



Are Extended Reality Interventions Effective in Helping Autistic Children to Enhance Their Social Skills? A Systematic Review

Rebecca Roberts¹ · James Stacey² · Susannah Jenner¹ · Eloise Maguire¹

Received: 16 November 2021 / Accepted: 5 April 2022 / Published online: 27 April 2022
© The Author(s) 2022

Abstract

Autistic children's social skills do not always align with those of their neurotypical peers and research suggests that this can negatively impact quality of life. This review aimed to assess the effectiveness of extended reality (XR) interventions in helping autistic children to enhance their social skills. Five electronic databases were systematically searched and seventeen studies were identified. The majority targeted social-emotional reciprocity and were of relatively low quality. There was insufficient evidence to determine whether effects were generalisable, sustained or important to autistic people. Research in this field is in its infancy and evidence of effectiveness should be viewed with caution. Future studies should aim for high-quality, theory-driven research, and involve autistic people to ensure meaningful outcomes.

PROSPERO ID: CRD42021229442

Keywords Autism · Social Skills · Augmented reality · Virtual reality · Assistive technology

Introduction

Autism and Social Differences

Effective social communication depends on a range of *social skills*, a term which is complex and has been variably defined across the literature, but is broadly considered to encompass verbal and non-verbal behaviours that enable positive social interactions (Gresham & Elliott, 1984). This multi-faceted concept includes innate skills, such as eye tracking and a preference for faces, as well as more complex, learned

behaviours, such as adjusting behaviour to suit particular social contexts (Spence, 2003). An exhaustive list of skills considered to be *social* is difficult to establish, due to the number of cognitive and behavioural processes drawn upon in any given social situation.

Autistic¹ children's social skills do not always align with those of neurotypical people. Their skills are different in terms of social-emotional reciprocity (e.g. social approach and initiation of interactions), nonverbal communicative behaviours (e.g. eye contact, and use of facial expressions), and the development, maintenance and understanding of relationships (e.g. reduced interest in peers and difficulties making friends); Diagnostic and Statistical Manual of Mental Disorders, 5th ed.; DSM-5; American Psychiatric Association [APA], (2013). These differences manifest differently for each individual (Happé et al., 2006) and can have negative secondary effects on wellbeing, learning and relationships. For example, social skills which do not align with neurotypical people have been shown to predict social anxiety (Bellini, 2006) and are associated with social rejection from peers (Ochs et al., 2001). This is not through lack of interest in

✉ Rebecca Roberts
rebecca.roberts@hmc.ox.ac.uk

James Stacey
james.stacey@oxfordhealth.nhs.uk

Susannah Jenner
susannah.jenner@hmc.ox.ac.uk

Eloise Maguire
eloise.ainger@hmc.ox.ac.uk

¹ Oxford Institute of Clinical Psychology Training and Research, Isis Education Centre, Warneford Hospital, University of Oxford, Oxford OX37JX, UK

² Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford OX37JX, UK

¹ Identity first language is used within this article as this is the preferred language of many adults on the autism spectrum, i.e. 'autistic individual' rather than 'individual with autism' (Kenny et al., 2016). However, there is a lack of information about the preferred language of autistic children.

social contact (Jaswal & Akhtar, 2019); autistic young people often express a desire for social interaction and report being more lonely than their neurotypical peers (Bauminger & Kasari, 2000). Furthermore, social differences in autistic people can have a negative effect on academic and occupational outcomes (Howlin, 2003; Welsh et al., 2001).

Theories of Social Difference in Autism

A number of theories might explain the social differences observed in autism. One theory proposes that autistic people can find it harder to make predictions about others' mental states (theory of mind) and so may find it difficult to encode socially relevant information (Baron-Cohen et al., 1985). Alternatively, it may be that difficulty understanding others' interpretations of the world is due to an underlying difference in executive functioning, such as rule learning (J. Russell, 1997). This can account for the finding that some autistic children struggle with theory of mind style tasks, even when the social element is removed (Russell et al., 1991). In particular, predictive processing might be different in autism. Pellicano and Burr (2012) propose that autistic people perceive the world more accurately, as they are less biased by prior experience when making predictions. At times this may be a strength, but could make some elements of social interaction more difficult, such as predicting when another person will speak (Stark et al., 2021). Another cognitive theory of autism is 'weak central coherence' (Frith, 1989) which suggests that autistic people have a preference for local processing rather than striving for overall meaning. This may lead to differences in the interpretation of social behaviours. However, more recent evidence suggests that this theory sits alongside social differences, rather than explains them (Happé & Frith, 2006). These cognitive theories are in line with evidence that social desire is not affected by autism (Happé & Frith, 1996).

In contrast, social motivation theory proposes that autistic people have differences in the psychological dispositions and biological mechanisms that bias neurotypical individuals to orient towards social stimuli, seek and derive pleasure from social interactions, and work to maintain social bonds (Chevalier et al., 2012). However, this contrasts with many accounts of autistic people, overlooks the role of other people's perceptions and responses, and ignores the many other explanations of social differences in autism (Jaswal & Akhtar, 2019). Furthermore, this theory has led to differences such as limited eye contact being prioritised as intervention targets, despite evidence for the positive effects of gaze aversion on concentration and regulation of emotions (Robledo et al., 2012).

Social Skills Interventions

A number of interventions have sought to teach social skills to autistic children, due to the secondary benefits on

wellbeing and quality of life shown by previous research (McConnell, 2002). A range of outcomes have been targeted, such as turn taking, initiating interaction, recognising emotions and understanding theory of mind (Begeer et al., 2011; Matson et al., 2007; Rieth et al., 2014). Well-established interventions include Social Stories™ (Karkhaneh et al., 2010), video modelling (Sng et al., 2014), pivotal response training (Koegel et al., 2016), social skills groups (Reichow et al., 2012) and peer-mediated behavioural interventions (Laushey & Heflin, 2000). There is increasing evidence that, at least in some contexts, these interventions can help autistic children to utilise neurotypical social skills but the duration, generalisability and precise mechanism of change remains unclear (McConnell, 2002). Furthermore, little acknowledgement has been given to the fact that it may be effortful, and perhaps emotionally draining for autistic people to operationalise these skills.

When considering social skills interventions, it is important to acknowledge that social differences in autism do not necessarily equate to a 'difficulty'. Many argue that difficulties arise due to a lack of accommodation of autistic differences in a neurotypical society (Brownlow, 2010), and emerging research is suggestive of a 'double empathy problem', whereby communication breakdowns between autistic and non-autistic people are a two-way issue resulting from disjuncture in reciprocity between the two differently disposed social actors (Edey et al., 2016; Milton, 2012; Shepard et al., 2016). Social skills training can therefore perpetuate the idea that autistic people need to adapt, without consideration of the role of the contexts in which autistic people live, learn and work in understanding, supporting and responding to autistic communication. However, many autistic people already find ways to overcome social challenges and seek support in doing this (Cresswell et al., 2019). Social skills interventions are one way in which services can support people to overcome these challenges, particularly in circumstances where it is not possible to adapt the environment e.g. public places.

Extended Reality Technologies

Over the last three decades, social skills interventions have begun to capitalise on the potential advantages of digital technologies. Computerised approaches can support social skills in situ (Vygotsky & Luria, 1994) or improve traditional training approaches, due to the attentional and motivational advantages of interactive technologies (Golan & Baron-Cohen, 2006). More recently, these advantages have been exploited using extended reality (XR), i.e. the merging of physical and virtual realities (Riva et al., 2016). This encompasses augmented reality (AR), in which virtual information is overlaid onto the real world, virtual reality (VR), in which users are fully immersed in a simulated environment and

mixed reality (MR), in which digital and real-world objects interact in real-time. By merging physical and virtual realities, XR could reduce the salience of potentially distressing stimuli, making interactions less overwhelming for some autistic people, thus providing an effective space to experiment with social skills (Thye et al., 2018). In virtual social settings, there is no risk of embarrassment in front of one's peers, yet the degree of realism may increase the likelihood of generalisability (Strickland et al., 1996).

Extended reality interventions have been successfully implemented in a number of therapeutic contexts including stroke rehabilitation (Laver et al., 2017) and exposure therapies for anxiety disorders (Bouchard et al., 2017; Vincelli et al., 2003). Applications for autistic people are also emerging, although the majority of interventions to date have used virtual learning environments (VLEs), a non-immersive technology in which users interact with simulated social situations using a desktop computer (e.g. Didehbani et al., 2016). More immersive technologies tend to require a head-mounted display (HMD) or Smartglasses, computer glasses that change what the wearer sees. Despite the potential challenges for those with tactile hypersensitivities, they have been shown to be well-tolerated, enjoyable and engaging in this population (Keshav et al., 2017; Newbutt et al., 2016), and effective in teaching other skills, such as taking public transport (Simoes et al., 2018).

Aims of this Review

Previous reviews have shown that XR interventions are feasible and sometimes beneficial for teaching autistic children a range of skills, although recognise that the quality of the literature is varied (e.g. Berenguer et al., 2020; Mesa-Gresa et al., 2018). However, no review to date has investigated the specific potential of XR for social skills. Furthermore, previous reviews of VR interventions have tended to include non-immersive technologies, such as VLEs. This review takes a focused approach by reviewing the evidence base for interventions in which one's perception of reality is altered. Social skills interventions have most commonly been implemented in children but we chose to include young adults, based on evidence that brain maturation, particularly that of the prefrontal cortex which coordinates the goal directed behaviour needed to implement social skills, continues to develop into a person's early twenties (Johnson et al., 2009). This is also in line with current plans to extend children's mental health services in England up to 25 years (NHS Long Term Plan, 2019). Due to the infancy of this research area, we chose not to restrict inclusion on the basis of study design.

