

Case Study 1: An Evidence-Based Practice Review Report***Theme: School (setting) based interventions for children with special educational needs (SEN)******Effectiveness of computer-based game interventions at improving social skills for children with Autistic Spectrum Conditions (ASC): a systematic review*****Section 1: Summary**

Children with autism often face social communication difficulties that may continue into adulthood. Play allows children to engage with peers, providing a unique context for the development of social communication in real-life settings. There has been a growing trend of play-based computer interventions at improving social skills (SS) for children with autism, especially during COVID-19 with remote learning. This paper aims to systematically review the evidence-base of five studies that were appraised using Kratochwill's (2003) coding protocol and the Gough (2007)'s weight of evidence framework. This review selected a total of five studies, of which four studies included were randomised controlled trials, and one study was quasi-experimental. All included studies focused on child related SS outcomes through parent or teacher questionnaire and behavioural observations. The majority of interventions targeted improvements on joint attention, initiation, social awareness, and emotion regulation. Overall, studies showed improvement in SS after intervention for school-aged children with autism. More robust research is needed for good-quality evidence-base in the field of computer-based interventions due to small sample size, lack of follow up, and reporting bias.

Section 2: Introduction

2.1 Children with Autism Spectrum Condition

Autism Spectrum Condition (ASC) is a neurodevelopmental condition that emerges in early childhood. ASC is characterised by a wide spectrum of difficulties in social communication and social interaction and restricted repetitive behaviours or interests (American Psychiatric Association, 2013). In the United Kingdom, children with ASC remain the most common type of special educational need (SEN) with an Education, Health and Care Plan (EHCP). According to the United Kingdom Statistics Authority (2021), there are 92,567 pupils with autism in the 2020/21 academic year (Roman-Urrestarazu et al., 2021). According to the largest national UK study to date (around 7 million pupils), the prevalence of children and young people with a clinical diagnosis of autism is estimated to be 1.76% of the population (Roman-Urrestarazu et al., 2021).

This review favours the term 'autism spectrum condition' (ASC), instead of 'autism spectrum disorder' (ASD) because the keyword 'condition' is less stigmatising as a type of disability. Both ASC and ASD have the same diagnostic process of presenting difficulty in social skills and communication, alongside with repetitive behaviour (Baron-Cohen et al., 2009). However, ASC approaches autism not as a disability of medical diagnosis, but expresses value for the neurodiversity of one's cognitive strength (Baron-Cohen et al., 2009).

2.2 ASC Social Skills

All children and young people with ASC experience challenges in social communication skills (O'Keeffe & McNally, 2021). Children with ASC may have developmental needs in verbal and nonverbal skills and in responding to joint attention.

These needs may lead to higher rates of social exclusion or bullying with peers (Chawarska et al., 2016; Roberts et al., 2020). Therefore, early and accessible interventions are recommended to improve SS of children on the spectrum to offset these undesirable developmental trajectories (Beaumont et al., 2021).

Historically, most psychological interventions targeting SS are underpinned by behavioural and social learning theories. For instance, a behavioural-based intervention, 'Social Stories' (Gray, 1991) helps to explain the complex and subtle social rules for children to adapt their behaviour. Research from social stories interventions have shown reductions in bullying and peer rejections (Jones & Bawazir, 2018). On the other hand, social learning theory emphasises the role of observational learning and modelling of different social behaviours. SS interventions based on this approach emphasise the importance of breaking down complex skills into small, attainable, explicit, and step-by-step tasks (Beaumont et al., 2021). An emerging theory from the work of Vygotsky (1967) recognises children's development of learning is based on real-life social interactions or play within the community.

2.3 Psychological Underpinnings of Computer-Based Game Interventions

2.3.1 Social Learning through Play

According to Vygotsky's cultural-historical theory, play is the important type of cognitive, emotional, and social development for children (Bodrova & Leong, 2015). Fostering a play-learning environment encourages children to acquire learning in a naturalistic context. This may include play-based social learning such as developing friendships and SS (Gray, 2011). Recently, there has been an interest in types of play

utilising digital platforms (Marsh et al., 2016). Researchers suggest computer games can offer children the opportunity to engage in social friendships, such as multi-player modes that help joint engagement and attention between peers.

2.3.2 Theory of Mind and Computer Games

Children with ASC have a heterogeneous range of symptoms. Although robust empirical literature suggests children with ASC display needs in Theory of Mind (ToM), they all vary in degrees of severity (Rosello et al., 2020). Children are exposed with daily social interactions that are unpredictable and ambiguous within a limited space of time (M Schaller & Rauh, 2017). The difficulties of recognising facial expressions, verbal and non-verbal contextual cues make social interactions challenging and cause congestion in processing social stimuli.

There have been recent attempts to teach components of ToM to children with ASC with computer-based games or virtual reality. These games aim to facilitate and improve social interactions in individuals with ASC (Grossard et al., 2017). They exist in multiple platforms, such as computers, robots, virtual reality, and iPads. Computer-based games allow children with social and communication difficulties to form relationships with peers without being overly reliant on adult assistance or external support (Atherton & Cross, 2021). This is demonstrated by a study where a majority of children were able to apply knowledge of social rules in the game, and interact with other peers without adult oversight (Lancy & Grove, 2017). Research suggests computer games have the capacity to provide perform diverse contexts, of which the majority are simulations of real-life (Wainer & Ingersoll, 2011). Therefore, computer-based social skills intervention may offer good educational tools for children to improve

their social interaction with peers (Wainer & Ingersoll, 2011). Several studies also indicate computer games could monitor and evaluate children's social skills (Li & Elmaghraby, 2014).

2.3.3 Executive Dysfunction and Computer Games

Children with ASC often experience executive functioning difficulties (Rosenthal et al., 2013). Computer games help to support and offset these needs for children because they are designed with built-in reward systems (e.g. point levels, progress bars) that help track progress (Atherton & Cross, 2021). These clear visual and auditory cues clearly define the expectations with immediate and frequent reinforcements. Having straightforward animated visual rules helps children to understand the well-defined structure and framework of the target or goal in the game. It is suggested that individuals with ASC often have strong visual processing skills and a predilection towards electronic media, which is particularly motivating in improving social skills with computer games (Shane & Albert, 2008). Especially to those who have strong preferences for routine, predictability, and repetition. These positive reinforcements help motivate children, including those with ASC, help completing the full intervention programme (Abd El-Sattar, 2008; Filsecker & Hickey, 2014). This suggests computer games may be more appealing to children with ASC because they are more predictable and controllable activities, compared to real-world interactions. There might be a limitation on social skills computer games reflecting daily interactions. However, it provides a basic and fundamental understanding of social interactions supported by various daily examples.

2.4 Rationale and Relevance to Educational Psychology Practice

Inclusive education has been the key focus of government initiatives (DoE, 2001, p. 360). However, difficulties with social communication remains a challenge for children with ASC in mainstream school settings (Harrower & Dunlap, 2001). As outlined in the SEND Code of Practice (2015), schools and educational psychologists (EPs) have the duty to promote inclusive practice in schools. Working systemically around the child promotes positive and sustainable change that marries up to the different demands of the environment.

Interventions targeted at developing SS usually require in-person participation. The growing trend of technology-enhanced interventions allows flexibility for participating in any settings, even at home. Especially with the social distancing restrictions imposed by COVID-19, many researchers developed interventions to be used remotely (Beaumont et al., 2021). Children with ASC have been found to benefit from interventions, including computer-based social skills interventions, which have improved their social communication due to explicit and repeated practice. Grossard et al. (2017) conducted a systematic review on computer-based video games as a tool to teach children with ASC social interactions and emotion recognition. The systematic review included 15 studies on social interactions and 16 studies on emotion recognition. Results suggested the computer games provided training on social interactions with a range of contexts and situations. However, studies are done mainly on clinically diagnosed autism, and they exhibit a significant degree of weak methodological design (e.g. computer intervention not described, small sample size, missing control group). This review excluded studies that require virtual environment. It is proposed that this review provides evidence on computer-based social skills interventions that are easily accessible and implemented in schools and home settings.

This allows greater feasibility for teachers and parents to use computer-based interventions.

Technology has the potential to provide affordable, accessible, and portable interventions in different settings, such as in schools (e.g. e-learning) or at home (Fletcher-Watson et al., 2016). Such SS computer-based interventions would be helpful for EPs to signpost to families and school staff (e.g. In-service training (INSET)) through targeted groups of children or a whole-school approach.

2.5 Review question

'How effective are computer-based game interventions at improving social skills for children with autistic spectrum conditions (ASC)?'

The aim of this present paper is to systematically review the evidence on the effectiveness of computer-based games as interventions to improve SS for children diagnosed with ASC. Only studies employing quantitative randomised control design and block design trials were selected for this review. The scope of the review will discuss the multitude of issues present in the literature (e.g. sample sizes, bias, measurement tool reliability, etc). Therefore, this paper aims to offer an accessible review of some existing literature on the use of computer games for children with autism and SS, including recommendations for future game-developers and researchers.

Section 3: Critical Review of the Evidence Base

3.1.1 Keywords and search terms

Online peer-reviewed journal searches were undertaken on PsycINFO, Web of Science, ERIC and SCOPUS (a full database search was conducted with no time limit) with additional ancestral searching. The latest search on 30 December 2021 used the following keywords and search terms shown in Table 1 for all databases. An example of the full electronic search strategy used is included in Appendix A.

Table 1: Search terms for the review

Autism		Intervention		Social skills		Child
Autis* OR Autism spectrum disorder* OR ASD OR Autism spectrum condition OR Asperger syndrome OR High functioning autis*	And	Game OR video game OR serious game OR online OR online game OR interactive OR interactive game OR mobile OR mobile game OR computer OR computer-based OR computer- based interventions OR CBI OR digital OR digital gamification OR gamification OR Play	And	Social skills OR Friendship OR social relationship OR peer relationship OR friend*	And	Child* OR school age* OR primary school OR elementary school OR youth OR secondary school

Note: These search terms were formulated according to four categories: autism, intervention, social skills, and child. The terms within a category were combined with 'OR' and terms in different categories were combined with 'AND'.

