Research in Autism Spectrum Disorders 66 (2019) 101412

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: www.elsevier.com/locate/rasd

A systematic review and meta-analysis of social emotional computer based interventions for autistic individuals using the serious game framework

Julia S.Y. Tang[a,b,*], Nigel T.M. Chen[a,b,c], Marita Falkmer[a,b,d], Sven Bölte[a,b,e,f], Sonya Girdler[a,b]

[a] Curtin Autism Research Group, School of Occupational Therapy, Social Work and Speech Pathology, Curtin University, Perth, Western Australia, Australia
[b] Cooperative Research Centre for Living with Autism (Autism CRC), Long Pocket Brisbane, Queensland, Australia
[c] School of Psychological Science, University of Western Australia, Perth, Western Australia, Australia
[d] School of Education and Communication, CHILD programme, Institute of Disability Research, Jönköping University, Jönköping County, Sweden
[e] Division of Neuropsychiatry Unit, Department of Women's and Children's Health, Center of Neurodevelopment Disorders at Karolinska Intitutet (KIND), Stockholm, Sweden
[f] Child and Adolescent Psychiatry, Center for Psychiatry Research, Stockholm County Council, Sweden

ARTICLE INFO

Number of reviews completed is 2

Keywords:
Autism
Innovative technology
Social skills
Serious games

ABSTRACT

Background and aim: Adopting the elements of the Serious Game framework has been hypothesised as a strategy to promote the efficacy of social emotional computer-based interventions (CBI) for autistic individuals. This systematic review aimed to review the application of Serious Game principles in current social emotional CBI targeting autistic individuals and evaluate the effect of these principles in remediating social emotional outcomes via meta-analysis.

Methods: Database searches identified 34 studies evaluating social emotional CBI with 17 controlled efficacy studies included in meta-regressions analyses. Narrative synthesis summarised the attributes of each CBI based on the five Serious Game principles; motivating storyline, goal directed learning, rewards and feedback, increasing levels of difficulty and individualisation.

Results: Based on the scores of the Serious Game assessment tool we developed, findings revealed on average a limited (45%) integration of Serious Game design principles in social emotional CBI for autistic individuals. Main findings from the meta-regressions of 17 controlled efficacy studies revealed a moderating effect of Serious Game design principles on the distant generalisation of social emotional skills and transferability of outcomes among autistic individuals. No significant moderating effects of Serious Game was found for close generalisation and maintenance outcomes.

Conclusion: Overall, findings suggest that the Serious Game design framework has utility in guiding the development of social emotional CBI which improve the social emotional skills of autistic individuals.

* Corresponding author at: Curtin Autism Research Group, School of Occupational Therapy, Social Work and Speech Pathology, Curtin University, Kent Street, Bentley, Western Australia, Australia.
E-mail address: julia.tang@postgrad.curtin.edu.au (J.S.Y. Tang).

https://doi.org/10.1016/j.rasd.2019.101412
Received 20 March 2019; Received in revised form 14 June 2019; Accepted 17 June 2019
Available online 02 July 2019
1750-9467/

1. Introduction

Difficulties in social communication and interaction is a primary behavioural hallmark of Autism Spectrum Disorder (ASD) (American Psychiatric Association, 2013), contributing to poor social integration and participation outcomes in adulthood (Howlin, Moss, Savage, & Rutter, 2013). Autistic individuals experience lifelong difficulty in developing and maintaining relationships, often requiring assistance to live independently and maintain employment (Levy & Perry, 2011; Magiati, Tay, & Howlin, 2014; Tobin, Drager, & Richardson, 2014). The cumulative impact of ASD is evident in the poor self-reported quality of life (Clark, Magill-Evans, & Koning, 2015; Howlin et al., 2013; Jonsson et al., 2017) and premature mortality of these individuals (Hirvikoski et al., 2016).