It was considered important to understand the effectiveness of these interventions because of the secondary benefits social skills can have on indicators of wellbeing, such as

increased peer interactions (Mandelberg et al., 2014) and reduced anxiety and low mood (Hillier et al., 2011; Schohl et al., 2014). Secondly, the realism of immersive VR offers a unique potential to improve upon existing social skills interventions which do not always generalise to real contexts (Bellini et al., 2007). Thirdly, XR is at a point of rapid expansion, with VR headsets becoming increasingly widespread. If social skills interventions are going to capitalise on this expansion, it is important to consider whether there is any good quality evidence for their effectiveness and whether investment in these technologies will actually lead to meaningful benefits for autistic young people.

To this end, this review aims to answer the following questions:

1. Is there any evidence that extended reality interventions are effective in helping autistic children and young people to enhance their social skills?
2. What is the quality of the evidence for observed effects?
3. To what extent have observed effects been shown to lead to meaningful change for autistic people?

Method

This quantitative systematic review uses a synthesis without meta-analysis methodology and information is reported in line with Cochrane guidance (Garritty et al., 2021). The protocol was pre-registered with PROSPERO (ID: CRD42021229442; available at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=229442). Five online databases (PsychInfo, Medline, Pubmed, EMBASE, CINAHL) were searched using key terms listed in Appendix 28. All retrieved datasets were downloaded into a Zotero library and exported to Excel for first author screening by title and abstract, then by full text. Another author (EM) screened 26.5% of studies at initial screening and 25% of studies at full text screening. All texts accepted at final screening were forward and backward citation searched using Scopus.

Screening and Selection

Studies were screened for eligibility according to the following inclusion criteria:

- The paper was published in English, in a peer-reviewed journal to ensure a minimum quality level.
- The paper was published after 1990, when AR was developed and VR headsets became widely commercially available.
- The paper included an active intervention using XR. Studies described as VR but which used a desktop computer or motion based video game were not included for

review as they do not to extend one's perception of reality.

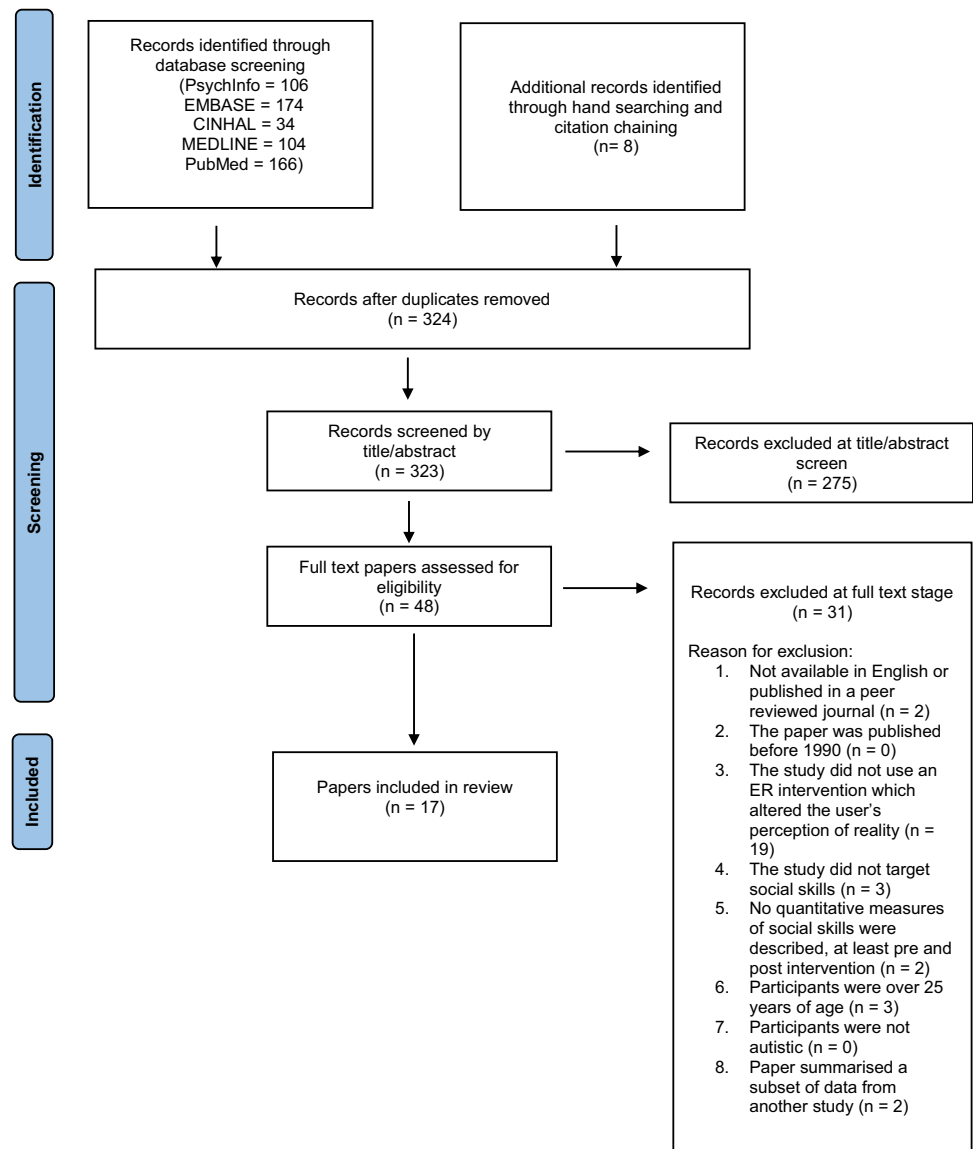
- The study targeted one or more social skill.
- The paper reported a quantitative measure of social skills, at least pre and post intervention. This included in vivo outcomes, as well as validated and idiosyncratic measures of social skills.
- Study participants were children and young adults up to the age of 25.
- Study participants were autistic.

Where papers reported on more than one study, they were included based on the study which met the inclusion criteria. Borderline cases were discussed amongst the researchers.

Study Selection

The initial search identified 592 records, 324 of which were retained after duplicates were removed (Fig. 1). Two hundred seventy-five studies were excluded at the abstract stage, with a 96.5% inter-rater reliability ($\kappa = .87$). Thirty-one studies were excluded at the full text screen, with a 91.7% inter-rater reliability ($\kappa = .83$). Disagreements were resolved through further screening at the full text stage and discussion with other authors. Seven additional records were identified through hand searching and citation chaining (Scopus search conducted May 3, 2021). A total of 17 studies were included in the review and a number of data items were extracted from each study. Where it was not possible to extract data directly, study authors were contacted, or available information was used to calculate an effect size. 50% of data extraction was duplicated by a second reviewer (EM).

Fig. 1 Study selection process



Data Synthesis

Data were synthesised and organised according to the social communication criteria of DSM-5's autism diagnostic criteria (APA, 2000), as the effect of XR on specific social skills was the primary area of interest. In line with Siddaway et al. (2019), findings were integrated and critiqued, not merely summarised. The high number of single case studies meant that it was not possible to conduct a meta-analysis but key information is summarised in tables, with a summary of effect (Cohen's *d*) reported where enough information was provided in the paper or could be obtained through correspondence with the study authors. All studies were included for synthesis, irrespective of quality, but due to the range of study quality, summary plots were not developed, to avoid misrepresentation of the evidence base.

Quality Assessment

The quality and risk of bias of each study was evaluated using a quality checklist for healthcare intervention studies, developed by Downs and Black (1998). The tool was selected due to its applicability to both randomised and non-randomised studies. All papers were rated by the first author and 30% were duplicated by another author (EM), with an interrater reliability of 90.4% ($\kappa = 0.86$). Discrepancies were resolved through discussion. Some items were not applicable to small-N research and so a summary score relative to the number of items assessed was generated for each paper to allow for direct comparison across studies. This was calculated by dividing the score obtained by the total possible score according to study design. A risk of bias score was generated by the same method, using only items assessing bias (14 to 26). A full summary of the assessment approach is provided in Appendix 31. Given that a number of items were not applicable to single case and small-N studies, the Single Case Experimental Design (SCED) Scale (Tate et al., 2008) was also used to qualitatively assess these studies (Appendix 32).

Results

Study Characteristics

Across the 17 studies (summarised in Table 1), 225 participants (15.5% female) aged 2 to 16 were included. Sample sizes ranged from one to 94 participants ($M = 13.24$, $SD = 18.02$). All participants were autistic however, only 11 studies reported that this was a clinical diagnosis according to DSM and only three confirmed diagnosis using a

screening tool: Ip et al. (2018) using the Childhood Autism Spectrum Test (Williams et al., 2005); Liu et al. (2017) using the Social Communication Questionnaire (Chandler et al., 2007); and Lorenzo et al. (2019) using the Autism Spectrum Inventory (Rivière, 2002). Nine studies specified an intelligence quotient (IQ), with a mean or lower limit of at least 70. The studies used a range of designs, although a majority (10) were small-N, with either multiple baseline or pre-post measures. Only five studies used a control (Crowell et al., 2020; Herrero & Lorenzo, 2020; Ip et al., 2018; Lorenzo et al., 2016, 2019). Seven studies explicitly reported being at the preliminary or feasibility stage of the research.

Fourteen unique interventions were evaluated (seven VR; six AR; one MR) and are described in Table 2. Seven adapted existing social skills interventions, such as video modelling (Chen et al., 2016), collaborative game play (Mora-Guiard et al., 2017) and Social Stories™ (Herrero & Lorenzo, 2020). Rationale for the use of XR, included distraction reduction; the potential for engagement and sustained attention; freedom to practice social interactions without risk; and customisability i.e. the capacity to emphasise social stimuli and adapt difficulty level. Interventions most often took place in schools, and ranged from one 15 minute session (Crowell et al., 2020) to eighty 25-min sessions (Lorenzo et al., 2013). Some studies gave rationale for these design decisions, such as short sessions due to attentional capacity, but the majority gave no justification for the frequency or duration of XR use. Adverse outcomes were not routinely assessed but one study commented that some children had difficulties with restlessness and tolerance of the headset (Ravindran et al., 2019) and another used 'reinforcement candy' to encourage 'emotional stability' during the intervention (Cheng et al., 2015).

