at:

<https://www.indiana.edu/~ebi/documents/workingfiles/EBImanual1.pdf>

(Accessed: 18 February 2018).

Krupicz, M. (2014). *Critical review: The efficacy of Reciprocal Imitation Training as a treatment for imitation deficits in young children with Autism Spectrum Disorder*. University of Western Ontario: School of Communication Sciences and Disorders.

Markodimitraki, M., Kypriotaki, M., Ampartzaki, M., & Manolitsis, G. (2013). Effects of context and facial expression on imitation tasks in preschool children with autism. *Early Child Development and Care*, 183, 1276–1292.

Mason, R. A., Rispoli, M., Ganz, J. B., Boles, M. B. & Orr, K. (2012). Effects of video modeling on communicative social skills of college students with asperger syndrome Intervention Name (description of study): Video Modeling. *Developmental Neurorehabilitation*, 15(6), 425-434.

McDuffie, A., Turner, L., Stone, W., Yoder, P., Wolery, M., & Ulman, T. (2007). Developmental correlates of different types of motor imitation in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37(3), 401-412.

Nikopoulos, C. K., & Keenan, M. (2003). Promoting social initiation in children with autism using video modeling. *Behavioral Interventions*, 18(2), 87-108.

Nikopoulos, C. K., & Keenan, M. (2004). Effects of video modeling on social initiations of children with autism. *Journal of Applied Behavior Analysis*, 37(1), 93-96.

Parker, R. I, Vannest, K. J. & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children*, 75, 135–150.

Sebastian, C., Blakemore, S., & Charman, T. (2009). Reactions to ostracism in adolescents with autism spectrum conditions. *Journal of Autism and Developmental Disorders, 39*(8), 1122–1130.

Stanton-Chapman, T. L., Denning, C. B., & Jamison, K R. (2012). Communication skill building in young children with and without disabilities in a preschool classroom. *The Journal of Special Education, 46*(2), 78-93.

Tantam, D. (2003). The challenge of adolescents and adults with Asperger syndrome. *Child and Adolescent Psychiatric Clinics of North America, 12*, 143–163.

Whittington-Barnish, A. K. (2012). *Research to practice: Evaluation of conversation skills video modeling intervention for adolescents with autism*. ProQuest Dissertations and Theses: ProQuest Dissertations Publishing.

Williams, S. K., Johnson, C., & Sukhodolsky, D. G. (2005). The role of the school psychologist in the inclusive education of school age children with autism spectrum disorders. *Journal of School Psychology, 43*, 117–136.

Williamson, R. L., Casey, L. B., Robertson, J. S. & Bugghey, T. (2013). Video Self-Modeling in Children with Autism: A Pilot Study Validating Prerequisite Skills and Extending the Utilization of VSM across Skill Sets. *Assistive technology, 25*(2), 63-71.

Appendix A

Table 7

Excluded Articles

Reference	Reason for Exclusion
Detar, W. J. (2013). Targeting Question-Asking Initiations through Video-Feedback to Improve Social Conversation in College Students with Autism Spectrum Disorders. ProQuest LLC, Ph.D. Dissertation, University of California, Santa Barbara.	1.Participants without learning difficulties. 2.Intervention delivered in a university setting.
Macpherson, K., Charlop, M. H. and Miltenberger, C. A. (Using Portable Video Modeling Technology to Increase the Compliment Behaviors of Children with Autism During Athletic Group Play. <i>Journal of Autism and Developmental Disorders</i> , 45(12), 3836-3845.	1.Participants are younger than 11.
Nikopoulos, C. K. and Nikopoulou-Smyrni, P. (2008). Teaching Complex Social Skills to Children with Autism; Advances of Video Modeling. <i>Journal of Early and Intensive Behavior Intervention</i> , 5(2), 30-43.	1.Participants are younger than 11.
Gül, S. O. (2016). The Combined Use of Video Modeling and Social Stories in Teaching Social Skills for Individuals with Intellectual Disability. <i>Educational Sciences: Theory and Practice</i> , 16(1), 83-107.	1.Participants did not have a diagnosis of Autism.
Zeedyk, S. M., Davies, C., Parry, S. and Caldwell, P. (2009). Fostering social engagement in Romanian children with communicative impairments: the experiences of newly trained practitioners of Intensive Interaction. <i>British Journal of Learning Disabilities</i> , 37(3), 186-196.	4.Qualitative study