3.1.2 Inclusion and exclusion criteria

Studies retrieved were included in the review if they met the criteria as detailed within Table 2. Figure 1 illustrates that a total of 223 articles were found through electronic database searches and other scoping searches. Of these, 94 were removed as duplicates with 129 studies remaining. Titles screening was then completed on those 129 studies, and a further 37 articles were excluded at that stage. Abstract screening

of the remaining 92 studies resulted in 42 studies being retained and 50 being excluded. A further 26 studies were excluded after initial full text screening as they did not meet the inclusion criteria outlined in Table 2. This left 16 articles for a thorough full text screen. After this, 4 articles remained. An addition 1 article was retrieved by hand search from reference list of relevant reviews. Therefore, 5 articles were included for the final review.

Table 2: Inclusion and Exclusion criteria

Study Feature	Inclusion Criteria	Exclusion Criteria	Rationale
1. Diagnosis	(a) Children met the criteria for Autism Spectrum Condition/ Disorder using the Diagnostic and Statistical Manual (DSM IV/V). Studies included children with autism on the spectrum.	(a) Children did not meet the criteria for Autism Spectrum Condition/ Disorder by a professional	(a) The review question focuses on the effectiveness of intervention for children with clinically diagnosed autism by a professional.
	(b) Nonverbal IQ of 80 or above from a standardised assessment tool (e.g. Wechsler Intelligence Scale for Children (WISC) and British Ability Scales)	(b) Studies did not report the cognitive abilities of children with a standardised assessment tool	(b) To confirm that cognitive functioning was within the normative range
	(c) Children with autism may have a co-existing diagnosis (e.g., ADHD)	(c) Children with a primary diagnosis that is not autism	(c) Nearly three-quarters of autistic children also have another psychiatric comorbidity, which this review aims to generalise the evidence-base into real-life scenarios.
2. Language	Study published in English	Published in another language other than English	Translation services are unavailable for this review due to limited resources
3. Country of Study	Study conducted within the Organisation for Economic Co-	Study conducted not in an OECD country	OECD countries have similar social infrastructure and policies that impact education and culture.

Study Feature	Inclusion Criteria	Exclusion Criteria	Rationale
	operation and Development (OECD) country		These societal factors may influence on parenting behaviours and school culture.
4. Participants	(a) School aged children and young people from Reception to Year 12 (ages 4-18)	(a) Children younger than age 4 and/ or older than age 18	(a) The review seeks to understand how computer-based interventions benefit school aged children
	(b) Children not undergoing any concurrent psychological treatment	(b) Children concurrently undergoing additional psychological treatment	(b) Concurrent treatments would confound the results observed
	(c) Children taking medication must be stable for at least 1 month before the start of the trial	(c) Children who take medication have new or changes in drug dosages during the study is not monitored.	(c) Unstable or unmonitored medication would confound the results observed
5. Play Computer-Based Intervention	The study includes technology or play computer-based intervention that adapts Whyte (2015)'s principles for an adequate investigate of serious games for autism:	The study includes an intervention that:	
	(a) Uses positive feedback rewards to reinforce learning. For example, feedback on	(a) does not use technology or computer-based intervention	(a) This review is only focusing on computer-based interventions. Other digital formats (e.g. virtual reality) are not included because this

Study Feature	Inclusion Criteria	Exclusion Criteria	Rationale
<p>response accuracy and cumulative point systems levels</p>	<p>(b) foster learning of targeted skills that are challenging for participants. The intervention includes an educational component of learning to improve real life outcomes</p>	<p>(b) is designed for the purpose of entertainment and does not have any key learning objectives.</p>	<p>review aims to implement interventions accessibly in school and home settings</p>
<p>(c) uses immersive storylines that is understood by school-aged children</p>	<p>(b) foster learning of targeted skills that are challenging for participants. The intervention includes an educational component of learning to improve real life outcomes</p>	<p>(c) does not have a narrative storyline in the game</p>	<p>(b) This review focuses on computer-based interventions that is a type of ‘Serious Game’. Serious game is not designed for entertainment purpose. Instead, serious game focuses on targeted educational objectives with specific evidence-based game mechanics to support learning and generalization of learning (Whyte et al., 2015).</p>
<p>(d) is motivating for children, includes rewards and feedback about goal progress. The intervention is gamified by increasing levels of difficulty</p>	<p>(c) uses immersive storylines that is understood by school-aged children</p>	<p>(d) does not include any rewards and feedback and does not give information about the goal progress. There is no level of difficulties in the serious game.</p>	<p>(c) Narrative storyline increases the enjoyment and motivation for children to learn by immersing into meaningful contexts. Children may develop emotional connections or relationships with the characters, providing a potential tool for enhancing social skills. Storylines enable children to experience and learn effectively with real life social situations</p>
	<p>(d) is motivating for children, includes rewards and feedback about goal progress. The intervention is gamified by increasing levels of difficulty</p>		<p>(d) This review aims to bring long-term outcomes, and achieving rewards to enhance intrinsic motivation is important for children to monitor their progress towards incremental and primary learning goals (Turan et al., 2016)</p>

Study Feature	Inclusion Criteria	Exclusion Criteria	Rationale
	(e) trains parents or teachers, so they could act as an intervention delivery agent with the children that includes no therapist-child contact	(e) includes contact between therapist or researcher and child during the intervention, even if parents or teachers are the primary implementers	(e) This review is only focusing on parent or teacher guided interventions with no child-therapist contact
6. Outcomes	One of the primary outcomes measures the impact of the intervention on the child’s social skills. Social skills outcome subtests, such as joint attention, initiation, social awareness, and emotion regulation.	None of the primary outcomes look at the impact of the intervention on the child’s social skills	This review seeks to evaluate the effectiveness of computer-based intervention on improving children with autistic conditions’ social skills
7. Type of Studies	(a) Empirical studies including randomised control trials (RCTs), quasi-experimental studies, and longitudinal studies. Studies include participants in either intervention or control group.	(a) The study employs a research study that has no intervention and control group. Studies may include case studies, descriptive studies, focused-group studies, and qualitative reports	(a) Participants in either intervention or control groups is to ensure the effect observed is from the intervention, not due to bias within the participants or other confounding factors. Empirical studies are included to critically review the evidence-base of quantitative studies
8. Type of Publications	Peer-reviewed journals published with no time limit	Non-peer reviewed journals, books, dissertations, conference paper,	Reviews on computer-based interventions would naturally be excluded in later years (e.g. 1900s) due to lack of advancement in technology. This

Study Feature	Inclusion Criteria	Exclusion Criteria	Rationale
		systematic review, and grey literature	review seeks to review any computer-based interventions in recent years (e.g. past 10 years) that would be effective and user-friendly for 21 st century or later computer users.

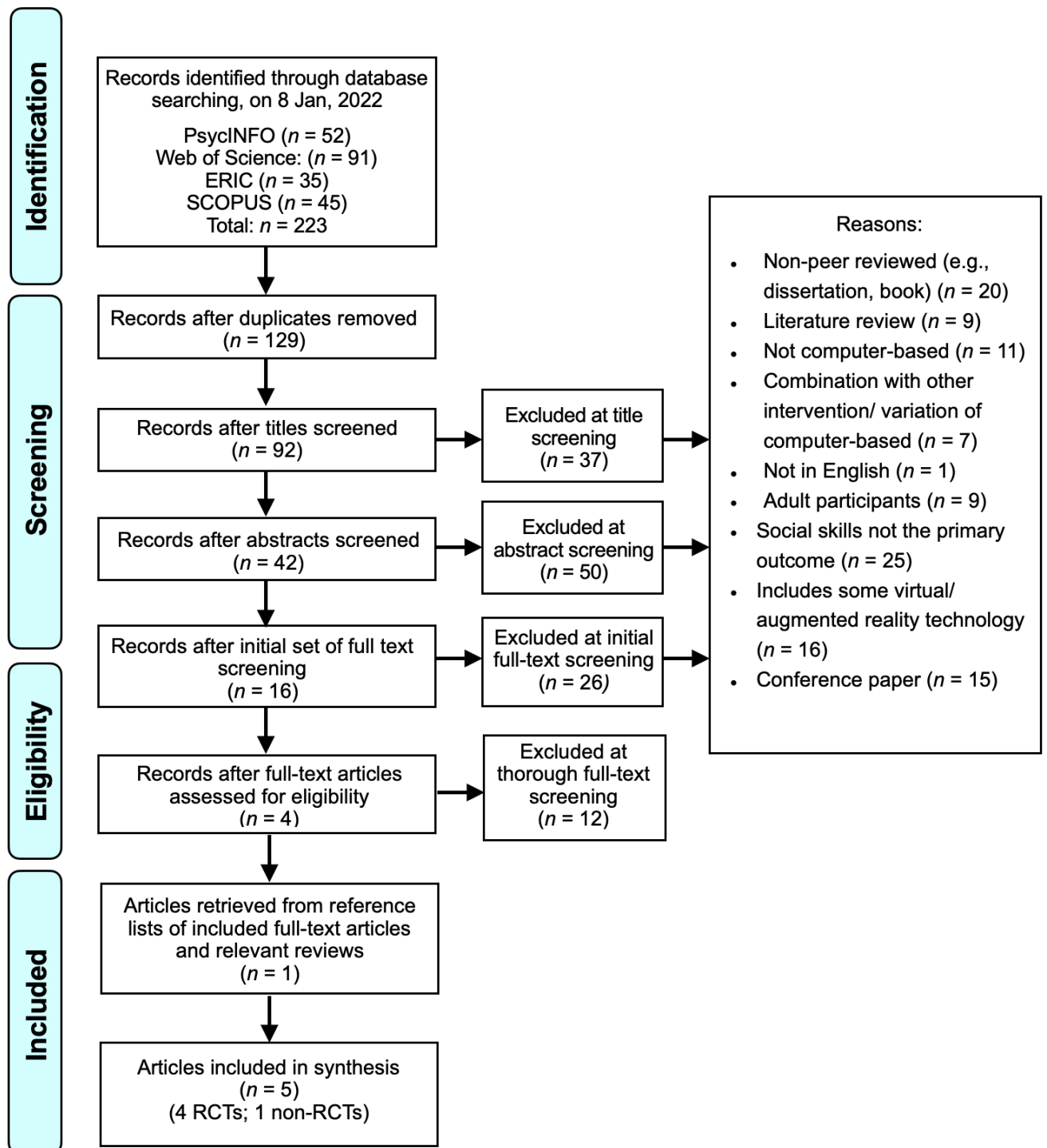


Figure 1: Chronological PRISMA Study Flow Diagram

Adapted from 'Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement', by D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, and The PRISMA Group, 2009, PLoS Med 6(6). Copyright by the Public Library of Science.