Outcomes

Social-Emotional Reciprocity

Fifteen studies targeted social-emotional reciprocity: six using AR, seven using VR and two using MR. Specific targets included identifying and responding to social greetings; motivation to communicate with a conversation partner; understanding and use of social initiations; and sharing emotions and affect with others. Given the multiple countries in which these studies took place, it is possible that reciprocal behaviours considered socially appropriate differed according to cultural norms. Participants ranged from 6 to 16 years, although the majority of participants were primary school aged.

AR interventions included Smartglasses, which gave real-time prompts in social situations (Liu et al., 2017; Sahin et al., 2018; Vahabzadeh et al., 2018), a motion based game,

Table 1 Summary of included studies

Study	Design	Country of study	Name of intervention (AR/VR/MR)	Autism diagnosis and assessment criteria	Age range of participants (years)	IQ	Total sample size included in analysis (% female)	Social skills targeted	Measure of social skills (validated?)
Liu et al. (2017)	Single case series, pre-post	USA	Brain Power System - Face Game and Emotion Game (AR)	Clinical diagnosis according to DSM-5	8–9	Not specified	2 (0%)	Social emotional reciprocity and non-verbal communication—social withdrawal and non-verbal communicative behaviours e.g. eye contact	ABC (✓) and caregiver report (x)
Sahin et al. (2018)	Single case study	USA	Empowered Brain - Face2Face Module (AR)	Clinical diagnosis but criteria not specified	13	Not specified	1 (0%)	Social emotional reciprocity—social awareness, cognition, communication, motivation and restricted interests	SRS-2 (✓)
Vahabzadeh et al. (2018)	Single case series, pre-post	USA	Empowered Brain - Face2Face Module (AR)	Clinical diagnosis according to DSM-IV or DSM-V.	6.7–8.8	Not specified	4 (0%)	Social emotional reciprocity—a range of behaviours indicating social reciprocity	ABC (✓)
Lee (2020)	Multiple baseline experimental design	Taiwan	Real-time skeletal tracking system (AR)	Clinical diagnosis according to DSM-IV-TR	7–9	$M = 102.33$	3 (33.3%)	Social emotional reciprocity—understanding of social greetings	Novel situation based questions (x); Idiosyncratic ratings of role play behaviours (x)
Lee et al. (2018)	Multiple baseline experimental design	Taiwan	Augmented reality concept map training system (AR)	Clinical diagnosis according to DSM-IV-TR	8–9	$M = 93.3$	3 (33.3%)	Social emotional reciprocity—understanding of social greetings	Novel situation based questions (x); idiosyncratic ratings of role play behaviours (x)

Table 1 (continued)

Study	Design	Country of study	Name of intervention (AR/VR/MR)	Autism diagnosis and assessment criteria	Age range of participants (years)	IQ	Total sample size included in analysis (% female)	Social skills targeted	Measure of social skills (validated?)
Lorenzo et al. (2019)	Case controlled, pre-post (non-randomised control)	Spain	Quiver Vision (AR)	Clinical diagnosis according to DSM-V and score > 30 on IAS	2–6	7 participants had a diagnosis of Developmental Delay, Intellectual Disability or Language Disorder	11 (9.1%)	Social emotional reciprocity—unclear target behaviour but domains of the IDEA include 'social relationship'; 'joint reference skills'; 'intersubjective and mentalistic skills'	IDEA ¹
Cheng et al. (2015)	Multiple baseline experimental design	Taiwan	3D-SU System (VR)	Diagnostic criteria not specified	10–12	$M = 82$	3 (0%)	Social emotional reciprocity—understanding of neurotypical social interactions	Novel situation-based questions (x)
Herrero & Lorenzo (2020)	Randomised controlled trial (parallel group design)	Spain	HMD Immersive VR System (VR)	Diagnostic criteria not specified	8–15	Not specified	14 (14.3% of experimental group, control not specified)	Social emotional reciprocity and non-verbal communication—social reciprocity, flexibility to change and non-verbal communication skills	Idiosyncratic ratings of in situ social reciprocity (x)
Ip et al. (2018)	Randomised controlled trial (parallel group design)	Hong Kong	Cave Automatic Virtual Environment (CAVE; VR)	Clinical diagnosis but diagnostic criteria unknown (mean score on Childhood Autism Spectrum Test above cutoff, 19.1)	7–10	95 (non-verbal)	72 (11.1%)	Social emotional reciprocity and understanding relationships—social reciprocity; emotion recognition; emotion regulation and everyday social skills e.g. maintaining relationships	PEP-3 (✓); Adapted Faces test (x); Eyes test (✓); ABAS-II (✓)

Table 1 (continued)

Study	Design	Country of study	Name of intervention (AR/VR/MR)	Autism diagnosis and assessment criteria	Age range of participants (years)	IQ	Total sample size included in analysis (% female)	Social skills targeted	Measure of social skills (validated?)
(Lorenzo et al., 2013)	Quasi-experimental, pre-post (no control)	Spain	Immersive Virtual Reality System (VR)	Asperger's syndrome, diagnostic criteria not specified	8–15	'Average Intelligence Levels' stated but not specified	20 (20%)	Social emotional reciprocity—no one primary outcome measure	Idiosyncratic ratings of in situ performance (x); Teacher ratings of classroom behaviours (x)
(Lorenzo et al., 2016)	Randomised controlled trial (parallel group design)	Spain	Immersive Virtual Reality System (VR)	Diagnosis according to DSM-V criteria	8–12	Not specified	40 (27.5%)	Social emotional reciprocity—a range of behaviours indicating social reciprocity	Idiosyncratic ratings of in situ performance (x); Teacher ratings of classroom behaviours (x)
Ravindran et al. (2019)	Single case series, pre-post	USA	Floreo's Joint Attention Module (VR)	Diagnosis confirmed by a family member	9–16	Not specified	12 (16.7%)	Social emotional reciprocity and non-verbal communication—frequency of interactions and eye contact	Joint attention assessment (x)
Crowell et al. (2020)	Randomised controlled trial (crossover ¹ design)	Spain	Lands of Fog (MR)	Clinical diagnosis according to DSM-IV-TR	8–12	> 70	18 (16.7% assumed)	Social emotional reciprocity—social initiations and responses in a collaborative game	Frequency of in situ social initiations and responses (x)
Mora-Guiard et al. (2017)	Quasi-experimental, repeated measures (no control)	Spain	Lands of Fog (MR)	Clinical diagnosis according to DSM-IV-TR	10–14	> 70	10 (0%)	Social emotional reciprocity—social initiations, acts and responses in a collaborative game	Frequency of in situ social initiations, acts and responses (x)
Tsai et al. (2020)	Multiple baseline experimental design	Taiwan	Cave Automatic Virtual Environment (CAVE; VR)	Clinical diagnosis according to DSM-5 criteria	7–9	$M = 87$	3 (0%)	Social emotional reciprocity—understanding of social initiations in response to others' expressions	Novel situation-based questions (x); idiosyncratic rating of role play behaviours (x)

Table 1 (continued)

Study	Design	Country of study	Name of intervention (AR/VR/MR)	Autism diagnosis and assessment criteria	Age range of participants (years)	IQ	Total sample size included in analysis (% female)	Social skills targeted	Measure of social skills (validated?)
Chen et al. (2015)	Multiple baseline experimental design	Taiwan	Augmented Reality Self-Facial Modelling Learning System (AR)	Clinical diagnosis according to DSM-IV-TR	10–13	$M = 101$	3 (33.3%)	Understanding relationships—aimed to improve emotion recognition	Idiosyncratic measure of ability to identify facial expressions and emotions (x)
Chen et al. (2016)	Multiple baseline experimental design	Taiwan	AR with Video Modelling Storybook Learning System (AR)	Clinical diagnosis according to DSM-IV-TR	11–13	$M = 103.66$	6 (17%)	Understanding relationships—aimed to improve emotion recognition	Idiosyncratic measure of ability to identify facial expressions and emotions (x)

¹It was not possible to determine whether this is a validated measure due to it having been published in Spanish

which provided opportunities to practice reciprocal behaviours (Lee, 2020), and a concept map enhanced with AR to aid learning (Lee et al., 2018). One study used an AR smartphone app to enhance child-therapist interactions, but the specific effects of AR were unclear (Lorenzo et al., 2019). VR interventions were less varied, with the majority simulating social situations in which participants could practice specific reciprocal behaviours (Cheng et al., 2015; Herrero & Lorenzo, 2020; Ip et al., 2018; Lorenzo et al., 2013, 2016; Ravindran et al., 2019). One study utilised a specific role playing game within the VR (Tsai et al., 2020). Conversely, the MR intervention (Crowell et al., 2020; Mora-Guard et al., 2017) used mixed reality objects to digitally emphasise the benefits of collaboration with a partner.