Appendix B

Table 8

Summary of Included Studies

Authors	N	Participants	Design	Measures of social skills	Intervention	Comparison condition	Adults delivering intervention
Ingersoll, Berger, Carlsen and Hamlin, 2017	19 (3 females, 16 males)	American full-time students (aged 12-20), at residential programmes for complex and multiple needs.	RCT. Randomised block design (between participants). Participants were stratified by age and paired.	Matson Evaluation of Social Skills for Individuals with Severe Retardation (MESSIER) Aberrant Behaviour Checklist-Residential (ABC-R).	Reciprocal Imitation Training Two individual, 10 minute sessions, 3 days a week for 5 months.	Equivalent number of 1:1 treatment sessions as treatment group but of usual support.	Teacher support staff received 2 weeks training and supervised by an author of the programme (Carlsen).
Ingersoll, Walton, Carlsen and Hamlin, 2013	4 (1 female, 3 males)	American students (aged 13-16), at a residential program for developmental disabilities.	Single-subject, multiple baseline design.	Direct observation using interval coding.	Reciprocal Imitation Training Between two and six individual, 20 minute	N/A	Teaching staff received 2 days training and video feedback on their implementation of RIT, from two authors of the

					sessions, per week for 10 weeks.		study (Ingersoll and Walton).
Mason, Rispoli, Ganz, Boles and Orr, 2012	1 (male)	American college student (age 19).	Multiple baseline single-case design.	Direct interval observation using likert scaling.	Video modelling Two, 50 minute sessions per week, for 5 weeks.	N/A	First author (Mason).
Whittington-Barnish, 2012	12 (4 females, 7 males)	American students (aged 14-18), at a public high school.	Single-subject, multiple baseline design.	Autism Social Skills Profile (ASSP) Social Responsiveness Scale (SRS)	Video modelling. One to two sessions per day, up to 6 sessions per week, for 4 weeks.	N/A	Research assistant, overseen by lead researcher (Whittington-Barnish).
Williamson, Casey, Robertson and Buggey, 2013	3 (1 female, 2 males)	American middle school students (aged 11-14), who received special education services.	Single-subject, multiple baseline design.	Direct observation using interval coding.	Video modelling. One individual session per day, for 3 weeks.	N/A	Trained school staff.

Appendix C

Weight of Evidence Ratings

Weight of Evidence A

Weight of Evidence A was assessed using two different coding protocols. Kratochwill et al.'s, (2003) protocol was used for group studies and Horner et al. (2005) for single case experimental designs. The coding protocols were adapted based on the review criteria and content of the studies within the review.

Group Studies

The Kratochwill et al., (2003) protocol was used for the Ingersoll, Berger, Carlsen and Hamlin (2017) study, as it employed a group design. The coding protocol was adapted in line with this review, removing: educational/clinical significance; identifiable components; replication; site of implementation and follow-up assessment.

Table 9

Alterations to Kratochwill et al., (2003) Coding Protocol

Section removed from protocol	Rationale for removal
Identifiable components	Review will consider the imitative components within interventions.
Replication	Comparison between each study's findings will be discussed within review.
Site of implementation	Inclusion criteria: all studies within the review are based in an educational setting.
Follow-up assessment	Group design study did not include follow-up assessment.
Outcomes (statistical analysis section included)	Outcomes are included within the effect size tables.

Table 10

Weight of Evidence A: Criteria for Group Studies

Criteria	Ratings
A: Measurement	3. All of the criteria are met: reliability of the measures is .85 or above; multiple sources and multiple methods of data collection are used; measures used are valid. 2. All of the criteria is met: reliability of the measures is .70 or above;