Table 3: Final studies included in this systematic review

1.	Beaumont, R., Walker, H., Weiss, J., & Sofronoff, K. (2021). Randomized controlled trial of a video gaming-based social skills program for children on the autism spectrum. <i>Journal of autism and developmental disorders</i> , 1-14.
2.	Hopkins, I. M., Gower, M. W., Perez, T. A., Smith, D. S., Amthor, F. R., Wimsatt, F. C., & Biasini, F. J. (2011). Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. <i>Journal of autism and developmental disorders</i> , 41(11), 1543-1555.
3.	Ben-Sasson, A., Lamash, L., & Gal, E. (2013). To enforce or not to enforce? The use of collaborative interfaces to promote social skills in children with high functioning autism spectrum disorder. <i>Autism</i> , 17(5), 608-622.
4.	Rice, L. M., Wall, C. A., Fogel, A., & Shic, F. (2015). Computer-assisted face processing instruction improves emotion recognition, mentalizing, and social skills in students with ASD. <i>Journal of autism and developmental disorders</i> , 45(7), 2176-2186.
5.	Beaumont, R., & Sofronoff, K. (2008). A multi-component social skills intervention for children with Asperger syndrome: The Junior Detective Training Program. <i>Journal of Child Psychology and Psychiatry</i> , 49(7), 743-753.

3.2 Weight of Evidence

In order to critically appraise the evidence-base for the included studies, Weight of Evidence (WoE) framework (Gough, 2007) was used. This WoE framework provides a systematic way to appraise in three main domains:

Weight of Evidence A (WoE) focuses on the judgement of the methodology quality employed in the study (Gough, 2007). This review used a modified version of the Kratochwill (2003) APA Task Force on Evidence Based Intervention in School

Psychology for group-based designs. The criteria used and their subsequent ratings are found in Appendix G, whilst all studies' coding protocol can be found in Appendix I.

Weight of Evidence B (WoE) reviews the relevance of the study design to answering the review question (Gough, 2007). The criteria and ratings are included in Appendix H, which is based on Petticrew and Roberts' (2003) typology of evidence criteria.

Weight of Evidence C consists of the study's relevance in relation to the review question (Gough, 2007). The criteria for each study is available in Appendix I.

All three separate WoE (A, B, C) were combined and produced an overall weighting score (Weight of Evidence D) that indicates the extent which each study adds value to the evidence-based research when answering the review question (Gough, 2007). Each study's weighting score ranged from 1-3, which a rating of ≤ 1.4 is 'low', 1.5-2.4 is 'medium' and ≥ 2.5 deemed 'high'. Details on the WoE D scores are included in Table 4 below.

Table 4: Overview of Weight of Evidence (WoE) ratings
(In order of WoE D ratings)

Study	WoE A: Methodological Quality	WoE B: Methodological Relevance	WoE C: Topic Relevance	WoE D: Overall Weight of Evidence
Beaumont et al. (2021)	2.8	3	3	2.9 (High)
Beaumont and Sofronoff (2008)	2.8	2	3	2.6 (High)
Hopkins et al. (2011)	1.8	2	3	2.3 (Medium)
Rice et al. (2015)	1.5	3	2	2.2 (Medium)
Ben-Sasson et al. (2013)	1.5	1	2	1.5 (Medium)

WoE rating descriptors

Overall Quality	Average Score
High	≥ 2.5
Medium	1.5 – 2.4
Low	≤ 1.4

3.2.1 Study Participants

In this review, five studies were included, of which 211 ASC participants were aged 6 to 15 years old from Australia, United States, and Israel. However, one study did not specify the recruitment process (Hopkins et al., 2011).

One study reported a significant portion of children with ASC had a co-morbid diagnosis, with the most common being ADHD (Beaumont et al., 2021). This reporting of comorbidity improves generalisability of children with ASC because comorbidity is common in clinical practice. All studies excluded children who were involved in another socio-emotional skills program or concurrent psychological treatment. Meeting the diagnosis status for ASC was one of the main criteria for WoE C, since this review evaluates the effects of computer-based interventions for children already have the threshold diagnosis of autism.

3.2.2 Study Design

Four studies included were randomised control designs, and one study was nonrandomised block design (Ben-Sasson et al., 2013). Three of the studies blinded their participants in control group with an alternative computer-based game (Beaumont et al., 2021; Hopkins et al., 2011; Rice et al., 2015). The participants were unaware whether they were in an intervention group or control group.

Only two studies included a follow up phase to assess maintenance of outcomes (Beaumont et al., 2021; Beaumont & Sofronoff, 2008), which supports the efficacy of the intervention, contributing to their higher WoE A rating for 'Educational/ Clinical Significance'. Beaumont et al., 2021 reported a 6 week follow up, and Beaumont and Sofronoff (2008)'s study reported a 6 week and 5 months follow up. However, the two studies that included follow up were conducted by the same author at different time

periods, therefore this should be interpreted with caution due to potential risk of bias in selection of the reported result.

3.2.3 Interventions

The interventions in this review varied in their content, format, and duration. One study (Beaumont et al., 2021) implemented 'Secret Agent Society' computer-based game intervention that targeted social and emotional skills. Group sessions of up to 3 parent participants were invited to a 10-week intervention, each 30 minutes, and a six-week follow-up.

'FaceSay' computer-based intervention was used in two studies (Hopkins et al., 2011; Rice et al., 2015). Post-test measures were collected two weeks after the final intervention session. Similarly, Rice et.al (2015)'s intervention implementation and data collection were done by research staff at the school site. Teachers were blinded to the intervention group, however the study did not report the number of sessions, duration of sessions, and duration of implementation. Time gap duration between pre and post-test was missing.

Ben-Sasson et al. (2012) used 'Collaborative Puzzle Game' (CPG), a game with 18 puzzles on a touch screen table. Children were asked to collaborate and move the puzzles together in pairs by touching the digital piece at the same time, then drag and release the puzzle. However, number and duration of sessions were unknown.

Lastly, 'The Junior Detective Program' (Beaumont & Sofronoff, 2008) was a computerized intervention with four components: group Ss training, parent training,

teacher handouts and a computer game. The aim was to improve children with ASC in engagement of reciprocal positive interactions and responding appropriately to others' behaviour.

All studies followed similar protocols in increasing the awareness of SSs, which included a social story, enhance facial and emotional understanding, joint attention, initiation, and social responses. Three interventions (Ben-Sasson et al., 2012; Hopkins et al., 2011; Rice et al., 2015) were conducted on school site, whereas two interventions were conducted online at home or in research labs, respectively in Beaumont et al. (2021) and Beaumont & Sofronoff (2008). Three interventions were conducted in schools, which teachers were trained to deliver the computer-based programs. However, Beaumont et al. (2021) and Beaumont & Sofronoff (2008)'s study that were conducted at home or research labs trained parents as the intervention-delivery agents. All studies had differing implementation of interventions, such as player mode (individual vs. multiplayer) and media representations (e.g. human avatars, puzzle pictures, cartoon characters). More details of the computer-based intervention are explored in Appendix D.

Limited studies reported measures to control intervention fidelity. All studies provided software training sessions to the children. Beaumont et al. (2012) received the highest fidelity rating on WoE A Intervention Fidelity with 'Strong Evidence'. The study included on-going weekly supervision for trouble-shoot technical implementation and child behaviour management tips to parents for the intervention (e.g. incidental teaching, praising desirable behaviour). The second highest fidelity rating was Beaumont & Sofronoff (2008)'s study, with a rating of 'Promising Evidence'. Both studies suggest

that the outcome results reflect a reliable uniform intervention among participants. This finding speaks the importance of a higher certainty in outcome results from the intervention itself, instead of individual differences in the intervention protocol. See Appendix G for rigorous appraisal of intervention fidelity.

3.2.4 Measures

Primary outcome measure of SSs questionnaire was included in all studies as a baseline and posttest measurement, and two studies (Beaumont et al., 2021; Beaumont & Sofronoff, 2008) included follow up measure of child outcomes. Beaumont et al. (2021) and Beaumont & Sofronoff (2008) used the Social Skills Questionnaire (SSQ; Spence 1995) parent and teacher report. However, these studies using SSQ did not mention the subtests or domains of measurement, which is unclear in what areas of SS the tool is measuring. These two studies also used the same Emotion Regulation and Social Skills Questionnaire (ERSSQ; Beaumont & Sofronoff 2008).

Ben-Sasson et al. (2013) Rice et al. (2015) used the same SS measurement but different year editions of Social Responsiveness Scale (SRS; Costantino & Gruber, 2002/2005). Ben-Sasson et al. (2013)'s SRS reported internal reliability and test-retest reliability. Finally, Hopkins et al. (2011) used the Social Skills Rating System (SSRS; Gresham & Elliot, 1990) with 38 items by parent report.

Besides SS questionnaire, behavioural observations were conducted in three studies (Hopkins et al., 2011; Ben-Sasson et al., 2013; Rice et al., 2015). SS observations

were done by two researchers blinded during recess or lunch time. Ben-Sasson et al. (2013) used Friendship Observation Scale (FOS; Bauminger et al. 2005). However, the measurement tool was created by the experimenter and based on an unpublished study in 2005, which impacted on a lower score in WoE A Measurement section. In summary, each of these studies spoke of the scales' reliability ratings from different SS measurement tools and observational methods, all reflected in their WoE A Measurement section. See Table 5 for concise summary on studies' primary outcome measurements.