Outcomes were most commonly measured by presenting participants with a range of social situations (sometimes in the form of Social Stories™) and asking them to describe, and/or demonstrate, how they might respond (Cheng et al., 2015; Herrero & Lorenzo, 2020; Lee, 2020; Lee et al., 2018; Tsai et al., 2020). These studies all concluded that social-emotional reciprocity could be improved using their specific XR intervention however, the criteria for rating responses was not validated or tested for age appropriateness using a non-autistic sample. Furthermore, one study reported that participants were given prompts when responding but did not specify a protocol for this (Cheng et al., 2015). Another reported improvements in social reciprocity but assessed behaviours using a novel and unvalidated joint attention assessment (Ravindran et al., 2019). A number of studies measured outcomes by observing changes in behaviours during the XR intervention. Again, it was unclear whether the coded behaviours were age-appropriate, or had any relation to social skills in the real world. Of studies measuring in situ behaviours, two concluded that behaviours improved after the intervention (Herrero & Lorenzo, 2020; Mora-Guard et al., 2017) and one found no impact of XR on social reciprocity when compared to an active control (Crowell et al., 2020). Others did not set clear hypotheses regarding primary outcome measures, making it difficult to draw conclusions on their efficacy (Lorenzo et al., 2013, 2016). Only four of the studies targeting social-emotional reciprocity utilised validated outcome measures (Ip et al., 2018; Liu et al., 2017; Sahin et al., 2018; Vahabzadeh et al., 2018): the Aberrant Behaviour Checklist (ABC; Aman et al., 1985), Social Responsiveness Scale (SRS-2; Constantino & Gruber, 2012) and PEP-3 Psychoeducational Profile (Schopler et al., 2005). All of these studies showed an improvement in scores before and after the intervention, with a maximum follow up of 2 weeks (Ip et al., 2018). However, only in one instance was this in comparison to a control (Ip et al., 2018). One measure was unsuitable for the age of the study participants (Ip et al., 2018) and another indicated a 100% reduction in social difficulties (Liu et al., 2017) despite theoretical understandings

Table 2 Description of AR, VR and MR interventions and their reported purpose

Intervention	Description
AR	
AR based Self-Facial Modelling Learning System	This system is designed for children to interact with while wearing masks which represent particular emotions. AR superimposes facial expressions onto a participant's view of themselves, to show the emotion most appropriate to the situation. This augmentation of the participant's view of themselves is thought to promote learning.
AR Video Modelling With Storybook Learning System	A tablet computer is used to augment a physical storybook with short video clips, which interactively emphasise social details. This draws attention to social elements of the story and is thought to be more engaging than traditional video modelling approaches.
Real Time Kinect Skeletal Tracking System	This system augments a virtual character with a user's body gestures and superimposes a range of real life backgrounds. This enables children to practice a range of body gestures which might be appropriate to different situations e.g. shaking hands or waving arms. The primary benefit is that it is real time but without the anxiety associated with real world interactions.
AR Concept Map Training System	A tablet computer is used to augment a physical concept map with 3D greeting animations and graphics to indicate relationships between people and degrees of closeness in relationships. The purpose is to help children to visually learn about basic social relations and appropriate greeting behaviours.
Empowered Brain/ Brain Power System	Children wear a pair of Smartglasses which can evaluate looking behaviours and superimpose digital text and images onto the world, so that social cues are given in real time according to the user's ability. It includes a number of applications: <i>Face Game (also termed Face2Face)</i> uses an AR game to draw attention to human faces with visual overlays such as arrows and cartoon masks which gradually fade. Users score points by directing their gaze towards their caregiver's face. <i>Emotion Game</i> asks children to identify the facial expression of the person in front of them by tilting their head towards the correct emoticon. Points are gained for correct answers.
Quiver Vision	A smartphone application is used to augment real life pictures so that they are perceived as 3D objects. In Lorenzo et al. (2019), participants engage in activities which involve interacting with the AR components alongside the therapist, such as touching different AR objects according to the therapist's instructions.
VR	
3D-SU System	Users view and interact with animated social events using an immersive VR headset. Based on their responses to questions about the events, users are rewarded with a message and applause or instructed to try again. Using a headset is thought to reduce distraction and encourage sustained attention towards social events.
Cave Automatic Virtual Environment (CAVE) VR	Users are immersed in a number of social scenarios within a 'cave' of multiple large projection screens. This allows children to practice their responses free from risk of embarrassment. Ip et al. (2018) employ this technology to allow children to practice responses in a range of situations. They are supported by a trainer if needed and a relaxation scenario is available if the experience becomes distressing. Tsai et al. (2020) enhance the CAVE with motion control cameras. Users are encouraged to interact with avatars within the environment in specific role playing games.
Floreo - Joint Attention Module	A smartphone is fitted to a Google Cardboard headset, creating a VR experience. Users work through a number of 'learning cards' which require the demonstration of joint attention skills. A virtual avatar initiates and responds to joint attention bids and can prompt where necessary. A monitoring therapist can observe using a tablet computer and provide further support.
Immersive Virtual Reality System	Users are immersed in a number of scenarios using two large projection screens and movement detection software. Lorenzo et al. (2013) use this system to allow children to rehearse their responses to a range of classroom based tasks such as initiating conversation with a classmate. In Lorenzo et al. (2016), participants follow an evaluator's guidance on how to behave in a range of social situations e.g. attending a birthday party. This approach is thought to be beneficial due to its realism and the capacity to present information clearly without irrelevant distractors.

Table 2 (continued)

Intervention	Description
HMD Immersive VR System	Users are immersed in two school environments which are viewed through a head mounted display. Children can socially interact with a number of avatars, whose responses are controlled by the researchers. The system is reported to be highly immersive and flexible to allow for adaptation and promote real world generalisability.
MR	
Lands of Fog	Users collaborate with a partner to hunt for insects in a dense layer of virtual fog using a physical butterfly net. Explorative and collaborative efforts are rewarded with an upbeat tune or new feature. The approach is thought to be effective in encouraging collaboration as users can clearly see the actions of themselves and others, unlike in traditional collaborative game play.

of autism as a lifelong neurodevelopmental difference characterised by differences in social communication. One study measured outcomes using the Autism Spectrum Inventory (IDEA; Rivière, 2002). It is unclear whether this is validated as the measure is in Spanish, but no improvements were observed when compared with an active control, and again, the theoretical justification for use of this measure was limited (Lorenzo et al, 2019).

Overall, a number of studies have targeted elements of social-emotional reciprocity and almost all claim evidence of effectiveness (see Table 3). However, there is significant risk of bias in how outcomes were measured and the better quality studies tended to be single case. It is therefore difficult to draw conclusions about the efficacy of XR for improving this aspect of social communication. Furthermore, despite some qualitative reports, the studies did not attempt to determine generalisability and lack of longer-term follow-up makes it difficult to determine the extent to which improvements were sustained.

Non-verbal Communicative Behaviours

No studies targeted non-verbal communication in isolation but three studies (two VR and one AR) included non-verbal behaviours as outcome measures. Herrero and Lorenzo (2020) used an idiosyncratic measure of ‘non-verbal behaviours’ (e.g. use of facial cues, imitation and gestures) to suggest that use of their VR intervention was beneficial for 7- to 12-year-olds. However, all but one of the participants were rated as having ‘fair’ or ‘good’ non-verbal communication prior to the intervention and the measure used was not validated. Two studies concluded that their respective interventions increased eye contact: Ravindran et al. (2019) coded the amount of eye contact used by the 9- to 16-year-old participants during a joint attention assessment and Liu et al. (2017) utilised an idiosyncratic caregiver measure.

Given the limited number of studies which have investigated non-verbal communicative behaviours, it is difficult

to draw conclusions about XR’s efficacy for this domain of social communication. This is compounded by use of unvalidated measures, small sample sizes and unclear theoretical rationale for anticipated improvements in non-verbal behaviours. Only one study measured generalisability (caregiver report).

Developing, Maintaining and Understanding Relationships

Two AR studies primarily targeted the development, maintenance and understanding of relationships (Chen et al., 2015, 2016) and one VR study included some relational outcome measures (Ip et al., 2018). The AR studies targeted emotion recognition which was measured according to participants’ ability to identify emotions in a story. According to an unvalidated assessment method, participants improved between baseline and follow up. Informal parent reports suggested that improvements corresponded with real world change in ability to identify emotions however, this was not formally measured. In contrast, Ip et al. (2018) found no evidence of improvements on this domain after use of VR, when compared with a waitlist control. They targeted the application of social skills to real life (e.g. ability to maintain relationships), measured using a domain of the Adaptive Behaviour Assessment System (ABAS-II; Harrison & Oakland, 2003), and emotion recognition, which was measured using the Eyes and Faces Tests (Baron-Cohen et al., 1997, 2001). Mixed outcomes from this small number of studies mean it is not possible to conclude whether XR is effective for improving relational social communication.

Related Skills

Two studies included outcomes which would theoretically benefit social skills but do not fall under the DSM-5 social communication criteria (APA, 2013), such as emotion regulation (Ip et al., 2018) and flexibility to changes during social situations (Herrero & Lorenzo, 2020). When compared with

Table 3 Summary of results

Study	Quality rating	Risk of bias rating	Effect of intervention on social skills (+ indicates improvement and – indicates lack of improvement)
Liu et al. (2017)	0.61	0.67	Observed reduction in social withdrawal after using the Brain Power System (+) Carer reported increases in non-verbal communicative behaviours post-intervention (+)
Sahin et al. (2018)	0.56	0.83	Observed reduction in ratings of social skills difficulties measured by the SRS-2 after using Empowered Brain (+)
Vahabzadeh et al. (2018)	0.61	0.67	Observed reduction in social withdrawal after using Empowered Brain, at two time points (+)
Lee (2020)	0.58	0.5	Statistically significant mean improvement in understanding of social greetings at maintenance in comparison to baseline (+)
Lee et al. (2018)	0.53	0.33	Statistically significant mean improvement in understanding of social greetings at maintenance in comparison to baseline (+)
Lorenzo et al. (2019)	0.67	0.62	No significant difference between pre- and post- scores across all dimensions of the IDEA (-) No significant difference in post-intervention ratings between the intervention group and non-digital control for all items compared (-)
Cheng et al. (2015)	0.56	0.5	Observed increases in ability to understand social initiations across baseline, intervention and maintenance (+)
Herrero & Lorenzo (2020)	0.31	0.23	Observed improvements in social reciprocity and inflexibility to change post-intervention for all participants (+) Observed improvements in non-verbal communication post-intervention for 6 of 7 participants (+)
Ip et al. (2018)	0.48	0.54	Statistically significant improvement social reciprocity post-intervention (+) ($d = 0.50$) Statistically significant improvement in emotion regulation post-intervention (+) ($d = 0.39$) No significant difference in emotion recognition (Faces Test) post-intervention (-) ($d = 0.26$) No significant difference in emotion recognition (Eyes Test) post-intervention (-) ($d = 0.19$) No significant difference in everyday social skills post-intervention (-) ($d = 0$) Statistically significant effect of group (intervention vs waitlist control) on improvements in emotion regulation (+) ($d = 0.54$) and social reciprocity (+) ($d = 0.67$)
Lorenzo et al. (2016)	0.37	0.46	No clear primary outcome measure. Study reported greater overall improvements in the intervention group, compared with the non-immersive control (+)
Ravindran et al. (2019)	0.42	0.33	Observed increases in frequency of interactions and amount of eye contact after using the VR system (+)
Crowell et al. (2020)	0.46	0.33	No significant difference in social initiations using MR, compared with non-digital control (-) ($d = 0.06$) Statistically significantly more social responses in control condition, compared with MR intervention (-) ($d = 0.14$)
Mora-Guiard et al. (2017)	0.39	0.33	Observed increases in initiations, acts and responses while using the MR system, between sessions 1 and 3 (+)
Tsai et al. (2020)	0.5	0.6	Observed improvements in understanding of social initiations between baseline and intervention for all participants. Sustained at maintenance for 2 of the 3 participants (+/-) Statistically significant increase in therapist ratings of role play performance at intervention and maintenance (+)
Chen et al. (2015)	0.5	0.43	Statistically significant mean improvement in emotion recognition at follow-up in comparison to baseline (+)
Chen et al. (2016)	0.5	0.43	Statistically significant mean improvement in emotion recognition at follow-up in comparison to baseline (+)

a waitlist control, the former showed an increase in emotion regulation ratings after use of VR, as measured by the PEP-3 (Schopler et al., 2005), and the latter showed improvements in flexibility to change across all participants in the intervention group, according to idiosyncratic measures.