Criteria	Ratings
	<p>multiple sources and multiple methods of data collection are used.</p> <p>1. Reliability of the measures is .50 or above.</p>
<p>B: Comparison group</p>	<p>3. All of the criteria are met: comparison group are given an alternative intervention that is equivalent to the intervention group; counterbalancing of change agents; group equivalence established; equivalent mortality with low attrition.</p> <p>2. Comparison group are given an alternative intervention that is not equivalent to the intervention group. Two additional criteria are met: counterbalancing of change agents; group equivalence established; equivalent mortality with low attrition.</p> <p>1. Comparison group are not given an alternative intervention. One additional criteria met: counterbalancing of change agents; group equivalence established; equivalent mortality with low attrition.</p>
<p>C: Statistical analysis</p>	<p>3. All criteria are met: use of appropriate units of analysis; familywise/experiment wise error rate controlled (if applicable); a sufficiently large N; 75% of primary outcomes must be significant.</p> <p>2. All criteria met: use appropriate units of analysis; familywise/experiment wise error rate controlled (if applicable); and a sufficiently large N.</p> <p>1. Study must meet all criteria: use appropriate units of analysis; familywise/experiment wise error rate controlled (if applicable).</p>
<p>D: Educational/clinical significance</p>	<p>3. Three out of four criteria met: categorical diagnosis data; outcomes assessed via continuous variables; subjective evaluation; social comparison.</p> <p>2. Two out of four criteria met: categorical diagnosis data; outcomes assessed via continuous variables; subjective evaluation; social comparison.</p>

Criteria	Ratings
E: Fidelity	<p>1. One out of four criteria met: categorical diagnosis data; outcomes assessed via continuous variables; subjective evaluation; social comparison.</p> <p>3. Information provided through written materials or formal training. Two additional criteria met: ongoing supervision/consultation; coding sessions; audio/video tapes and use of a manual.</p> <p>2. Information provided through written materials or formal training; fidelity measured through one of the above criteria.</p> <p>1. One of the above criteria or use of a manual should be used.</p>

Table 11

Weight of Evidence A Ratings (Group Studies)

Author	Dimensions					Total WoE A
	A	B	C	D	E	
Ingersoll, Berger, Carlsen and Hamlin, 2017	3	3	1	3	3	2.6

Single Case Experimental Studies

For the studies which employed a single case experimental design, the Horner et al. (2005) protocol was used: Ingersoll, Walton, Carlsen and Hamlin, (2013); Mason, Rispoli, Ganz, Boles and Orr (2012); Whittington-Barnish, (2012) and Williamson, Casey, Robertson and Buggiey, (2013). The quality indicators used as the criteria for Weight of Evidence A are detailed in Table 1 of the Horner et al. (2005, p.174) paper. The descriptors provided numerical ratings between 1 and 3 for each category. Ratings for each criteria were summed and divided by 7 (total number of criteria) to provide a total Weight of Evidence A score for each study.

Table 12

Weight of Evidence A: Criteria for single case experimental studies

Criteria	Ratings
A: Description of Participants and Setting	3. All 3 of the following criteria are fulfilled: participants are described with sufficient detail to allow others to select individuals with similar characteristics (e.g., age, gender, disability,

Criteria	Ratings
	<p>diagnosis); the process for selecting participants is described with replicable precision; critical features of the physical setting are described with sufficient precision to allow replication.</p> <p>2. Two out of the criteria above are fulfilled.</p> <p>1. One of the criteria above is fulfilled.</p>
B: Dependent Variable	<p>3. All of the following criteria are fulfilled: dependent variables are described with operational precision; each dependent variable is measured with a procedure that generates a quantifiable index; measurement of the dependent variable is valid and described with replicable precision; dependent variables are measured repeatedly over time; 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%).</p> <p>2. Three to four of the above criteria are fulfilled.</p> <p>1. One to two of the above criteria are fulfilled.</p>
C: Independent Variable	<p>3. All of the following criteria are fulfilled: independent variable is described with replicable precision; independent variable is systematically manipulated and under the control of the experimenter; overt measurement of the fidelity of implementation for the independent variable is highly desirable.</p> <p>2. Two of the above criteria are fulfilled.</p> <p>1. One of the above criteria is fulfilled.</p>
D: Baseline	<p>3. All of the following criteria are fulfilled: the study includes a baseline phase that provides repeated measurement of a dependent variable; the study 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; baseline conditions are described with replicable precision.</p> <p>2. Two of the above criteria are fulfilled.</p> <p>1. One of the above criteria is fulfilled.</p>
E: Experimental Control/Internal Validity	<p>3. All of the following criteria are fulfilled: the design provides at least three demonstrations of experimental effect at three different points in time; the design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses); the results document a pattern that demonstrates experimental control.</p> <p>2. Two of the above criteria are fulfilled.</p> <p>1. One of the above criteria is fulfilled.</p>
F: External Validity	<p>3. Experimental effects are replicated across 3 or more participants and also either settings or materials to establish external validity.</p>