Alterations in the social emotional skills of autistic individuals are underpinned by various social cognitive processes such as difficulties in mentalising, joint attention and emotion recognition (Baron-Cohen, 2000; Harms, Martin, & Wallace, 2010; Mathersul, McDonald, & Rushby, 2013; Uljarevic & Hamilton, 2013; Wong & Kasari, 2012). Difficulties integrating essential emotional cues are reportedly another feature characterising ASD (Guillon, Hadjikhani, Baduel, & Rogé, 2014), which become increasingly evident with increasing task demands (Harms et al., 2010). Current approaches aimed at remediating the social emotional skills of autistic individuals commonly occur in highly structured group environments lead by expert clinicians (Turner-Brown, Perry, Dichter, Bodfish, & Penn, 2008; White, Keonig, & Scahill, 2007). Social emotional group interventions are typically framed in line with the typical learning styles of autistic individuals (Baron-Cohen, 2006; Olsson et al., 2017; White et al., 2007), explicitly targeting individual skills through overt instruction, modelling, role-playing and continuous feedback (Reichow, Steiner, & Volkmar, 2012). While these teaching approaches have demonstrated preliminary effect in improving social emotional skills, further evidence is needed before clear conclusions can be drawn in regard to their efficacy (Choque Olsson et al., 2017; Gates, Kang, & Lerner, 2017; Seida et al., 2009; White et al., 2007). However, group approaches may not be universally appropriate for autistic individuals, as they are resource consuming, practically challenging and require participants to be in the same geographical location (Rao, Beidel, & Murray, 2008). Given this context, there is a need to consider additional modalities for teaching fundamental social emotional skills to autistic individuals.

Recent research exploring the utility of computer-based intervention (CBI) in remediating the social emotional skills of autistic individuals (Fletcher-Watson, 2014; Goodwin, 2008) suggests CBI delivered in home environments are both accessible and cost-effective (Goodwin, 2008). Autistic individuals often have a noted affinity for computer based activities (Mazurek, Shattuck, Wagner, & Cooper, 2012; Shane & Albert, 2008), which may provide an opportunity to learn and engage with complex social emotional skills in a safe, structured and predictable environment (Kapp, 2012).

CBI employ a variety of methods aiming to ameliorate the social emotional difficulties of autistic individuals (Fletcher-Watson, 2014), including embedding photographs and videos targeting social skills through repetition (Golan & Baron-Cohen, 2006; Silver & Oakes, 2001), combined with interactive feedback systems informing players of their progress, guiding them through the game (Ramdoss et al., 2012). Some CBI target social emotional skills via narratives ranging in complexity from simplistic themed based games (Faja et al., 2012) to intricate narratives reflecting real life scenarios (Beaumont & Sofronoff, 2008).

Recent years have seen an increase in research evaluating the effects of CBI for autistic individuals (Fletcher-Watson, 2014; Grossard et al., 2017; Grynszpan, Weiss, Perez-Diaz, & Gal, 2014; Ramdoss et al., 2012). At present, there has been one meta-analysis consisting of 14 controlled trials evaluating the overall efficacy of CBI for autistic individuals (Grynszpan et al., 2014). This meta-analysis, including nine studies focusing specifically on remediating social emotional skills, revealed promising evidence supporting CBI in improving skills of autistic individuals, reporting an overall medium effect size ($d = 0.47$). Given the proliferation of CBI in the recent years, an updated evaluation of the efficacy of CBI in remediating social emotional skills is timely. While a previous systematic review concluded that CBI showed some promise in remediating social emotional difficulties of autistic individuals (Ramdoss et al., 2012), these effects were primarily limited to improving 'close generalisation' skills, that is skills directly targeted within the CBI and assessed using a similar format to the intervention itself (Golan & Baron-Cohen, 2006; Swettenham, 1996). Conversely, the effects of CBI in improving the 'distant generalisation' of targeted skills, assessed via formats different to the intervention were more limited (Golan & Baron-Cohen, 2006; Ramdoss et al., 2012; Swettenham, 1996). To date, reviews have aggregated the results of close and distant generalisation social emotional outcomes (Grynszpan et al., 2014), failing to consider how interventions may differentially impact on these outcomes. Meta-analytic evaluation would be further beneficial in understanding the efficacy of social emotional CBI in improving the close and distant generalisation outcomes, particularly given that the ultimate goal of CBI is to transfer improvements in skills in everyday functioning (Ramdoss et al., 2012).

The ability of CBI to target social emotional skills largely depends on their capacity to promote player engagement (Catalano, Luccini, & Mortara, 2014). Current evaluation of these interventions have reported issues associated with attrition and compliance (Golan & Baron-Cohen, 2006; Heimann, Nelson, Tjus, & Gillberg, 1995), issues at least partially attributable to the inattention paid to strategies promoting player engagement (Goh, Ang, & Tan, 2008). Motivating tasks maximise the learning outcomes of CBI, being linked with higher levels of compliance (Habgood & Ainsworth, 2011) and social emotional gains (Beaumont & Sofronoff, 2008). A recent literature review of CBI in ASD suggested that personalisation, immediate feedback and realistic game features potentially increased intervention effectiveness (Fletcher-Watson, 2014). However, the authors acknowledged challenges in synthesising the literature resulting from heterogeneity in study design and limited reporting of design processes (Fletcher-Watson, 2014; Grossard et al., 2017). Further research is needed to identify the most salient design features of these interventions.