Table 5: Primary outcome measurements of social skills questionnaire and social skills observation

Studies	Social Skills Questionnaire			Social Skills Observation	
		Parent Report	Teacher Report		
Beaumont et al. (2021)	Social Skills Questionnaire (SSQ; Spence 1995)	✓	✓	<ul style="list-style-type: none"> Rate how often a child displays 30 social skills on a 3-point Likert scale 	N/A
	Emotion Regulation and Social Skills Questionnaire (ERSSQ; Beaumont & Sofronoff 2008)	✓	✓	<ul style="list-style-type: none"> 26 item scale on specific emotion recognition, emotion regulation and social skills taught in the intervention 	
Hopkins et al. (2011)	Social Skills Rating System (SSRS; Gresham & Elliot, 1990)	✓		<ul style="list-style-type: none"> 38 items on cooperation, assertion, responsibility, and self-control 	Social Skills Observation (during recess/ free time) <ul style="list-style-type: none"> Two 5-min observations Positive social interaction, low-level social interaction, and negative social interaction 2 observers (blind)
Ben-Sasson et al. (2013)	Social Responsiveness Scale (SRS; Constantino & Gruber, 2005)	✓		<ul style="list-style-type: none"> 65 items on social impairments and repetitive behaviours 5 subscales: social cognition, social awareness, social communication, social 	Friendship Observation Scale (FOS; Bauminger et al. 2005) <ul style="list-style-type: none"> Goal-directed behaviours, prosocial behaviour, conversation, and nonverbal interaction Unknown number of observers or whether observers are blinded

Studies	Social Skills Questionnaire		Social Skills Observation		
	Parent Report	Teacher Report			
				motivation, and autistic mannerisms	
Rice et al. (2015)	Social Responsiveness Scale, Second Edition (SRS-2; Costantino & Gruber, 2002)	✓	✓	<ul style="list-style-type: none"> 65 items on social impairment 5 subscales: social cognition, social awareness, social communication, social motivation, and restricted interests and repetitive behaviours 	Social Skills Observations (during recess and lunch time) <ul style="list-style-type: none"> 10 min observations Positive interactions: spontaneous initiations and positive interactions. Negative interactions: unpleasant social behaviours (e.g. physical or verbal aggressiveness) 2 observers (blind)
Beaumont & Sofronoff (2008)	Social Skills Questionnaire (SSQ: Spence, 1995a)	✓	✓	<ul style="list-style-type: none"> 30 items on child's social behaviour in the past four weeks. 	N/A
	Emotion Regulation and Social Skills Questionnaire (ERSSQ; Beaumont & Sofronoff 2008)		Unknown	<ul style="list-style-type: none"> 27 item scale of a 5-point Likert Scale on specific emotion recognition, emotion regulation and social skills taught in the intervention 	

3.2.5 Results Outcomes and Effect Sizes

All studies yield a medium or large effect sizes from SS computer-based intervention (see Table 6). These comparison of effect sizes were drawn from different methodological designs and SS computer-based interventions. Majority of studies reported effect sizes as n^2p , Wilcoxon signed-rank test, and Spearman's Rho correlation, which were converted or calculated to Cohen's d using the effect size calculator provided from Campbell Collaboration (Wilson, n.d.).

The largest effect sizes of social skills questionnaire were seen in Ben-Sasson et al. (2013)'s study, however only within-group main effect of the Social Responsiveness Scale (Constantino & Gruber, 2005) subscales was reported. It is uncertain how effective the intervention is compared between with control group. Two studies used the same measurement tool of Social Skills Questionnaire (SSQ; Spence 1995) reported large effect (Beaumont et al., 2021; Beaumont & Sofronoff, 2008). SSQ-Parent report showed larger effect than SSQ-Teacher report in both studies in pre-posttests. While these findings are encouraging, interpreting results with caution is warranted because parents were both the intervention delivery agents and evaluators in both studies, making them more susceptible to responder bias. As seen in Beaumont et al. (2021)'s study, the treatment condition on SSQ-Parent and ERSSQ-Parent have significant effect of improvements compared to Teacher reports. In contrary, Beaumont and Sofronoff (2008)'s follow-up measurement showed stronger treatment effect on teacher-report relative to parent-report measures. Significant effect sizes from teacher report follow-up at 6-weeks should also interpret with caution as there is a high attrition rates of teacher reports from post-test (26 children teacher reports) to follow up (19 children teacher reports). In the follow-up test, less than 20%

of sample teacher reports were included in the statistical analysis. Moreover, teacher reports scores remained to one and a half standard deviations below the normative mean ($M = 52.28$, $SD = 10.09$) at post-treatment. Therefore, the improvement of clinically or educationally significant from teacher report in social skills measure is questionable, which impacted on the WoE A: Primary Outcomes Statistically Significant criterion.

Mixed results were shown for SS observation. Three studies reported SS observations, however there is a wide range of effect from insignificant to significant reports ($d = 1.76$; Ben-Sasson et al., 2013). Only Hopkins et al. (2011) showed large effect size of SS observations, whereas other studies showed insignificant effect.

Table 6: Primary Outcome Measures' Effect Sizes and Descriptors

Study	Number of Participants	Primary Outcome Measure	Internal Consistency/ Reliability	Comparison (Control group)	Statistical Test	Statistical Comparison	Effect Size (Cohen's d)	Descriptor*	WoE D
Beaumont et al. (2021)	70	Social Skills Questionnaire (SSQ; Spence 1995)	$\alpha = .88$ (parent) and $.94$ (teacher)	Virtual jigsaw puzzles	Mixed-model ANOVA	Group x time interaction	Parent = 1.04^* ($p < .0005$) Teacher = $.51^*$ ($p = .04$)	Large Medium	High
		Emotion Regulation and Social Skills Questionnaire (ERSSQ; Beaumont & Sofronoff 2008)	$\alpha = .95$ (parent) and $.94$ (teacher)		Mixed-model ANOVA	Group x time interaction	Parent = 1.21^* Teacher = $.20$	Large Small	
Hopkins et al. (2011)	49	Social Skills Rating System (SSRS; Gresham & Elliot, 1990)	$\alpha = .87-.90$ Test re-test reliability = $.97$	11 children with autism were asked to play with <i>Tux Paint</i> , a drawing software for children (www.tuxpaint.org)	ANCOVA	Group x time interaction	0.29^*	Small	Medium
		Social Skills Observation (during recess or free time)	Cohen's kappa (Cohen, 1960) between two raters: 1. Positive social interaction = $.95$		ANCOVA	Group x time interaction	1.34^*	Large	

Study	Number of Participants	Primary Outcome Measure	Internal Consistency/ Reliability	Comparison (Control group)	Statistical Test	Statistical Comparison	Effect Size (Cohen's d)	Descriptor*	WoE D
			2. Low-level social interaction = .74						
			3. Negative social interaction = .86						
Ben-Sasson et al. (2013)	12	Friendship Observation Scale (FOS, Bauminger et al. 2005)	Missing	Free Play (FP) mode: children play puzzle pieces independently	Wilcoxon signed-rank test	Treatment vs. Control group for Positive Social (PSI) subscale	1.76* (p = .02)	Large	Medium
					Wilcoxon signed-rank test	Treatment vs. Control group for Negative Social Interaction (NSI)	Insignificant (Missing effect size)	N/A	
					Spearman's Rho correlation	Within group-treatment group, on the following subtests:			
		Social Responsiveness Scale (SRS; Constantino & Gruber, 2005)	α= .97 Test re-test reliability = .85						
						1. Social Awareness	1.46* (p = .04)	Large	
						2. Social Communication	1.25* (p = .08 marginally significant)	Large	

Study	Number of Participants	Primary Outcome Measure	Internal Consistency/ Reliability	Comparison (Control group)	Statistical Test	Statistical Comparison	Effect Size (Cohen's d)	Descriptor*	WoE D
Rice et al. (2015)	31	Social Responsiveness Scale, Second Edition (SRS-2; Costantino & Gruber, 2002)	Missing	SuccessMaker, a reading program tailored to participant's reading level in grades K-8.	ANCOVA	3. Expressive Communication Group x time interaction	1.43* ($p = .05$) Parent = Missing Teacher = 0.79*	Large N/A Large	Medium
						ANCOVA	Group x time interaction	Insignificant (Unable to calculate effect size due to insufficient information)	
Beaumont and Sofronoff (2008)	49	Social Skills Questionnaire (SSQ: Spence, 1995a)	Missing	Wait-list control	Repeated-measures ANOVA	Group x time interaction	Pre-Post: Teacher = 1.8* Parent = 1.06* Follow up: Teacher = 1.32* (6-week) Parent = 0.51* (5-month)	Large Large Large Medium	High

Study	Number of Participants	Primary Outcome Measure	Internal Consistency/ Reliability	Comparison (Control group)	Statistical Test	Statistical Comparison	Effect Size (Cohen's d)	Descriptor*	WoE D
		Emotion Regulation and Social Skills Questionnaire (ERSSQ)	$\alpha = .89$		Repeated-measures ANOVA	Group x time interaction	0.46	Small	

* = indicate significant main effect

Effect size thresholds descriptors from Cohen's d (1992): 0.2 = small effect size, 0.5 = medium effect size, and 0.8 = large effect size.