Quality of Included Studies

Quality ratings ranged from 0.31 to 0.67 (where possible scores are between 0 and 1) with an average quality rating of 0.50. Risk of bias scores ranged from 0.17 to 0.83 with

an average of 0.47. Although the adapted method of scoring makes it difficult to give a precise assessment of quality, comparisons with previous applications of this checklist suggests that the studies were generally of poor to fair quality (Hooper et al., 2008). Higher scores were typically given for clarity of reporting, such as describing the main aims, interventions and findings. The majority of studies took place in representative locations, and there was little evidence of non-compliance. Lower quality studies tended to have non-representative samples, unclear recruitment strategies and lack of reporting on attrition. Furthermore, several studies did not attempt to blind those involved in rating outcomes. It was particularly notable that no studies gave justification for their selected sample size.

Further quality appraisal using the SCED Scale indicated that the majority of small-N studies were truly experimental in design (7 of 10) and conducted sufficient sampling at baseline and treatment phases; some of these studies also performed statistical analysis. However, a lack of precise and repeatable outcome measures was a significant weakness across small-N studies. In particular, outcomes tended to be rated by one, non-independent assessor and where there were multiple assessors, interrater reliability was not calculated. No attempt was made to show generalisation across settings or therapists.

Discussion

This systematic review identified 17 studies that investigated the potential of XR interventions to enhance the social skills of autistic children. To date, the majority of interventions have targeted social-emotional reciprocity, with relatively little attention given to the non-verbal and relationship aspects of social-communication. The overall quality of the research is relatively low, perhaps reflecting the infancy of the research area. The majority of studies have small sample sizes and a number are not truly experimental, making it difficult to draw firm conclusions about the efficacy of XR, particularly in comparison to current, less costly social skills interventions. Significant heterogeneity means it is not possible to determine the ‘active ingredient’ of XR interventions. Furthermore, only limited attempts have been made as yet to determine generalisability and there is limited exploration of whether the statistically significant changes observed by many of the studies were also clinically significant i.e. led to improvements in everyday social or occupational skills, or other important areas of functioning.

Limitations of the Current Literature

Given the relative novelty of this research, it is important to consider the limitations of the current evidence base, to

be improved upon in future research. In particular, current evidence is limited in its ethical considerations, theoretical grounding, robustness of study design and sample representativeness. It is also unclear as yet, how easily and cost-effectively XR interventions could be scaled to clinical practice.

Ethical Considerations

Social skills interventions are primarily important because of the negative secondary effects of social difficulties and so quality of life has been an important outcome measure across the literature (e.g. Baghdadli et al., 2013; Mitchel et al., 2010). However, none of the studies included in this review measured quality of life, nor did they consider the importance of the intervention target to the participants themselves. Krasny et al. (2003) emphasise that interventions are most efficacious when children understand the relevance of the social skill and work towards individualised goals. Future interventions could capitalise on the customisability of XR to tailor interventions to individual needs. It is perhaps telling that the majority of interventions have targeted social reciprocity which, by definition, impacts those around autistic people. There are few attempts to directly target development and maintenance of relationships, despite this being what some autistic people have expressed as most important (Cresswell et al., 2019). Some studies did use participatory research methods, i.e. the incorporation of the views of autistic people into the design and implementation of the research (Cornwall & Jewkes, 1995), although the extent of involvement is somewhat limited. Increased meaningful involvement from autistic people within this field of research would help to ensure outputs are relevant and beneficial to the people for whom they are designed (Fletcher-Watson et al., 2019).

Future research should also make a more comprehensive attempt to measure adverse events, particularly given that one study which did comment on negative outcomes, noted problems with restlessness and discomfort, while another reported using positive reinforcement to keep children engaged. Moreover, commercially available VR headsets typically include safety notices regarding their use in children, emphasising the need for robust assessment of potential harm. Future studies could improve this through quantitative measurement of participant experience, such as whether the tool was straightforward to use and physically and sensorially comfortable, and analysis of the relationship to outcomes i.e. whether more positive experiences are associated with greater improvements. As well as the potential adverse effects of XR, future studies should consider the wider impact of teaching children to implement neurotypical social skills. Many autistic people already invest significant time and energy in monitoring and modifying their behaviour in order to align with social norms, known as

‘social camouflaging’. Evidence suggests that this hiding of one’s true self can be mentally and physically draining, and requires excessive concentration, self-control and discomfort (Hull et al., 2017; Mandy, 2019). Further exploration is needed of whether social skills training simply teaches children to mask their true selves, and if so, serious attention should be given to the potential deleterious impact on wellbeing. Perhaps greater research focus could instead be given to how settings, such as schools, could adapt to become more accommodating for autistic people, rather than placing all emphasis on changing autistic children to manage in a neurotypical world. This would be more consistent with the shared responsibility for communication between autistic and non-autistic people highlighted by the double empathy problem (Milton, 2012).

Theoretical Grounding

Despite the number of theoretical explanations for social differences in autism, the studies reviewed in this research give relatively little theoretical justification to the design of their interventions. As a result, XR has been applied in a number of different ways, often with no clear rationale for why design decisions were made. Greater attention to theory when designing interventions might further improve efficacy. This has proved beneficial for non-XR studies, such as the *Mind Reading* intervention which draws on the empathising-systemising theory by exploiting participants’ relative strength in systematising when teaching emotion recognition (Golan & Baron-Cohen, 2006). Limited consideration of the mechanism by which change might occur may also account for lack of specificity of some of the research. In particular, some studies measured a number of outcomes with no clear indication of how these related to the purpose or design of the intervention. Theoretical understandings of the mechanism of change could also inform intervention duration, which varied significantly across the studies. While it seems theoretically unlikely that clinically significant changes could be made in the time of the shortest intervention (15 min), strong theoretical justification should be given before subjecting children to high numbers of intervention sessions.

Study Design

Quality appraisal highlighted a number of methodological limitations of the current research. For example, the majority of studies did not attempt to blind the individuals who coded behavioural outcomes and in fact, behaviours were often rated by researchers or parents who knew the purpose of the study, both of whom are likely to have been highly invested in positive outcomes. In particular, parents

of autistic children have been shown to report positive effects when told their child is receiving treatment, even in the absence of an intervention (Jones et al., 2017). It is also possible that improvements occurred as a result of repetition of the outcome measures themselves. For example, a number of studies used questions based on Social Stories™ to measure outcomes, and as a social skills intervention in itself, this may have contributed to improvements. Similarly, when outcomes were measured by in vivo performance, improvements may have been due to practice effects, rather than social skill development. Greater use of validated measures of social skills, applied by raters who are blinded to the study purpose and intervention allocation, would improve some of these issues. Future research should also ensure that potential confounders, such as whether participants are receiving any other therapies, are fully considered.

Appropriately for the stage of research, the majority of studies were small-N. However, in larger trials, no details were given about the determination of sample size, and so it may be that those which did not find an effect were underpowered. In one study, a significant proportion of the sample were excluded from analysis based on age, however, the purpose of this is not explained, diminishing the validity of their results. To determine effectiveness, more large-scale trials are required: these should be sufficiently powered and ideally pre-registered, to reduce the likelihood of data dredging and subsequent false positives. More experimental, controlled studies would also make it easier to determine, not only whether XR interventions are effective, but whether they offer any real improvement to social skills, over and above current methods.

Representativeness of Sample

Current research is limited in the extent to which participants are representative of the autistic population as a whole. Around 25% of autistic people are female (Loomes et al., 2017) but only five of the included studies met this threshold. Similarly, up to 50% of the autistic population are estimated to have an Intellectual Disability (ID; Charman et al., 2011; Loomes et al., 2017) yet a number of the included studies excluded participants with an ID or stated a mean IQ within the average range. Limited justification is given for this and it compounds evidence that people with ID are consistently excluded from autism research (G. Russell et al., 2019). Future research should seek to recruit participants who truly reflect the autistic population as a whole, or provide strong justification for exclusion.

Limitations of the Review

The findings of this review should be viewed in the context of its limitations. Firstly, an effect size could only be given

for a limited number of studies making it impossible to compare the efficacy of different interventions using meta-analysis. Instead, this study used a systematic but non-analytic approach to synthesis. Vote counting and summary plots were avoided to reduce the potential misrepresentation of efficacy, given the number of low quality studies. This could perhaps have been improved by only including studies which meet a particular quality threshold. However, this was not deemed feasible for this review given the small number of studies. As the research develops and more experimental, controlled studies are carried out, more systematic methods of synthesis will be possible. Furthermore, publication bias will have significantly impacted the number of studies showing positive effects. Grey literature was not included, to ensure included studies were of peer-review quality, but this may have increased the positive results bias. When determining studies to be included, there were some borderline cases, where insufficient descriptions of study interventions made it difficult to establish whether the XR was immersive. This was resolved through discussion amongst researchers but, could have been improved by contacting study authors for more information.