The Serious Game framework, previously applied in the fields of health and education has been proposed as relevant for developing CBI in ASD (Whyte, Smyth, & Scherf, 2015). Serious Game design frameworks guide the design phase of CBI, providing guidelines for their development and structure, facilitating the usability and playability of interventions for end-users (Rooney, 2012). There are several existing Serious Game frameworks within the field of ASD (Mitgutsch & Alvarado, 2012; Rooney, 2012;

Yusoff, 2010), drawing on pedagogical frameworks to facilitate the learning experience within these games, including the framework described by Yusoff (2010). Yusoff's (2010) framework was applied in a review evaluating the design quality of Serious Games for autistic individuals (Grossard et al., 2017), concluding there was wide variability in the application of Serious Game design principles, with few games emphasising the playability of CBI.

Whyte et al. (2015) articulated a Serious Game framework outlining the features likely to enhance the motivational aspects of CBI, with the aim of maximising learning opportunities in an engaging environment and generalising skills to everyday contexts (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Whyte et al., 2015). The principles of this framework consider the inclusion of immersive narratives, short and long term goals for targeted skills, providing meaningful feedback and rewards, personalisation and graded difficulty of tasks (Whyte et al., 2015). While the current framework articulates the design principles of Serious Games, clearer operational definitions of these principles would support a more consistent approach to CBI development and evaluation. A systematic approach to evaluating the application of the Serious Game design framework would also enable investigation of the relationship between application of these principles and the efficacy of CBI in ASD.

The overarching aim of the present review was to evaluate, via meta-analysis, the potential impact of Serious Game design in enhancing the effect of social emotional CBI in ASD. The present review firstly sought studies evaluating CBI targeting social emotional skills of autistic individuals, providing a narrative synthesis of intervention features, assessing them against the principles of the Serious Game framework (Whyte et al., 2015) with a tool developed specifically for this review. Secondly, this review aimed to retrieve controlled studies evaluating social emotional CBI, systematically examining the potential moderating influence of Serious Game design principles on the effect of these interventions in remediating social emotional difficulties of autistic individuals.

2. Methods

2.1. Eligibility criteria

The selection of the articles was conducted in two stages. The first stage aimed to locate studies evaluating social emotional CBI for autistic individuals. The inclusion and exclusion criteria for this stage are outlined in Table 1.

Subsequently, studies retrieved in stage one were evaluated to determine eligibility for a meta-analysis. Studies were included in the meta-analysis if they employed a control group based design involving autistic participants. As a strategy to increase homogeneity of CBI included in the meta-analysis, only those intervention focusing on social and behavioural approaches were included with interventions utilising neurophysiological methods of delivery such as neurofeedback training excluded.

2.2. Information sources and search strategy

Five electronic databases, CINAHL Plus (EBSCOhost), Embase (Ovid), ERIC (Proquest), Medline (Ovid), PsychInfo (Ovid) and PudMed were used to locate relevant articles. The main keywords search terms were "autism", "social", "emotion", and "computer", which were combined with relevant subject headings and Boolean operators, as presented in Table 2. Search strategies for each of the database are outlined in Appendix A. The search was limited to articles published from 1990 to 2018, replicating the search strategy of the previous review (Ramdoss et al., 2012). No language restrictions were placed on the searches. The search was conducted on

Table 1
Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Participants of studies	Participants diagnosed with ASD under the Diagnostic Statistical Manual of Mental Disorders 5 (DSM-5) or IV (DSM-IV) (American Psychiatric Association, 2013), International Classification of Diseases 10 (ICD-10) (World Health Organization, 1992) or the older versions. This inclusion accommodated for the adolescents and adults' population as individuals in this age group were mostly diagnosed under the earlier diagnostic classification. Therefore, studies targeting participants with Autistic Disorder, Asperger's Syndrome and Pervasive Developmental Disorder Not Otherwise Specified were included.	Individuals with Rett syndrome as the primary focus were excluded as the symptoms associated with this condition are related to a specific gene mutation.
Type of interventions	Studies evaluating the effect of education and therapy interventions targeting social emotional skills delivered via computerised devices as the primary platform. As per a previous review (Ramdoss et al., 2012), social emotional skills were defined as skills for initiating and maintaining social interactions, including several precursors skills to theory of mind, e.g. emotion recognition, joint attention, imitation and social reciprocity.	Excluded interventions with limited interaction between the user and the computerised platform, in that feedback was not adjusted based on user input. Therefore, programs with minimal user input such as DVD programs and video modelling were excluded. Studies involving advanced technology including virtual reality environments and robotics were excluded in order to focus on traditional screen-based computer software. Interventions targeting social skills for vocational training were excluded.
Types of studies	Articles were published in English, and were experimental in design, including randomised controlled trials, quasi experimental design, pre and post-test design and case studies.	Grey literature such as dissertations were excluded.