Cronbach alpha (α) $\geq .9$ Excellent, .9-0.8 Good, 0.8-0.7 Acceptable, 0.7-0.6 Questionable, 0.6-0.5 Poor, 0.5 > Unacceptable

Section 4: Conclusion

4.1 Discussion

To our knowledge, this review is the first independent systematic review on the effectiveness of computer-based game interventions to improve social skills for school-aged children with Autistic Spectrum Conditions (ASC). The current review synthesized five studies, of which four were randomised controlled trials, and one quasi-experimental design. All studies included a comparison control group, and only two studies included a follow up measurement (Beaumont et al., 2021; Beaumont & Sofronoff, 2008). Two studies received an Overall Weight of Evidence rating of High (Beaumont et al., 2021; Beaumont & Sofronoff, 2008), and three studies with Medium (Ben-Sasson et al., 2013; Hopkins et al., 2011; Rice et al., 2015) (see Table 4). Primary outcomes of social skills levels were measured by (1) parent and teacher questionnaires, and (2) behavioural observations (e.g., school recess/ lunch time). SS questionnaire findings were encouraging from parent and teacher reports, however it should be interpreted with caution. This is because parents or teachers were the intervention delivery agents and evaluators, which are more susceptible to responder and reporting bias. Therefore, the study findings of parent and teacher self-report shown more significant improvements, compared to school setting behavioural observations by blinded researchers.

4.2 Limitations and Future Research

One of the main limitations in this review is the overall small sample size of 211 participants from the included five studies. Having a small sample size suggests a weak relationship between the intervention and outcome measures. This is because

sample size is a core determinant of power, enabling the detection of statistical, clinical or educational significance more accurately (West et al., 2002). Moreover, it may appear that these studies report a medium to high effect size on primary outcomes. However, the primary outcome of SS varied with a large range in effect sizes; $d = 0.29$ (Hopkins et al., 2011) to $d = 1.80$ (Beaumont & Sofronoff, 2008). Although these effect sizes were statistically significant, no studies reported confidence intervals. The failure to report confidence intervals suggest an unreliable statistic outcome of effect sizes being over or underestimated.

At first glance, two randomised control trials (RCT) rated High in Overall Weight of Evidence (WoE D) (Beaumont et al., 2021; Beaumont & Sofronoff, 2008). However, both studies were conducted by the same author (Beaumont, R.) who created the computer-based intervention and the primary outcome measurement of 'Emotion Regulation and Social Skills Questionnaire'. This may increase the risk of bias in the blinding of outcome assessment and selective reporting. Moreover, these two RCTs are the only studies that included follow-up measurement and reported significant effect. However, both studies should be interpreted with caution because there is a high attrition rate especially in the follow up phase, which leads to attrition bias. Therefore, research on computer-based intervention in improving SS is still in its infancy. More research is needed with larger and powered studies, especially during follow-up, to ascertain the true effectiveness.

All studies' SS outcomes were answered by parents or teachers. There is a need to explore the impact of the intervention from the child's perspective. Involving children

and young people in discussion to understand the effectiveness of the support or intervention is key for evaluation (SEND Code of Practice, 2015).

Although the included studies were conducted in Australia, USA, and Israel, none of the studies were completed within the UK which may reduce the generalisability. There is a need for further well-designed RCT, with sufficient statistical power based on other countries to expand the external validity of the current evidence base. Furthermore, all studies published on computer-based interventions to improve SS for children with ASC report significant improvement. Future research on computer-based intervention should publish insignificant studies because statistically insignificant results are as important as significant results. This helps to minimise interpretive bias in systematic reviews (Hewitt et al., 2008). Future reviews may also include unpublished work that meets inclusion and exclusion criteria, to reduce publication bias or the 'file drawer problem' (Rosenthal, 1979).

4.3 Implications and Recommendations for Educational Context and Educational Psychology Practice

This review sheds light on developing game-based computer interventions to support children with an ASC level of SS. Technology could provide a therapeutic intervention for families and schools who face geographical or financial barriers in accessing the original face-to-face program. The included studies have shown implementation of intervention is feasible in school and home settings.

All studies included quantitative measures of SS, however a dynamic assessment by an EP within computer-based intervention could further understanding of the social context of learning around the child's system. Therefore, EPs may suggest a bank of strategies for the child in real-life social contexts.

Nevertheless, one of the biggest challenges EPs face in implementing SS intervention is gaining commitment from parents and teachers. SS improvement is often supported indirectly by parents and teachers, therefore this review supports the implementation of gamified computer-based interventions at home and school. However, in response it could be argued whether computerised interactions can replicate human interaction such as shared attention and non-verbal communication. Such innovative SS support requires further research into whether computer-based interventions could adapt and differentiate various real-life interactions. Therefore, this review examined the potential in play-based computer interventions for children with ASC, due to its affordability and accessibility of services. There is a pressing need for further research, especially in the global crisis of restricted face-to-face contact during COVID-19.

References

- Abd El-Sattar, H. K. H. (2008). A Novel Interactive Computer-Based Game Framework: From Design to Implementation. 2008 International Conference Visualisation. <https://doi.org/10.1109/VIS.2008.29>
- Atherton, G., & Cross, L. (2021). The Use of Analog and Digital Games for Autism Interventions. *Frontiers in Psychology*, 3049. <https://doi.org/10.3389/fpsyg.2021.669734>
- Beaumont, R., & Sofronoff, K. (2008). A multi-component social skills intervention for children with Asperger syndrome: The Junior Detective Training Program. *Journal of Child Psychology and Psychiatry*, 49(7), 743-753. <https://doi:10.1111/j.1469-7610.2008.01920.x>
- Beaumont, R., Walker, H., Weiss, J., & Sofronoff, K. (2021). Randomized controlled trial of a video gaming-based social skills program for children on the autism spectrum. *Journal of autism and developmental disorders*, 1-14. <https://doi.org/10.1007/s10803-020-04801-z>
- Ben-Sasson, A., Lamash, L., & Gal, E. (2013). To enforce or not to enforce? The use of collaborative interfaces to promote social skills in children with high functioning autism spectrum disorder. *Autism*, 17(5), 608-622. <https://doi.10.1177/1362361312451526>
- Bodrova, E., & Leong, D. J. (2015). Vygotskian and Post-Vygotskian Views on Children's Play. *American Journal of Play*, 7(3), 371-388.
- Chawarska, K., Ye, S., Shic, F., & Chen, L. (2016). Multilevel differences in spontaneous social attention in toddlers with autism spectrum disorder. *Child development*, 87(2), 543-557. <https://doi:10.1111/cdev.12473>

Department for Education & Department of Health (2015). Special educational needs and disability code of practice: 0 to 25 Years. Available at:

<https://www.gov.uk/government/publications/send-code-of-practice-0-to-25>

(Accessed: 20th June 2022).

Filsecker, M., & Hickey, D. T. (2014). A multilevel analysis of the effects of external rewards on elementary students' motivation, engagement and learning in an educational game. *Computers & Education*, 75, 136-148.

<https://doi.org/10.1016/j.compedu.2014.02.008>

Fletcher-Watson, S., Petrou, A., Scott-Barrett, J., Dicks, P., Graham, C., O'Hare, A., Pain, H., & McConachie, H. (2016). A trial of an iPad™ intervention targeting social communication skills in children with autism. *Autism*, 20(7), 771-782.

<https://doi.org/10.1177/1362361315605624>

Gray, P. (2011). The Special Value of Children's Age-Mixed Play. *American Journal of Play*, 3(4), 500-522.

Grossard, C., Grynspan, O., Serret, S., Jouen, A.-L., Bailly, K., & Cohen, D. (2017). Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (ASD). *Computers & Education*, 113, 195-211.

<https://doi.org/10.1016/j.compedu.2017.05.002>

Harrower, J. K., & Dunlap, G. (2001). Including children with autism in general education classrooms: A review of effective strategies. *Behavior modification*, 25(5), 762-784.

Hewitt, C. E., Mitchell, N., & Torgerson, D. J. (2008). Listen to the data when results are not significant. *Bmj*, 336(7634), 23-25.

- Jones, P., & Bawazir, R. (2018). A theoretical framework on using social stories with the creative arts for individuals on the autistic spectrum. *International Journal of Medical and Health Sciences*, 11(9), 533-541.
- Lancy, D. F., & Grove, M. A. (2017). Marbles and Machiavelli: the role of game play in children's social development. *American Journal of Play*, 3(4).
- Li, Y., & Elmaghraby, A. S. (2014). A framework for using games for behavioral analysis of autistic children. 2014 Computer Games: AI, Animation, Mobile, Multimedia, Educational and Serious Games (CGAMES),
- M Schaller, U., & Rauh, R. (2017). What difference does it make? Implicit, explicit and complex social cognition in autism spectrum disorders. *Journal of autism and developmental disorders*, 47(4), 961-979. <https://doi.org/10.1007/s10803-016-3008-x>
- Marsh, J., Plowman, L., Yamada-Rice, D., Bishop, J., & Scott, F. (2016). Digital play: A new classification. *Early Years*, 36(3), 242-253. <https://doi.org/10.1080/09575146.2016.1167675>
- O'Keeffe, C., & McNally, S. (2021). A systematic review of play-based interventions targeting the social communication skills of children with Autism Spectrum Disorder in educational contexts. *Review Journal of Autism and Developmental Disorders*, 1-31. <https://doi.org/10.17605/OSF.IO/FYEQS>
- Roberts, J. M., Paynter, J., Trembath, D., Westerveld, M., & Williams, K. (2020). *Designing a Longitudinal Study of Development of Students With Autism: Focus on Learning, Education, and Meaningful Outcomes*. SAGE Publications Ltd.
- Roman-Urrestarazu, A., van Kessel, R., Allison, C., Matthews, F. E., Brayne, C., & Baron-Cohen, S. (2021). Association of race/ethnicity and social disadvantage

with autism prevalence in 7 million school children in England. *JAMA pediatrics*, e210054-e210054.

<https://doi.org/10.1001/jamapediatrics.2021.0054>

Rosello, B., Berenguer, C., Baixauli, I., García, R., & Miranda, A. (2020). Theory of mind profiles in children with autism spectrum disorder: Adaptive/social skills and pragmatic competence. *Frontiers in psychology*, *11*, 567401.

<https://doi.org/10.3389/fpsyg.2020.567401>

Rosenthal, M., Wallace, G. L., Lawson, R., Wills, M. C., Dixon, E., Yerys, B. E., & Kenworthy, L. (2013). Impairments in real-world executive function increase from childhood to adolescence in autism spectrum disorders.