The conclusions that can be drawn from this review are also constrained by the limited amount of data, both in terms of number of studies and overall sample size. The maximum age of participants across studies was 16 and so it is unclear whether XR interventions could be effective for helping older teenagers and young adults to enhance their social skills. Additionally, only VR studies which altered one's perception of reality were included and so the efficacy of immersive VR could not be compared to that of VLEs, which may be more cheaply and easily implemented in clinical practice.

Conclusion

The potential to digitally alter how one perceives the world offers exciting new possibilities in developing effective and engaging social skills interventions for autistic children. However, enthusiasm for XR interventions should be viewed with caution. While there is some indication that interventions are feasible to implement with autistic children, limited evidence exists for their effectiveness in bringing about meaningful, longstanding improvements in everyday functioning. Furthermore, no assessment has been made of the potential emotional cost of autistic children implementing behaviours which do not come naturally to them. This review demonstrates the need for theoretically grounded interventions, designed with the interests of autistic people at the forefront. Controlled trials and larger sample sizes, as well as other improvements to study design, are required to draw

firm conclusions about the efficacy of XR interventions and their generalisability, before potentially high cost scaling to routine clinical services is considered.

Appendix 1

Search terms*

-
1. 'Extended Reality' OR 'Virtual Reality' OR 'Virtual Environment' OR 'Augmented Reality' OR 'Augmented Environment' OR 'Simulat* Reality' OR 'Mixed Reality' OR 'VR' OR 'AR' OR 'XR' OR 'MR'
 2. 'Child*' OR 'Young' OR 'Youth*' OR 'Adolescen*' OR 'Student*' OR 'Teen*' OR 'Boy*' OR 'Girl*' OR 'P?ediatric*' OR 'Minor*'
 3. "Social Skill*" OR "Social Interacti*" OR "Social Competence" OR "Social Responsiveness" OR 'Social Motivation' OR 'Social Abilit*' OR 'Social Awareness' OR 'Conversation Skill*' OR 'Conversation Abilit*' OR 'Friendship Skill*' OR 'Peer interacti*' OR 'Emotion* Recognition' OR 'Facial Recognition' OR 'Perspective Taking' OR 'Turn Taking' OR 'Turn-Taking' OR 'Theory of Mind' OR 'Communicat*' OR 'Interact*'
 4. 'Autis*' OR 'Asperger*' OR 'ASD' OR 'ASC' OR 'Pervasive Developmental Disorder' OR 'PDD' OR 'HFA'
 5. 1 AND 2 AND 3 AND 4
-

*Keyword and thesaurus searches were also conducted for each of the above terms, where possible within each database.

Terms considered but excluded

'TOM', 'Toddler*', 'Baby', 'Babies', 'Newborn*', 'Preschool*', 'Preschool*'

Appendix 2

The Downs and Black Checklist (below) was used to assess study quality. The scoring of item 27, referring to the power of the study, was modified. Instead of rating according to an available range of study powers, we rated whether or not a power calculation was performed. Items which were not applicable to single case and small-N studies are marked*. Items which were not applicable to studies which did not use statistical analysis are marked †. Questions 14 and 22 were not applicable to studies which did not have a control intervention. All other items were rated as *Yes* (scored as 1) or *No* (scored as 0) or *Unable to Determine* (scored as 0). Items 14 to 26 specifically assess bias and so were used to generate a risk of bias score.

1. *Is the hypothesis/aim/objective of the study clearly described?*
 2. *Are the main outcomes to be measured clearly described in the Introduction or Methods section? If the main outcomes are first mentioned in the Results section, the question should be answered no.*
 3. *Are the characteristics of the participants included in the study clearly described? In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given.*
 4. *Are the interventions of interest clearly described? Treatments and placebo (where relevant) that are to be compared should be clearly described.*
 5. *Are the distributions of principal confounders in each group of subjects to be compared clearly described? A list of principal confounders is provided.*
 6. *Are the main findings of the study clearly described? Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions.*
 7. *Does the study provide estimates of the random variability in the data for the main outcomes? In non-normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered Yes.*
 8. *Have all important adverse events that may be a consequence of the intervention been reported? This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).*
-
9. *Have the characteristics of patients lost to follow up been described? This should be answered yes where there were no losses to follow up or where losses were so small that findings would be unaffected by their inclusion. This should be answered no, where the study does not report the number of patients lost to follow up.**
 10. *Have actual probability values been reported (e.g. 0.035 rather than <.05) for the main outcomes except where the probability value is less than 0.001? †*
 11. *Were the subjects asked to participate in the study representative of the entire population from which they were recruited? The study must identify the source population for participants and describe how the participants were selected. Participants would be representative if they comprised the entire source population, an unselected sample of consecutive participants, or a random sample. Random sampling is only feasible where a list of all members of the relevant population exists. Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.*
 12. *Were those subjects who were prepared to participate representative of the entire population from which they were recruited? The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.*

13. *Were the staff, places and facilities where the patients were treated, representative of the treatment the majority of patients receive? For the question to be answered yes the study should demonstrate that the intervention was representative of that in use in the source population. The questions should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.*

14. *Was an attempt made to blind study subjects to the intervention they have received? For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.**

15. *Was an attempt made to blind those measuring the outcomes of the intervention?*

16. *If any of the results of the study were based on 'data dredging', was this made clear? Any analysis that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.*

17. *In trials and cohort studies, do the analyses adjust for different lengths of follow up of patients, or in the case-control studies, is the time period between the intervention and outcome the same for cases and controls? Where follow-up was the same for all study patients, the answer should be yes. If different lengths of follow-up were adjusted for by, example, survival analysis the answer should be yes, studies where differences in follow-up are ignored should be answered no.***

18. *Were the statistical tests used to assess the main outcomes appropriate? The statistical techniques must be appropriate to the data. For example non-parametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the questions should be answered yes.*

19. *Was compliance with the intervention/s reliable? Where there was non-compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.*

20. *Were the main outcome measures used accurate (valid and reliable)? For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrated the outcomes measures are accurate, the question should be answered as yes.*

21. *Were the participants in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population? For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and case control studies where there is no information concerning the source of patients included in the study.* †*

22. *Were study participants in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time? For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.**

23. *Were study subjects randomised to intervention groups?* Studies which state patients were randomised should be answered yes except where method of randomisation would not ensure random allocation. For example, alternate allocation would score no because it is predictable.* †

24. *Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?* All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.* †

25. *Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?* This question should be answered no for trials if: the main conclusion of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In non-randomised studies, if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.* †

26. *Were losses of patients to follow up taken into account?* If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow up was too small to affect the main findings, the question should be answered yes.

27. *Is there any evidence that a power calculation or reasonable equivalent was used to determine sample size?*

Appendix 3

The SCED scale (below) was used to qualitatively assess the quality of the single case and small-N studies included in this review. All items are scored as *Yes* or *No*.

1. Clinical history was specified. Must include age, sex, aetiology and severity?
2. Target behaviours. Precise and repeatable measures that are operationally defined. Specify measure of target behaviour.
3. Design 1: 3 phases. Study must be either A-B-A or multiple baseline.
4. Design 2: Baseline (pre-treatment phase). Sufficient sampling was conducted.
5. Design 3: Treatment phase. Sufficient sampling was conducted.
6. Design 4: Data record. Raw data points were reported.
7. Observer bias: Inter-rater reliability was established for at least one measure of target behaviour.
8. Independence of assessors.
9. Statistical analysis.
10. Replication: either across subjects, therapists or settings.
11. Evidence for generalisation.

Acknowledgements Many thanks to Dr Sarah Coles and the Oxford Institute of Clinical Psychology Training and Research for their contribution to this work.

Author Contribution Rebecca Roberts, James Stacey, and Susannah Jenner contributed to the study conception and design. Literature searching, data extraction, and analysis were performed by Rebecca Roberts and Eloise Maguire. The first draft of the manuscript was written by Rebecca Roberts, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985). The aberrant behavior checklist: A behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency, 89*(5), 485–491.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Baghdadli, A., Brisot, J., Henry, V., Michelon, C., Soussana, M., Rattaz, C., & Picot, M. C. (2013). Social skills improvement in children with high-functioning autism: A pilot randomized controlled trial. *European Child & Adolescent Psychiatry, 22*(7), 433–442. <https://doi.org/10.1007/s00787-013-0388-8>
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition, 21*(1), 37–46. [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The ‘Reading the Mind in the Eyes’ Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 42*(2), 241–251. <https://doi.org/10.1017/S0021963001006643>
- Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a ‘language of the eyes’? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4*(3), 311–331. <https://doi.org/10.1080/713756761>
- Bauminger, N., & Kasari, C. (2000). Loneliness and friendship in high-functioning children with autism. *Child Development, 71*(2), 447–456. <https://doi.org/10.1111/1467-8624.00156>
- Begeer, S., Gevers, C., Clifford, P., Verhoeve, M., Kat, K., Hoddenbach, E., & Boer, F. (2011). Theory of mind training in children with autism: A randomized controlled trial. *Journal of Autism and Developmental Disorders, 41*(8), 997–1006. <https://doi.org/10.1007/s10803-010-1121-9>
- Bellini, S. (2006). The development of social anxiety in adolescents with autism spectrum disorders. *Focus on Autism and Other*