Criteria	Ratings
	2. Experimental effects are replicated across 3 or more participants. 1. Experimental effects are replicated by inclusion of at least 2 participants.
G: Social validity	3. All of the following criteria are fulfilled: the dependent variable is socially important; the magnitude of change in the dependent variable resulting from the intervention is socially important; implementation of the independent variable is practical and cost effective; social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts. 2. Two to three of the above criteria are fulfilled. 1. One of the above criteria is fulfilled.

Table 13

Weight of Evidence A Ratings (Single Case Experimental Studies)

Authors	Dimensions							Total WoE A
	A	B	C	D	E	F	G	
Ingersoll, Walton, Carlsen and Hamlin, (2013)	3	3	2	3	3	3	3	2.9
Mason, Rispoli, Ganz, Boles and Orr (2012)	3	2	1	3	3	2	1	2.1
Whittington-Barnish, (2012)	3	3	3	3	3	2	2	2.7
Williamson, Casey, Robertson and Buggiey, (2013)	3	3	2	3	3	3	1	2.6

Weight of Evidence B

Weight of Evidence B was scored for methodological relevance based on review-specific criteria. Group studies and single case designs required different criteria to determine WoE B.

Table 14

Weight of Evidence B Criteria (Group Studies)

Criteria	Ratings	Rationale
A. Randomisation	3. Participants are randomly assigned to the intervention or control group.	Randomisation reduces the effects of pre-existing group differences.

Criteria	Ratings	Rationale
	2. Participants are randomly assigned to treatment conditions in groups. 1. There is no randomisation in the assignment of participants to treatment conditions.	
B. Comparison Group	3. The control group receives alternative support to the intervention group, but in the same format. 2. The control group receives alternative support to the intervention group, but in a different format. 1. No alternative intervention is provided to the control group.	Providing the control group with support in the same format as the intervention group reduces the effects of social support on the intervention results.

WoE B scores for group studies were based on the sum of criteria A and B which was divided by two.

Table 15

Weight of Evidence B Ratings (Group Studies)

Author	Dimensions		Total WoE B
	A	B	
Ingersoll, Berger, Carlsen and Hamlin, 2017	3	3	3

Table 16

Weight of Evidence B Criteria (Single Case Design)

Criteria	Ratings	Rationale
Single Case Design	3. All criteria must be met: multiple baseline design; three demonstrations of intervention effect. 2. Three demonstrations of intervention effect.	Based on Horner et al.'s (2005) criteria.

Criteria	Ratings	Rationale
	1. Intervention effect demonstrated less than three times.	

WoE B scores for Single Case Designs were calculated by assigning scores of: 3 (high methodological relevance); 2 (medium methodological relevance); 1 (low methodological relevance).

Weight of Evidence C

Weight of Evidence C was based on topic relevance and calculated using review-specific criteria.

Table 17

Weight of Evidence C Criteria

Criteria	Ratings	Rationale
A. Implementer of intervention	3. Intervention is delivered by academic setting staff. 2. Intervention is delivered by both academic staff and researchers. 1. Intervention is delivered by researchers only.	Review aims to identify effective school-based interventions.
B. Identification of participants	3. Participants are selected through an explicit screening process to identify social skill development needs. 2. Participants are identified by staff based on their perception of the young person's social skill development needs. 1. No clear rationale for selection of participants.	Review aims to identify imitative interventions that increase social skills, therefore participants should have additional social skill needs.
C. Purpose of intervention	3. Intervention was planned to increase social skills. 2. Intervention was planned to address other communication skills.	Review aims to identify whether imitative interventions increase social skills.

Criteria	Ratings	Rationale
	1. No clear rationale for intervention.	
D. Setting	3. Intervention is carried out in a school setting. 2. Intervention is carried out in other academic settings e.g. college or university. 1. Intervention is carried out in any other setting e.g. home or clinic.	Review aims to identify effective school-based interventions.