Neuropsychology, *27*(1), 13. <https://doi.org/10.1037/a0031299>

Rosenthal, R. (1979). The file drawer problem and tolerance for null results.

Psychological bulletin, *86*(3), 638.

Shane, H. C., & Albert, P. D. (2008). Electronic screen media for persons with autism spectrum disorders: Results of a survey. *Journal of autism and developmental disorders*, *38*(8), 1499-1508. <https://doi.org/10.1007/s10803-007-0527-5>

Turan, Z., Avinc, Z., Kara, K., & Goktas, Y. (2016). Gamification and education: Achievements, cognitive loads, and views of students. *International journal of emerging technologies in learning*, *11*(7).

<http://dx.doi.org/10.3991/ijet.v11i07.5455>

Vygotsky, L. S. (1967). Play and its role in the mental development of the child.

Soviet psychology, *5*(3), 6-18.

Wainer, A. L., & Ingersoll, B. R. (2011). The use of innovative computer technology for teaching social communication to individuals with autism spectrum

disorders. *Research in Autism Spectrum Disorders*, 5(1), 96-107.

<https://doi.org/10.1016/j.rasd.2010.08.002>

West, S., King, V., Carey, T. S., Lohr, K. N., McKoy, N., Sutton, S. F., & Lux, L.

(2002). Systems to rate the strength of scientific evidence. *Evidence report/technology assessment (Summary)*(47), 1-11.

Whyte, E. M., Smyth, J. M., & Scherf, K. S. (2015). Designing serious game interventions for individuals with autism. *Journal of autism and developmental disorders*, 45(12), 3820-3831. <https://doi.org/10.1007/s10803-014-2333-1>

Wilson, D. B., Ph.D. (n.d.). Practical Meta-Analysis Effect Size Calculator

[Online calculator]. Retrieved June 21, 2022,

from <https://campbellcollaboration.org/research-resources/effect-size-calculator.html>

Appendix

Appendix A: Search strategy in PsycINFO

- 1 exp autism spectrum disorders/ (48991)
- 2 (autism or autis*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word] (62746)
- 3 exp computer games/ (8341)
- 4 (video game or game or mobile or computer-based or gamification or serious game or app or iPad or tablet or computer or digital or online game or CBI).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word] (232915)
- 5 (computer* adj3 (intervention* or technique* or therap* or treat* or game*)).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word] (15121)
- 6 exp friendship/ (10253)
- 7 (Social skill* or social relationship or peer relationship or friend*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures, mesh word] (77877)
- 8 1 or 2 (62746)
- 9 3 or 4 or 5 (233463)
- 10 6 or 7 (77877)
- 11 8 and 9 and 10 (52)

Appendix B: List of excluded studies following full paper screening

Article	Exclusion criteria number(s) (See Table 2, p.12)
Grossard, C., Grynspan, O., Serret, S., Jouen, A. L., Bailly, K., & Cohen, D. (2017). Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (ASD). <i>Computers & Education</i> , 113, 195-211.	8 – Study was a systematic review. 5(a) This study utilised virtual reality environment that is different from this review’s purpose. The purpose of this review’s play-based computer intervention is to be easily implemented at home and at school
Bernardini, S., Porayska-Pomsta, K., & Smith, T. J. (2014). ECHOES: An intelligent serious game for fostering social communication in children with autism. <i>Information Sciences</i> , 264, 41-60.	7 – Study did not have a control group
Gallup, J., & Serianni, B. (2017). Developing friendships and an awareness of emotions using video games: Perceptions of four young adults with autism. <i>Education and Training in Autism and Developmental Disabilities</i> , 52(2), 120-131.	1(a) and 1(b) – Study did not mention whether children met the clinical significance of autism and cognitive range
O’Hagan, S., & Hebron, J. (2017). Perceptions of friendship among adolescents with autism spectrum conditions in a mainstream high school resource provision. <i>European Journal of Special Needs Education</i> , 32(3), 314-328.	5(b) – Study did not focus on an educational designed serious game
Chung, U. S., Han, D. H., Shin, Y. J., & Renshaw, P. F. (2016). A prosocial online game for social cognition training in adolescents with high-functioning autism: an fMRI study. <i>Neuropsychiatric disease and treatment</i> , 12, 651.	4(c) – Participants reported to have medical seizures and new medication during the study
Friedrich, E. V., Suttie, N., Sivanathan, A., Lim, T., Louchart, S., & Pineda, J. A. (2014). Brain-computer interface game applications for combined neurofeedback and	6 – Primary outcome did not focus on social skills

biofeedback treatment for children on the autism spectrum. *Frontiers in neuroengineering*, 7, 21.

5(b) – Targeted learning objectives are not clear in the game

Thomeer, M. L., Smith, R. A., Lopata, C., Volker, M. A., Lipinski, A. M., Rodgers, J. D., ... & Lee, G. K. (2015). Randomized controlled trial of mind reading and in vivo rehearsal for high-functioning children with ASD. *Journal of autism and developmental disorders*, 45(7), 2115-2127.

6 – The focus of social skills was a secondary outcome

Barakova, E. I., Gillesen, J. C., Huskens, B. E., & Lourens, T. (2013). End-user programming architecture facilitates the uptake of robots in social therapies. *Robotics and Autonomous Systems*, 61(7), 704-713.

5(a) – The study utilised other means of technology, such as robots and virtual reality

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Meir-Goren, N., ... & Golan, O. (2017). 'Emotiplay': a serious game for learning about emotions in children with autism: results of a cross-cultural evaluation. *European child & adolescent psychiatry*, 26(8), 979-992.

6 – Primary outcome did not focus on social skills

Malinverni, L., Mora-Guiard, J., Padillo, V., Valero, L., Hervás, A., & Pares, N. (2017). An inclusive design approach for developing video games for children with Autism Spectrum Disorder. *Computers in Human Behavior*, 71, 535-549.

5(a) – The study utilised other means of technology, such as robots and virtual reality

Mairena, M. Á., Mora-Guiard, J., Malinverni, L., Padillo, V., Valero, L., Hervás, A., & Pares, N. (2019). A full-body interactive videogame used as a tool to foster social initiation conducts in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 67, 101438.

5(a) – The study utilised other means of technology, such as robots and virtual reality

Friedrich, E. V., Sivanathan, A., Lim, T., Suttie, N., Louchart, S., Pillen, S., & Pineda, J. A. (2015). An effective neurofeedback intervention to improve social interactions in children with autism spectrum disorder. *Journal of autism and developmental disorders*, 45(12), 4084-4100.

6 – Primary outcome did not focus on social skills

5(c) – The game did not include a narrative storyline

Appendix C: Mapping the Field

Study	Design	Sample				Country	Autism Diagnosis	Implementation Method	Intervention Method	Computer Intervention
		Size	Control; intervention group	Male %	Age range					
Beaumont et al. (2021)	RCT	70	35;35	86	7-12 (<i>M</i> = 9.89, <i>SD</i> = 1.37)	Australia (Queensland, New South Wales, Victoria)	<ul style="list-style-type: none"> • DSM-V • Cognitive abilities in average range: Full Scale IQ or Verbal and Perceptual indices 	<ul style="list-style-type: none"> • 150-minute therapist-led parent webinar 	<ul style="list-style-type: none"> • 10 sessions, 30 minutes each • Weekly online support and monitoring were provided 	<i>Secret Agent Society</i>
Hopkins et al (2011)	RCT	49	24;25	90	6-15 (<i>M</i> = 10.17)	United States (Alabama)	<ul style="list-style-type: none"> • DSM-IV • CARS • Cognitive abilities in average range (mental age between 6-10 years old) 	<ul style="list-style-type: none"> • Unknown 	<ul style="list-style-type: none"> • 12 sessions, 6 weeks • 10-25 min per session 	<i>FaceSay</i>
Ben-Sasson et al (2013)	Non-randomised block design	12	6;6	100	8-11 (<i>M</i> = 9.28, <i>SD</i> = 0.94)	Israel	<ul style="list-style-type: none"> - DSM-IV-TR (2000) - ADOS Module 3 (Lord et al., 2000) - Nonverbal IQ of 80 or above from WISC-R (Wechsler, 1974) 	<ul style="list-style-type: none"> • Unknown 	<ul style="list-style-type: none"> • 6 sessions, one puzzle per session 	<i>Collaborative Puzzle Game</i>

Study	Design	Sample			Country	Autism Diagnosis	Implementation Method	Intervention Method	Computer Intervention
Rice et al (2015)	RCT	31	15;16	90	5-11 (<i>M</i> = 7.7)	United States (California)	<ul style="list-style-type: none"> • DSM-IV • WISC-III or WISC-IV • FSIQ >70 (<i>M</i> = 101, <i>SD</i> = 14.45) 	<ul style="list-style-type: none"> • Unknown number of sessions • 6 weeks 	FaceSay
Beaumont and Sofronoff (2008)	RCT	49	23; 26	90	7.5-11 (<i>M</i> = 9.7)	Australia (Queensland)	<ul style="list-style-type: none"> • DSM-IV-TR • WISC-III pro-rated IQ score of 85 or above • CAST (Scott, Baron-Cohen, Bolton & Brayne, 2002) 	<ul style="list-style-type: none"> • 2-hour training session with the parents by the chief investigator • 8 sessions • 1 session per week 	Junior Detective Training Program