- Developmental Disabilities*, 21(3), 138–145. <https://doi.org/10.1177/10883576060210030201>
- Bellini, S., Peters, J. K., Benner, L., & Hopf, A. (2007). A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education*, 28(3), 153–162. <https://doi.org/10.1177/07419325070280030401>
- Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., Loranger, C., & Roucaut, F. X. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. *The British Journal of Psychiatry: The Journal of Mental Science*, 210(4), 276–283. <https://doi.org/10.1192/bjp.bp.116.184234>
- Brownlow, C. (2010). Re-presenting autism: The construction of ‘NT syndrome’. *The Journal of Medical Humanities*, 31(3), 243–255. <https://doi.org/10.1007/s10912-010-9114-4>
- Chandler, S., Charman, T., Baird, G., Simonoff, E., Loucas, T., Meldrum, D., Scott, M., & Pickles, A. (2007). Validation of the social communication questionnaire in a population cohort of children with autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46(10), 1324–1332. <https://doi.org/10.1097/chi.0b013e31812f7d8d>
- Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: Data from the Special Needs and Autism Project (SNAP). *Psychological Medicine*, 41(3), 619–627. <https://doi.org/10.1017/S0033291710000991>
- Chen, C.-H., Lee, I.-J., & Lin, L.-Y. (2015). Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. *Research in Developmental Disabilities*, 36, 396–403. <https://doi.org/10.1016/j.ridd.2014.10.015>
- Chen, C.-H., Lee, I.-J., & Lin, L.-Y. (2016). Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. *Computers in Human Behavior*, 55(Part A), 477–485. <https://doi.org/10.1016/j.chb.2015.09.033>
- Cheng, Y., Huang, C.-L., & Yang, C.-S. (2015). Using a 3D immersive virtual environment system to enhance social understanding and social skills for children with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 30(4), 222–236. <https://doi.org/10.1177/1088357615583473>
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16(4), 231–239. <https://doi.org/10.1016/j.tics.2012.02.007>
- Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale—second edition (SRS-2)*. Western Psychological Services.
- Cornwall, A., & Jewkes, R. (1995). What is participatory research? *Social Science & Medicine*, 41(12), 1667–1676. [https://doi.org/10.1016/0277-9536\(95\)00127-S](https://doi.org/10.1016/0277-9536(95)00127-S)
- Crowell, C., Sayis, B., Benitez, J. P., & Pares, N. (2020). Mixed reality, full-body interactive experience to encourage social initiation for autism: Comparison with a control nondigital intervention. *Cyberpsychology, Behavior and Social Networking*, 23(1), 5–9.
- Didehbani, N., Allen, T., Kandalraft, M., Krawczyk, D., & Chapman, S. (2016). Virtual Reality Social Cognition Training for children with high functioning autism. *Computers in Human Behavior*, 62, 703–711. <https://doi.org/10.1016/j.chb.2016.04.033>
- Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, 52(6), 377–384. <https://doi.org/10.1136/jech.52.6.377>
- Edey, R., Cook, J., Brewer, R., Johnson, M. H., Bird, G., & Press, C. (2016). Interaction takes two: Typical adults exhibit mind-blindness towards those with autism spectrum disorder. *Journal of Abnormal Psychology*, 125(7), 879–885. <https://doi.org/10.1037/abn0000199>
- Fletcher-Watson, S., Adams, J., Brook, K., Charman, T., Crane, L., Cusack, J., Leekam, S., Milton, D., Parr, J. R., & Pellicano, E. (2019). Making the future together: Shaping autism research through meaningful participation. *Autism*, 23(4), 943–953. <https://doi.org/10.1177/1362361318786721>
- Frith, U. (1989). *Autism: Explaining the Enigma*. Blackwell.
- Garrity, C., Gartlehner, G., Nussbaumer-Streit, B., King, V. J., Hamel, C., Kamel, C., Affengruber, L., & Stevens, A. (2021). Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. *Journal of Clinical Epidemiology*, 130, 13–22. <https://doi.org/10.1016/j.jclinepi.2020.10.007>
- Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology*, 18(2), 591–617. <https://doi.org/10.1017/S0954579406060305>
- Gresham, F. M., & Elliott, S. N. (1984). Assessment and classification of children’s social skills: A review of methods and issues. *School Psychology Review*, 13(3), 292–301.
- Happé, F., & Frith, U. (1996). The neuropsychology of autism. *Brain*, 119(4), 1377–1400. <https://doi.org/10.1093/brain/119.4.1377>
- Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5–25. <https://doi.org/10.1007/s10803-005-0039-0>
- Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, 9(10), 1218–1220. <https://doi.org/10.1038/nn1770>
- Harrison, P. L., & Oakland, T. (2003). *Adaptive Behavior Assessment System Second Edition* ((2nd ed.). ed.). The Psychological Corporation.
- Herrero, J. F., & Lorenzo, G. (2020). An immersive virtual reality educational intervention on people with autism spectrum disorders (ASD) for the development of communication skills and problem solving. *Education and Information Technologies*, 25(3), 1689–1722. <https://doi.org/10.1007/s10639-019-10050-0>
- Hillier, A. J., Fish, T., Siegel, J. H., & Beversdorf, D. Q. (2011). Social and vocational skills training reduces self-reported anxiety and depression among young adults on the autism spectrum. *Journal of Developmental and Physical Disabilities*, 23(3), 267–276. <https://doi.org/10.1007/s10882-011-9226-4>
- Hooper, P., Jutai, J. W., Strong, G., & Russell-Minda, E. (2008). Age-related macular degeneration and low-vision rehabilitation: A systematic review. *Canadian Journal of Ophthalmology*, 43(2), 180–187. <https://doi.org/10.3129/i08-001>
- Howlin, P. (2003). Longer-term educational and employment outcomes. In M. Prior (Ed.), *Learning and behavior problems in Asperger syndrome* (pp. 269–293). Guilford Press.
- Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., & Mandy, W. (2017). “Putting on My Best Normal”: Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, 47(8), 2519–2534. <https://doi.org/10.1007/s10803-017-3166-5>
- Ip, H. H. S., Wong, S. W. L., Chan, D. F. Y., Byrne, J., Li, C., Yuan, V. S. N., Lau, K. S. Y., & Wong, J. Y. W. (2018). Enhance emotional and social adaptation skills for children with autism spectrum disorder: A virtual reality enabled approach. *Computers & Education*, 117, 1–15. <https://doi.org/10.1016/j.compedu.2017.09.010>
- Johnson, S. B., Blum, R. W., & Giedd, J. N. (2009). Adolescent maturity and the brain: The promise and pitfalls of neuroscience research in adolescent health policy. *The Journal of Adolescent Health*, 45(3), 216–221. <https://doi.org/10.1016/j.jadohealth.2009.05.016>

- Jones, R. M., Carberry, C., Hamo, A., & Lord, C. (2017). Placebo-like response in absence of treatment in children with Autism. *Autism Research*, 10(9), 1567–1572. <https://doi.org/10.1002/aur.1798>
- Karkhaneh, M., Clark, B., Ospina, M. B., Seida, J. C., Smith, V., & Hartling, L. (2010). Social Stories™ to improve social skills in children with autism spectrum disorder: A systematic review. *Autism: The International Journal of Research and Practice*, 14(6), 641–662. <https://doi.org/10.1177/1362361310373057>
- Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), 442–462. <https://doi.org/10.1177/1362361315588200>
- Keshav, N. U., Salisbury, J. P., Vahabzadeh, A., & Sahin, N. T. (2017). Social communication coaching smartglasses: well tolerated in a diverse sample of children and adults with autism. *JMIR MHealth and UHealth*, 5(9), 140. <https://doi.org/10.2196/mhealth.8534>
- Koegel, L. K., Ashbaugh, K., & Koegel, R. L. (2016). Pivotal response treatment. In R. Lang, T. B. Hancock, & N. N. Singh (Eds.), *Early intervention for young children with autism spectrum disorder* (pp. 85–112). Springer International Publishing. https://doi.org/10.1007/978-3-319-30925-5_4
- Krasny, L., Williams, B. J., Provencal, S., & Ozonoff, S. (2003). Social skills interventions for the autism spectrum: Essential ingredients and a model curriculum. *Child and Adolescent Psychiatric Clinics of North America*, 12(1), 107–122. [https://doi.org/10.1016/s1056-4993\(02\)00051-2](https://doi.org/10.1016/s1056-4993(02)00051-2)
- Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders*, 30(3), 183–193. <https://doi.org/10.1023/a:1005558101038>
- Laver, K. E., Lange, B., George, S., Deutsch, J. E., Saposnik, G., & Crotty, M. (2017). Virtual reality for stroke rehabilitation. *The Cochrane Database of Systematic Reviews*, 11, CD008349. <https://doi.org/10.1002/14651858.CD008349.pub4>
- Lee, I.-J., Chen, C.-H., Wang, C.-P., & Chung, C.-H. (2018). Augmented reality plus concept map technique to teach children with ASD to use social cues when meeting and greeting. *The Asia-Pacific Education Researcher*, 27(3), 227–243. <https://doi.org/10.1007/s40299-018-0382-5>
- Liu, R., Salisbury, J. P., Vahabzadeh, A., & Sahin, N. T. (2017). Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching. *Frontiers in Pediatrics*, 5, 145. <https://doi.org/10.3389/fped.2017.00145>
- Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(6), 466–474. <https://doi.org/10.1016/j.jaac.2017.03.013>
- Lorenzo, G., Gómez-Puerta, M., Arráez-Vera, G., & Lorenzo-Lledó, A. (2019). Preliminary study of augmented reality as an instrument for improvement of social skills in children with autism spectrum disorder. *Education and Information Technologies*, 24(1), 181–204. <https://doi.org/10.1007/s10639-018-9768-5>
- Lorenzo, G., Lledó, A., Pomares, J., & Roig, R. (2016). Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders. *Computers & Education*, 98, 192–205. <https://doi.org/10.1016/j.compedu.2016.03.018>
- Lorenzo, G., Pomares, J., & Lledó, A. (2013). Inclusion of immersive virtual learning environments and visual control systems to support the learning of students with Asperger syndrome. *Computers & Education*, 62, 88–101. <https://doi.org/10.1016/j.compedu.2012.10.028>
- Mandelberg, J., Laugeson, E. A., Cunningham, T. D., Ellingsen, R., Bates, S., & Frankel, F. (2014). Long-term treatment outcomes for parent-assisted social skills training for adolescents with autism spectrum disorders: The UCLA PEERS Program. *Journal of Mental Health Research in Intellectual Disabilities*, 7(1), 45–73. <https://doi.org/10.1080/19315864.2012.730600>
- Mandy, W. (2019). Social camouflaging in autism: Is it time to lose the mask? *Autism*, 23(8), 1879–1881. <https://doi.org/10.1177/1362361319878559>
- Matson, J. L., Matson, M. L., & Rivet, T. T. (2007). Social-skills treatments for children with autism spectrum disorders: An overview. *Behavior Modification*, 31(5), 682–707. <https://doi.org/10.1177/0145445507301650>
- McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders*, 32(5), 351–372. <https://doi.org/10.1023/A:1020537805154>
- Milton, D. E. M. (2012). On the ontological status of autism: The ‘double empathy problem’. *Disability & Society*, 27(6), 883–887. <https://doi.org/10.1080/09687599.2012.710008>
- Mitchel, K., Regehr, K., Reaume, J., & Feldman, M. (2010). Group social skills training for adolescents with Asperger Syndrome or high functioning autism. *Journal on Developmental Disabilities*, 16(2), 52–63.
- Mora-Guiard, J., Crowell, C., Pares, N., & Heaton, P. (2017). Sparking social initiation behaviors in children with autism through full-body Interaction. *International Journal of Child-Computer Interaction*, 11, 62–71. <https://doi.org/10.1016/j.ijcci.2016.10.006>
- Newbutt, N., Sung, C., Kuo, H.-J., Leahy, M. J., Lin, C.-C., & Tong, B. (2016). Brief Report: A pilot study of the use of a virtual reality headset in autism populations. *Journal of Autism and Developmental Disorders*, 46(9), 3166–3176. <https://doi.org/10.1007/s10803-016-2830-5>
- Ochs, E., Kremer-Sadlik, T., Solomon, O., & Sirota, K. G. (2001). Inclusion as social practice: Views of children with autism. *Social Development*, 10(3), 399–419. <https://doi.org/10.1111/1467-9507.00172>
- Pellicano, E., & Burr, D. (2012). When the world becomes ‘too real’: A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510. <https://doi.org/10.1016/j.tics.2012.08.009>
- Ravindran, V., Osgood, M., Sazawal, V., Solorzano, R., & Turnacioglu, S. (2019). Virtual reality support for joint attention using the Floreo joint attention module: Usability and feasibility pilot study. *JMIR Pediatrics and Parenting*, 2(2), e14429. <https://doi.org/10.2196/14429>
- Reichow, B., Steiner, A. M., & Volkmar, F. (2012). Social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Campbell Systematic Reviews*, 8(1), 1–76. <https://doi.org/10.4073/csr.2012.16>
- Rieth, S. R., Stahmer, A. C., Suhrheinrich, J., Schreibman, L., Kennedy, J., & Ross, B. (2014). Identifying critical elements of treatment: Examining the use of turn taking in autism intervention. *Focus on Autism and Other Developmental Disabilities*, 29(3), 168–179. <https://doi.org/10.1177/1088357613513792>
- Rivière, A. (2002). *IDEA: Inventario de Espectro Autista*. Fundec.
- Russell, G., Mandy, W., Elliott, D., White, R., Pittwood, T., & Ford, T. (2019). Selection bias on intellectual ability in autism research: A cross-sectional review and meta-analysis. *Molecular Autism*, 10(1), 9. <https://doi.org/10.1186/s13229-019-0260-x>
- Russell, J. (1997). How executive disorders can bring about an inadequate ‘theory of mind’. In *Autism as an executive disorder* (pp. 256–304). Oxford University Press.
- Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The ‘windows task’ as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology*,