Table 18
Weight of Evidence C Ratings

Authors	Dimensions				Total WoE C
	A	B	C	D	
Ingersoll, Berger, Carlsen and Hamlin, 2017	3	3	3	3	3
Ingersoll, Walton, Carlsen and Hamlin, (2013)	3	2	3	3	2.8
Mason, Rispoli, Ganz, Boles and Orr (2012)	1	2	3	2	2
Whittington-Barnish, (2012)	1	3	3	3	2.5
Williamson, Casey, Robertson and Bugghey, (2013)	3	1	2	3	2.3

Weight of Evidence C scores were based on the sum of criteria A, B, C and D which was divided by three.

Weight of Evidence D scores were based on the sum of WoE A, WoE B and WoE C scores which was divided by three.

Appendix D

Example of Coding Protocol for Group Designs

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

Full Study Reference in proper format:

Ingersoll, B., Berger, N., Carlsen, D. and Hamlin, T. (2017). Improving social functioning and challenging behaviors in adolescents with ASD and significant ID: A randomized pilot feasibility trial of reciprocal imitation training in a residential setting. *Developmental Neurorehabilitation*, 20(4), 236-246.

Intervention Name (description of study): Reciprocal Imitation Training

- Type of Publication:
- Book/Monograph
- Journal Article
- Book Chapter
- Other (specify):

1. General Characteristics

A. General Design Characteristics

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

- Completely randomized design
- Randomized block design (between participants, e.g., matched classrooms)
- Randomized block design (within participants)
- Randomized hierarchical design (nested treatments)

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

- Nonrandomized design
- Nonrandomized block design (between participants)
- Nonrandomized block design (within participants)
- Nonrandomized hierarchical design
- Optional coding for Quasi-experimental designs

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

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)

- Very high (explicitly stated)
- N/A
- Unknown/unable to code

B. Participants

Total size of sample (start of study): 20

Intervention group sample size: 10

Control group sample size: 10

C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment
- Unknown

D. Stage of Program

- Model/demonstration programs
- Early stage programs
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure
- Prior exposure
- Unknown

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

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

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

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

- Yes
- No

Unknown/unable to code

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

Yes

No

N/A

Unknown/unable to code

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

Yes

No

N/A

Unknown/unable to code

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

Yes validated with specific target group

In part, validated for general population only

No

Unknown/unable to code

Overall Rating for measurement: 3

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

B. Comparison Group

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

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

Attention placebo

Intervention element placebo

Alternative intervention

Pharmacotherapy

No intervention

Wait list/delayed intervention

Minimal contact

Unable to identify type of comparison

B2 Overall confidence of judgment on type of comparison group

Very low (little basis)

Low (guess)

Moderate (weak inference)

High (strong inference)

Very high (explicitly stated)

Unable to identify comparison group

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

- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

B4 Group equivalence established (select one of the following)

- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality

- Low attrition (less than 20 % for post)
- Low attrition (less than 30% for follow-up)
- Intent to intervene analysis carried out?
Findings_____

Overall rating for Comparison group: 3

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

C. Appropriate Statistical Analysis

Analysis 1: Independent t-test

- Appropriate unit of analysis
- Familywise/experimenter wise error rate controlled when applicable
- Sufficiently large N

Analysis 2: Mann Whitney U test

- Appropriate unit of analysis
- Familywise/experimenter wise error rate controlled when applicable
- Sufficiently large N

Analysis 3: Wilcoxon Signed Rank test

- Appropriate unit of analysis
- Familywise/experimenter wise error rate controlled when applicable
- Sufficiently large N

Overall rating for Statistical Analysis: 1

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

D. Educational/Clinical Significance

Study must provide evidence in support of the clinical significance for:

- Categorical Diagnosis Data
- Outcomes Assessed Via Continuous Variables
- Subjective Evaluation
- Social Comparison

Overall rating for educational/clinical significance

 3

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

E. Fidelity

Evidence of adherence is measured through

- Ongoing supervision/consultation
- Coding sessions
- Audio/Video tapes
- Use of a manual
- There is a descriptions of any adaptations

Overall rating for fidelity 3

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

Summary of Evidence

Indicator	Overall evidence rating 0-3	Description of evidence Strong Promising Weak No/limited evidence Or Descriptive ratings