Appendix D: Computer-based intervention for included studies

Study	Computer-based intervention	Age	Capabilities	Origin/ developer	Device	Minimum system requirement	Price	Player mode
Beaumont et al. (2021)	<i>Secret Agent Society</i>	5 years or older	412 emotion samples in three main applications (e.g. Emotions library, learning centre and game zone)	Golan and Baron-Cohen (2006)	Computer	Macintosh Edition: OS 9.2 or later with 16 MB Ram; Windows Edition: OS XP or later, Pentium 3 or faster processors recommended, 32 MB RAM, 2.5 GB hard drive disk space; 800 * 600 PSR with 16-bit colour monitor, Sound card and speakers required.	Available in CD-ROM or DVD version; Single user license costs \$125 (USD) and site license costs \$495 (USD)	Individual
Hopkins et al (2011) and Rice et al (2015)	<i>FaceSay</i>	6-14	Interactive approach of video-realistic avatars (humans and animals) to teach face and emotion recognition skills	Symbionica, LLC, San Jose, CA	Touch screen tablet (recommended) or Computer	Windows XP, Windows Vista, Windows 7 and Windows 2000. 1GB for For XP and 2000 2GB for Vista and Windows 7. 300mb of free disk space. CPU Speed 1.75GHZ or faster.	\$349 USD (5 student/ client license); \$299 USD for 1 computer; \$79 USD home edition for 2 children	Individual
Ben-Sasson et al (2013)	<i>Collaborative Puzzle Game</i>	Unknown	18 puzzle pictures	Unknown	DiamondTouch (DT) interface	Unknown	Unknown	Multiplayer mode playing together
Beaumont and Sofronoff (2008)	<i>Junior Detective Training Program</i>	8-11	Raise awareness of their feelings and emotions presented by characters and relate the presented scenario to their own life experiences	Beaumont and Sofronoff	Computer	A CD-ROM based application	Not available commercially	With parents

Appendix E: Summary of participants’ recruitment process from included studies

Study	Recruitment/ selection process
Beaumont et al. (2021)	Autism practitioners and clinics. Flyers to rural and regional areas, advertisements in websites
Hopkins et al (2011)	“From several sources”- Unclear reporting
Ben-Sasson et al (2013)	Special classroom within a public school
Rice et al (2015)	Advertised within the school district in Ventura County. All primary school students.
Beaumont and Sofronoff (2008)	Local newspaper advertisement (Queensland Asperger Syndrome Support Network newsletter), and autism practitioners

Appendix F: Amendments made to the Kratochwill (2003) Coding Protocol

APA Task Force Coding Protocol using Kratochwill (2003) has been used in this review. Details of amendments to the protocol suggested by reasons are outlined below.

Section heading	Section removed/ modified	Rationale
1. General Characteristics	B7 : Coding	Only used for qualitative research
	B8 : Interactive Process Followed	Only used for qualitative research
2. Key features for Coding Studies and Rating Level of Evidence/Support	Section C	This information is being reported in the Mapping the Field table - Primary outcomes are being discussed in the paper and secondary outcomes are not relevant to the review question
	Section D3 – Subjective Evaluation	Since the intervention is undertaken by the parents or teachers to impact their children, the participants in the study are not the target population – hence the category was changed to indicate that subjective behaviour change was evaluated by the parents or teachers ‘the participants’ who were in direct contact with their child – who in this case the intervention outcomes were targeted towards.
	Section D4: Social Comparison	The goal of the research question is whether the intervention is effective at improving social skills, not how these could compare with children with no autism

	Section E: Identifiable components	The interventions selected in these studies were manualised approaches and components are not separate
	Section G: Replication	No replications done in any of the studies included in this review
	Section H: Site of Implementation	Not relevant to the review question and does not impact on the methodology of the studies being included in this review
III. Other Descriptive or Supplemental Criteria to Consider	Section A2: Participant characteristics specified for treatment and control	Reported in the 'Mapping the Field' table and discussed in the review
	Section A4: Receptivity	This information is reported in the review and was not assessed to impact quality of methodology
	Section A5.3: Generalization across persons	Not relevant to study
	Section B: Length of Intervention	Reported in the 'Mapping the Field' table and discussed in the review
	Section C: Intensity/Dosage of Intervention	Reported in the 'Mapping the Field' table and discussed in the review
	Section D. Dosage Response	Not relevant to review question
	Section E: Program Implementer	Reported in 'Primary Outcome' table
	Section H. Cost Analysis Data	Not relevant to review question
	Section J. Feasibility	Not reported in the included studies and not relevant to review question

Appendix G: Weight of Evidence A (WoE A)

Weight of Evidence A is analysed using a modified Kratochwill (2003) coding protocol in relation to measurement, comparison group, appropriate statistical analysis, primary outcome statistically significant, implementation fidelity, and educational/ clinical significance.

The modified criterion is derived from Kratochwill’s (2003) coding protocol in the section, Key Features for Coding Studies and Rating Level of Evidence/ Support. Table 6 outlines the average WoE A score ranges for included studies.

Table 1: Criteria for Measurement

Weighting	Criteria
High (3)	<ul style="list-style-type: none"> • Reliability $\geq .85$ (for all primary outcome measures) • Multiple measurement sources used • Multiple measurement methods used
Medium (2)	<ul style="list-style-type: none"> • Reliability $\geq .7$ (for at least 75% of primary outcome measures) • Multiple measurement sources OR Multiple measurement methods used
Low (1)	<ul style="list-style-type: none"> • Reliability $\geq .5$ (for at least 50% of primary outcome measures) • One source or one method of data is used

Table 2: Criteria for Comparison group

Weighting	Criteria
High (3)	<ul style="list-style-type: none"> • An active control group was used (e.g., alternate treatment or attention placebo) • There is group equivalence by random assignment • Equivalent attrition/mortality between treatment and control groups at post and follow-up • Change agents have been counterbalanced

Medium (2)	<ul style="list-style-type: none"> • A ‘no intervention’ control group was used (e.g. Waitlist or no intervention) • At least 2 of the following are present: Equivalent groups OR counterbalancing of change agents OR equivalent mortality with low attrition • In the case of inequivalent mortality rates between the conditions – then no significant difference must be reported between the groups
Low (1)	<ul style="list-style-type: none"> • A control group is used • At least 1 of the following: Equivalent grouping OR equivalent mortality with low attrition OR counterbalancing of change agents. • In the case of inequivalent mortality rates between the conditions – then no significant difference must be reported between the groups

Table 3: Criteria for Implementation Fidelity

Weighting	Criteria
High (3)	<ul style="list-style-type: none"> • Ongoing supervision/ consultation, coding sessions, or audio/video tapes • Use of manual, including information must be provided to the implementers either: (1) written materials involving an exact and detailed procedures and sequences used or (2) a formal training session that includes the exact and detailed procedures and sequences used
Medium (2)	<ul style="list-style-type: none"> • Ongoing supervision/ consultation, coding sessions, or audio/video tapes • Use of manual, including information must be provided to the implementers either: (1) written materials involving a broad principles and description procedures and sequences used or (2) a formal training session that includes a broad principles and description procedures and sequences used
Low (1)	<ul style="list-style-type: none"> • Ongoing supervision/ consultation, coding sessions, or audio/video tapes <u>OR</u> use of manual

Table 4: Criteria for Educational/ Clinical significance

Weighting	Criteria
High (3)	<ul style="list-style-type: none"> Evidence of support seen in all 3 criteria evaluated during the post-test or follow up phases for most of the participants
Medium (2)	<ul style="list-style-type: none"> Evidence of support seen in 2 out of the 3 criteria evaluated during the post-test or follow up phases for most of the participants
Low (1)	<ul style="list-style-type: none"> Evidence of support seen in 1 out of the 3 criteria evaluated during the post-test or follow up phases for most of the participants.

Table 5: WoE A overall weighting scores for study

(In order of overall WoE A ratings)

Study	Measure	Comparison Group	Fidelity	Educational/ Clinical Significance	Overall WoE A
Beaumont et al. (2021)	3	2	3	3	2.8 (High)
Beaumont and Sofronoff (2008)	3	2	3	3	2.8 (High)
Hopkins et al (2011)	2	2	1	2	1.8 (High)
Rice et al (2015)	2	2	1	1	1.5 (Medium)
Ben-Sasson et al (2013)	2	1	1	2	1.5 (Medium)

Table 6: Average WoE A score ranges

Overall Quality	Average Score
High	≥ 2.5
Medium	1.5 – 2.4
Low	≤ 1.4

Appendix H: Weight of Evidence B (WoE B)

WoE B reviews the methodological relevance to the review question (Gough, 2007). The criteria are adapted from evidence hierarchies (Guyatt et al., 2008), suggesting some research designs are more appropriate for certain studies. Below are the criteria for the included studies in this review. Each study must meet all criteria section to receive the rating.

Weighting	Criteria
High (3)	<p>To get this rating, the study must meet 3 out of 4 criteria</p> <ul style="list-style-type: none"> • There is a randomly assigned control group • Participants are blinded for assignment in experimental or control group through a stringent procedure that is stated in detail • Pre-, Post, and Follow-up scores are reported for both treatment and control groups for primary outcome measure • All outcome measures to test effectiveness are clearly stated, giving both reliability and validity of the measurement in primary outcomes
Medium (2)	<p>To get this rating, the study must meet 3 out of 4 criteria</p> <ul style="list-style-type: none"> • There is a randomly assigned waitlist control group • Participants are blinded for assignment in experimental or control group through a stringent procedure • Pre- and Post- scores are reported for both treatment and control groups for primary outcome measure • All outcome measures to test effectiveness are clearly stated, giving reliability OR validity of the measurement in primary outcomes
Low (1)	<p>To get this rating, the study must meet 3 out of 4 criteria</p> <ul style="list-style-type: none"> • There is a control group without details stated • Participants are not randomly assigned to experimental or control group • Pre- and Post- scores are reported • The outcome measures to test effectiveness are reported

Rationale for the criteria used in WOE B

Criteria	Rationale
Use of control group	Since this systematic review is evaluating the effectiveness of a particular intervention, studies must include a control group. This is because control groups help to confirm that the study results are due to manipulation of independent variable/ intervention, rather than extraneous variables. Presence of control group allow comparison of the true effectiveness from the intervention group. An active control group is superior to a wait list control group.
Random assignment	Blinding of randomised control trial is considered high quality design to measure effectiveness of an intervention (Petticrew & Roberts, 2003)
Reporting results	Primary outcome measures from pre-, post- and multiple follow up points are reported, with statistical data.
Description of measures	The reliability and validity of outcome measures are clearly stated in the studies add high credibility to the research, which highlights the methodological design of what is trying to be measured is doing so reliably (e.g. Inter-rater reliability, Cronbach alpha- internal consistency).