- 9(2), 331–349. <https://doi.org/10.1111/j.2044-835X.1991.tb00881.x>
- Schohl, K. A., Van Hecke, A. V., Carson, A. M., Dolan, B., Karst, J., & Stevens, S. (2014). A replication and extension of the PEERS intervention: Examining effects on social skills and social anxiety in adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 44(3), 532–545. <https://doi.org/10.1007/s10803-013-1900-1>
- Schopler, E., Lansin, M. D., Reichler, R. J., & Marcus, L. M. (2005). *Psychoeducational Profile Third Edition (PEP-3)* ((3rd ed.). ed.). Pro-ed Incorporation. <https://doi.org/10.1037/152601-000>
- Sheppard, E., Pillai, D., Wong, G. T.-L., Ropar, D., & Mitchell, P. (2016). How easy is it to read the minds of people with autism spectrum disorder? *Journal of Autism and Developmental Disorders*, 46(4), 1247–1254. <https://doi.org/10.1007/s10803-015-2662-8>
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70, 747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- Sng, C. Y., Carter, M., & Stephenson, J. (2014). A review of video modelling and scripts in teaching conversational skills to individuals with autism spectrum disorders. *Review Journal of Autism and Developmental Disorders*, 1(2), 110–123. <https://doi.org/10.1007/s40489-013-0010-5>
- Spence, S. H. (2003). Social skills training with children and young people: Theory, evidence and practice. *Child and Adolescent Mental Health*, 8(2), 84–96. <https://doi.org/10.1111/1475-3588.00051>
- Strickland, D., Marcus, L. M., Mesibov, G. B., & Hogan, K. (1996). Brief report: Two case studies using virtual reality as a learning tool for autistic children. *Journal of Autism and Developmental Disorders*, 26(6), 651–659. <https://doi.org/10.1007/BF02172354>
- Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) scale. *Neuropsychological Rehabilitation*, 18(4), 385–401. <https://doi.org/10.1080/09602010802009201>
- Thye, M. D., Bednarz, H. M., Herringshaw, A. J., Sartin, E. B., & Kana, R. K. (2018). The impact of atypical sensory processing on social impairments in autism spectrum disorder. *Developmental Cognitive Neuroscience*, 29, 151–167. <https://doi.org/10.1016/j.dcn.2017.04.010>
- Tsai, W.-T., Lee, I.-J., & Chen, C.-H. (2020). Inclusion of third-person perspective in CAVE-like immersive 3D virtual reality role-playing games for social reciprocity training of children with an autism spectrum disorder. *Universal Access in the Information Society*, 2, 375–389. <https://doi.org/10.1007/s10209-020-00724-9>
- Vincelli, F., Anolli, L., Bouchard, S., Wiederhold, B. K., Zurloni, V., & Riva, G. (2003). Experiential cognitive therapy in the treatment of panic disorders with agoraphobia: A controlled study. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, 6(3), 321–328. <https://doi.org/10.1089/109493103322011632>
- Vygotsky, L. S., & Luria, A. (1994). Tool and symbol in child development. In R. Van Der Veer & J. Valsiner (Eds.), *The Vygotsky Reader*. Blackwell Publishers.
- Welsh, M., Parke, R. D., Widaman, K., & O’Neil, R. (2001). Linkages between children’s social and academic competence: A longitudinal analysis. *Journal of School Psychology*, 39(6), 463–482. [https://doi.org/10.1016/S0022-4405\(01\)00084-X](https://doi.org/10.1016/S0022-4405(01)00084-X)
- Williams, J., Scott, F., Stott, C., Allison, C., Bolton, P., Baron-Cohen, S., & Brayne, C. (2005). The CAST (Childhood Asperger Syndrome Test): Test accuracy. *Autism*, 9(1), 45–68. <https://doi.org/10.1177/1362361305049029>
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., Text Revision).
- Berenguer, C., Baixauli, I., Gomez, S., Andres, M. de E. P., & De Stasio, S. (2020). Exploring the impact of augmented reality in children and adolescents with autism spectrum disorder: A systematic review. *International Journal of Environmental Research and Public Health*, 17(17). <https://doi.org/10.3390/ijerph17176143>
- Cresswell, L., Hinch, R., & Cage, E. (2019). The experiences of peer relationships amongst autistic adolescents: A systematic review of the qualitative evidence. *Research in Autism Spectrum Disorders*, 61. <https://doi.org/10.1016/j.rasd.2019.01.003>
- Jaswal, V. K., & Akhtar, N. (2019). Being versus appearing socially uninterested: Challenging assumptions about social motivation in autism. *Behavioral and Brain Sciences*, 42. <https://doi.org/10.1017/S0140525X18001826>
- Lee, I.-J. (2020). Kinect-for-windows with augmented reality in an interactive roleplay system for children with an autism spectrum disorder. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2019.1710851>
- Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J.-A., & Gil-Gómez, J.-A. (2018). Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: An evidence-based systematic review. *Sensors*, 18(8). <https://doi.org/10.3390/s18082486>
- NHS. (2019). *NHS Long Term Plan*. <https://www.longtermplan.nhs.uk/>
- Riva, G., Baños, R. M., Botella, C., Mantovani, F., & Gaggioli, A. (2016). Transforming experience: The potential of augmented reality and virtual reality for enhancing personal and clinical change. *Frontiers in Psychiatry*, 7. <https://doi.org/10.3389/fpsy.2016.00164>
- Robledo, J., Donnellan, A. M. P., & Strandt-Conroy, K. (2012). An exploration of sensory and movement differences from the perspective of individuals with autism. *Frontiers in Integrative Neuroscience*, 6. <https://doi.org/10.3389/fnint.2012.00107>
- Sahin, N. T., Abdus-Sabur, R., Keshav, N. U., Liu, R., Salisbury, J. P., & Vahabzadeh, A. (2018). Case study of a digital augmented reality intervention for autism in school classrooms: Associated with improved social communication, cognition, and motivation via educator and parent assessment. *Frontiers in Education*, 3. <https://doi.org/10.3389/educ.2018.00057>
- Simoes, M., Bernardes, M., Barros, F., & Castelo-Branco, M. (2018). Virtual travel training for autism spectrum disorder: Proof-of-concept interventional study. *JMIR SERIOUS GAMES*, 6(1). <https://doi.org/10.2196/games.8428>
- Stark, E., Stacey, J., Mandy, W., Kringelbach, M. L., & Happé, F. (2021). Autistic cognition: Charting routes to anxiety. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2021.03.014>
- Vahabzadeh, A., Keshav, N. U., Abdus-Sabur, R., Huey, K., Liu, R., & Sahin, N. T. (2018). Improved socio-emotional and behavioral functioning in students with autism following school-based smartglasses intervention: Multi-stage feasibility and controlled efficacy study. *Behavioral Sciences*, 8(10). <https://doi.org/10.3390/bs8100085>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.