General Characteristics		
Design		Randomised block design (between participants)
Type of programme		Intervention
Stage of programme		Established
Concurrent/ historical intervention exposure		Unknown
Key features		
Measurement	2	Strong evidence
Comparison group	3	Strong evidence
Appropriate Statistical Analysis	2	Promising evidence
Educational/clinical significance	3	Strong evidence
Fidelity	3	Strong evidence

Example of Coding Protocol for Single Case Research

[From The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education, Council for Exceptional Children, Horner, R. H., Carr, E. G., Halle, J., Odom, S. and Wolery, M. (2005)]

Full Study Reference in proper format:

Ingersoll, B., Walton, K., Carlsen, D. and Hamlin, T. (2013). Social Intervention for Adolescents with Autism and Significant Intellectual Disability: Initial Efficacy of Reciprocal Imitation Training. *American Journal on Intellectual and Developmental Disabilities*, 118(4), 247-261.

Intervention Name (description of study): Reciprocal Imitation Training

Type of Publication:

Book/Monograph

Journal Article

Book Chapter

Other (specify):

Section A: Description of Participants and Setting

A1. Participants are described with sufficient detail to allow others to select individuals with similar characteristics; (e.g., age, gender, disability, diagnosis, **race/ethnicity, medication**).

Yes

No

N/A

Unknown/Unable to Code

A2. The process for selecting participants is described with operational precision.

Yes

No

N/A

Unknown/Unable to Code

A3. Critical features of the physical setting are described with sufficient precision to allow replication.

Yes

No

- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3

Section B: Dependent Variable

B1. Dependent variables are described with operational precision.

- Yes
- No
- N/A
- Unknown/Unable to Code

B2. Each dependent variable is measured with a procedure that generates a quantifiable index.

- Yes
- No
- N/A
- Unknown/Unable to Code

B3. Measurement of the dependent variable is valid and described with replicable precision.

- Yes
- No
- N/A
- Unknown/Unable to Code

B4. Dependent variables are measured repeatedly over time.

- Yes
- No
- N/A
- Unknown/Unable to Code

B5. Data are collected on the reliability or inter-observer agreement associated with each dependent variable, and IOA levels meet minimal standards.

- Yes
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3

Section C: Independent Variable

- C1. Independent variable is described with replicable precision.
- Yes
 - No
 - N/A
 - Unknown/Unable to Code
- C2. Independent variable is systematically manipulated and under the control of the experimenter.
- Yes
 - No
 - N/A
 - Unknown/Unable to Code
- C3. Overt measurement of the fidelity of implementation for the independent variable is highly desirable.
- Yes
 - No
 - N/A
 - Unknown/Unable to Code

Overall Rating of Evidence: 2

Section D: Baseline

- D1. 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.
- Yes
 - No
 - N/A
 - Unknown/Unable to Code
- D2. Baseline conditions are described with replicable precision.
- Yes
 - No
 - N/A
 - Unknown/Unable to Code

Overall Rating of Evidence: 3

Section E: Experimental Control/Internal Validity

E1. The design provides at least three demonstrations of experimental effect at three different points in time.

- Yes
- No
- N/A
- Unknown/Unable to Code

E2. The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).

- Yes
- No
- N/A
- Unknown/Unable to Code

E3. The results document a pattern that demonstrates experimental control.

- Yes
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3

Section F: External Validity

F1. Experimental effects are replicated across participants, settings, or materials to establish external validity.

- Yes
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3

Section G: Social Validity

G1. The dependent variable is socially important.

- Yes

- No
- N/A
- Unknown/Unable to Code

G2. The magnitude of change in the dependent variable resulting from the intervention is socially important.

- Yes
- No
- N/A
- Unknown/Unable to Code

G3. Implementation of the independent variable is practical and cost effective

- Yes
- No
- N/A
- Unknown/Unable to Code

G4. Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.

- Yes
- No
- N/A
- Unknown/Unable to Code

Overall Rating of Evidence: 3

Average WoE A across the 7 judgement areas:

Sum of X / N = 20 / 7

X = individual quality rating for each judgement area

N = number of judgement areas

Overall Rating of Evidence: High (2.9)