Qualitative descriptors of WoE B ratings

Overall Quality	Average Score
High	≥ 2.5
Medium	1.5 – 2.4
Low	≤ 1.4

Weight of Evidence B

Study	Overall WoE B
1. Beaumont et al., 2021	3
2. Hopkins et al., 2011	2
3. Ben-Sasson et al., 2013	1
4. Rice et al., 2016	3
5. Beaumont and Sonfronoff, 2008	2

Appendix I: Weight of Evidence C (WoE C)

WoE C reviews the topic relevance from the research question. A review of question specific weighting, which evaluates the study based on the focus of the review (Gough, 2007)

Rationale for the criteria used in WOE C

Weighting	Criteria
High (3)	<ul style="list-style-type: none"> • The study clearly looks specifically on the effectiveness of intervention to bring positive change or improvement to the child’s social skills • Children met the DSM IV/V criteria for autism spectrum condition • The computer-based game intervention is specifically designed to improve children’s social skills • The study has a follow up phase that looks at the maintenance of intervention effects • Social skill is the primary focus and outcome measure of the study • Child behaviour measures are used to look at the impact of social skills on behaviour and parent views of social skills change is measured • A highly adapted, trained, and supportive computer-based guided intervention is used

To get this rating, the study must meet at least 6 out of 7 criteria

Medium (2)	<ul style="list-style-type: none"> • The study clearly looks specifically on the effectiveness of intervention to bring positive change or improvement to the child’s social skills • Children met the DSM IV/V criteria for autism spectrum condition • The computer-based game intervention is specifically designed to improve children’s social skills • The study has a follow up phase that looks at the maintenance of intervention effects • Social skill is the primary focus and outcome measure of the study • Child behaviour measures are used to look at the impact of social skills on behaviour and parent views of social skills change is measured • A highly adapted, trained, and supportive computer-based guided intervention is used
------------	--

To get this rating, the study must meet at least 4 out of 7 criteria

- Low (1)
- The study clearly looks specifically on the effectiveness of intervention to bring positive change or improvement to the child’s social skills
 - Children met the DSM IV/V criteria for autism spectrum condition
 - The computer-based game intervention is specifically designed to improve children’s social skills
 - The study has a follow up phase that looks at the maintenance of intervention effects
 - Social skills is the primary focus and outcome measure of the study
 - Child behaviour measures are used to look at the impact of social skills on behaviour and parent views of social skills change is measured
 - A highly adapted, trained, and supportive computer-based guided intervention is used

To get this rating, the study must meet at least 2 out of 7 criteria

- No evidence (0)
- No evidence is found in the study meets any of the criteria with regards to the focus for this systematic review.

Rationale for the criteria used in WOE C

Criteria	Rationale
Focus on study effectiveness	This is the main focus of the review that looks specifically at the effectiveness of intervention align with the review question.
Autism Diagnosis/ Condition	The review seeks to answer whether this intervention is effective for children who are clinically diagnosed with autism condition.
Design of computer-based game intervention	The aim of review reports the effect of computer-based game intervention with the focus and intention of improving social skills. Therefore, the most relevant studies’ intervention must target on improving behaviours of social skills.
Follow-up phase	The effectiveness on the duration from follow-up phase indicates whether the intervention maintains effect over time after the intervention has ceased.
Primary focus on social skills	The review is concerned with the effectiveness of improving children with autistic spectrum conditions’ social skills
Behaviour measures on social skills	The review clearly defines the behavioural outcome measures of social skills in children. The social skills

Intervention description	behavioural measures are valid, realisable and widely-used measure. A detailed description of the intervention assists future replication.
--------------------------	---

Weight of Evidence C

Study	Overall WoE C
1. Beaumont et al., 2021	3
2. Hopkins et al., 2011	3
3. Ben-Sasson et al., 2013	2
4. Rice et al., 2016	2
5. Beaumont and Sonfronoff, 2008	3

Qualitative descriptors of WoE C ratings

Overall Quality	Average Score
High	≥ 2.5
Medium	1.5 – 2.4
Low	≤ 1.4

Appendix J: Coding Protocols for Weight of Evidence A

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

Coding Protocol

Name of Coder: Anon. Date: 14 Jan, 2022

Full Study Reference in proper format: Beaumont, R., Walker, H., Weiss, J., & Sofronoff, K. (2021).

Randomized controlled trial of a video gaming-based social skills program for children on the autism spectrum. Journal of autism and developmental disorders, 1-14.

Intervention Name (description of study): 'Secret Agent Society' (SAS), computer game-based social skills program

Study ID Number: 01

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

1. General Characteristics

A. General Design Characteristics

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

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

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

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

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

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- N/A
- Unknown/unable to code

B. Participants

Total size of sample (start of study): 70

Intervention group sample size: 35

Control group sample size: 35

C. Type of Program

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

D. Stage of Program

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

E. Concurrent or Historical Intervention Exposure

- Current exposure
- Prior exposure
- Unknown

Author stated no prior exposure.

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

A. Measurement Criteria

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

- Yes
- No
- Unknown/unable to code

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

- Yes
 - 1. Social Skills Questionnaire (SSQ; Spence 1995)
 - 2. Emotion Regulation and Social Skills Questionnaire (ERSSQ; Beaumont and Sofronoff 2008)
 - 3. Spence Children’s Anxiety Scale- Parent (SCAS-P; Nauta et al. 2004)
 - 4. Eyberg Child Behavior Inventory- Parent (ECBI; Eyberg and Pincus 1999)
 - 5. Program Satisfaction Ratings
- No
- N/A
- Unknown/unable to code

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

- Yes
 - Primary Outcome:
 - 1. Parent-reported social skills (including emotion-regulation skills)
 - Secondary outcomes:
 - 1. Teacher-reported social skills (including emotion-regulation skills).
 - 2. Parent-reported anxiety.
 - 3. Parent-reported child behavior.
- No
- N/A
- Unknown/unable to code

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

- Yes validated with specific target group (not specifically for ASD)
 - Cronbach’s alphas:
 - 1. SSQ: .88 (parent) and .94 (teacher)
 - 2. ERSSQ: .95 (parent) and .94 (teacher)
 - 3. SCAS: .92 (parent)
 - 4. ECBI: .90 (parent)
- In part, validated for general population only
- No
- Unknown/unable to code

Overall Rating for measurement 3

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

B. Comparison Group

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

- Typical intervention (typical intervention for that setting, without additions that make up the intervention being evaluated)
- Attention placebo
- Intervention element placebo
- Alternative intervention
- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
- Unable to identify type of comparison

B2 Overall confidence of judgment on type of comparison group

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- Unable to identify comparison group

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

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

B4 Group equivalence established (select one of the following)

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

B5 Equivalent mortality

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

Findings _____ p. 3645 "additional analyses were conducted using the intention-to-treat principle via the 'ast observation carried forward' (LOCF) method"

Overall Rating for comparison group 2

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

C. Primary Outcomes are Statistically Significant

D1 Evidence of appropriate statistical analysis for **primary outcomes**

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

D2 Percentage of **primary outcomes** that are significant (select one of the following)

- Significant primary outcomes for at least 75% of the total primary outcome measures for each key construct
- Significant primary outcomes for between 50% and 74% of the total primary outcome measures for each key construct
- Significant primary outcomes for between 25% and 49% of the total primary outcome measures for any key construct.

Overall rating for Primary Outcomes 3

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

D. Implementation Fidelity

E1 Evidence of Acceptable Adherence

Eg. "Challenges to program fidelity (eg. lack of child engagement, time management) were problem-solved during the weekly parent coaching sessions"

- E1.1 Ongoing supervision/consultation
- E1.2 Coding intervention sessions/lessons or procedures
- E1.3 Audio/video tape implementation
 - E1.3.1 Entire intervention
 - E1.3.2 Part of intervention

E2. Manualization (select all that apply)

- E 2.1 Written material involving a detailed account of the exact procedure and the sequence they are to be used.
- E2.2 Formal training session that includes a detailed account of the exact procedures and the sequence in which they are to be used.
- E2.3 Written material involving an overview of broad principles and a description of the intervention phases. p. 3641, Table 3
- E2.4 Formal or informal training session involving an overview of broad principles and a description of the intervention phases.

- E3. Adaptation procedures are specified yes no unknown

eg. interventions to fit varying contexts, cultures, and participants while maintaining essential components p. 3649, Appendix 1

Overall rating for Implementation Fidelity: 3

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

E. Educational/ Clinical Significance

Outcome Variables:	Pre-test	Post-test	Follow up
--------------------	----------	-----------	-----------

<p>D2. Outcome Assessed via continuous variables</p> <p>social skills level (behaviour + cog)</p>		<p>Positive change in percentage of participants showing clinical improvement from pre to post test</p> <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown</p>	<p>Positive change in percentage of participants showing clinical improvement from posttest to follow up</p> <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown</p>
<p>D3. Subjective Evaluation: The importance of behaviour change is evaluated by parents in direct contact with the child</p>	<p>Importance of behaviour change is evaluated</p> <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown</p>	<p>Importance of behaviour change from pre to post-test is evaluated positively by the parents in direct contact with the target child</p> <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown</p>	<p>Importance of behaviour change from post-test to follow up is evaluated positively by the parents in direct contact with the target child</p> <p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Unknown</p>

<p>Overall rating for Educational/ Clinical Significance: <u>3</u></p> <p>3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence</p>
--

Summary of Evidence

Indicator	Overall evidence rating 0-3	Description of evidence Strong Promising Weak No/limited evidence Or Descriptive ratings
General Characteristics		
Design		Strong
Type of programme		Intervention: 'Secret Agent Society'
Stage of programme		Established program since 2010, across more than 9 countries
Concurrent/ historical intervention exposure		No. Children were excluded from the study if they were involved in another socio-emotional skills program.
Key features		
Measurement	3	Strong
Comparison group	2	Promising
Primary outcome statistically significant	3	Strong
Implementation fidelity	3	Strong
Educational/ Clinical significance	3	Strong