Table 2
Search strategy.

Search strategy	
1.	(autis* OR Asperger* OR ASD* OR "pervasive development disorder*" OR PDD)
2.	(Soci* OR "social conversation" OR "social interaction" OR "emotion*" OR "face*" or "facial*" OR feeling* OR "expression*" OR "turn taking" OR "eye contact" OR "joint attention" OR "empath*" OR "theory of mind")
3.	(Computer* OR software* OR technolog* OR "computer based" OR "computer assisted" OR "computer game*" OR "video game*" OR "serious game*" OR "simulation game*" OR "online game*" OR "game based learning")
4.	1 AND 2 AND 3
5.	Limit to 1990-2018

27th and 28th of July 2016 and updated on the 20th of October 2017 and 17th of December 2018.

2.3. Data collection and analysis

2.3.1. Study selection

Citations from the search results were imported to EndNote referencing manger where duplicates were eliminated. Screening at the title and abstract level was completed by the first author (JT), eliminating articles clearly irrelevant to the aim of this review. Studies primarily focusing on individuals without an ASD diagnosis and/or interventions targeting health behaviours, rather than social emotional skills were excluded. Full text articles were then retrieved and initially assessed by the first author (JT) for inclusion. For inter-rater reliability, 20% of the total search results were randomly selected and reviewed by two additional reviewers (MF and SG) based on the inclusion and exclusion criteria, achieving Cohen's kappa score of $r_{k} = .86–.88$, indicative of strong agreement between the reviewers (McHugh, 2012). Any disagreements were resolved via discussion.

2.3.2. Data extraction

According to the guidelines outlined in the Cochrane Handbook for Systematic Reviews (Higgins & Green, 2011), data extraction included; study design, description of participants including total number, mean age, gender and diagnostic information, intervention features relating to targeted skill domain, and outcome measures relating to Serious Game features and social emotional intervention outcomes. Information on social emotional CBI were primarily derived from the journal articles themselves. Where possible, a Google search was employed to retrieve further information from published sources of the game, freely downloadable game content, video demonstrations and author websites. Available information for each game is provided in Appendix B.

2.4. Summary measures

2.4.1. Serious game outcomes

The level to which the social emotional CBI included in this review applied the Serious Game principles was evaluated using a specifically develop Serious Game Assessment Tool informed by a review by Whyte et al. (2015), delineating five key principles for enhancing the motivation and learning experience of autistic individuals and facilitating the generalisation of skills. The first guiding principle describes narratives as a useful tool into contextualising and supporting immersive learning experiences, allowing for opportunities to interact with characters within the game. The second principle highlights the importance of structuring goals incrementally towards a long term objective. The third principle points to the importance of both feedback and reward systems in players obtaining information on their progress and collecting tangible rewards in sustaining their motivation for continued play. Incrementally increasing the level of difficulty and individualising the level of challenge represent the fourth and fifth Serious Game principles respectively, aiming to minimise player's frustration and ensuring that the game is appropriately pitched.

The Serious Game Assessment tool developed for this review included an ordinal classification system, quantifying the level to which each CBI applied the Serious Game principles. Full incorporation of a principle into the CBI was given the maximum score of two points. One point was given for partial incorporation and no points was given if the principle was absent within the game. Scores for each of the five game design principles were subsequently summed and converted to a percentage score relative to the maximum obtainable score of 10, with higher scores indicative of a greater incorporation of Serious Game principles. Scores of 49% and below, 50–59%, 60–69%, 70–79% and 80% and above, were taken to reflect limited, average, good, strong and very strong application of Serious Game principles, respectively. The final operational definitions and scoring matrix of the Serious Game Assessment Tool are presented in Table 3. Inter-rater reliability was assessed and achieved a kappa agreement score of $r_{k=}.96$. Further details on the development of the tool are available in Appendix C.

2.4.2. Social emotional outcomes

Continuous social emotional outcome measures were extracted, including the pre and post means and standard deviations for intervention and control groups for meta-analyses. Social emotional outcomes included assessments of social interactions or precursors of theory of mind skills such as eye contact, joint attention, emotion imitation and face and emotion recognition. These were then categorised into outcomes relating to close generalisation, distant generalisation, transferability to other skills and maintenance or engagement outcomes. Definitions for each outcome are outlined in Table 4.

Table 3
Serious game design evaluation- Scoring matrix.

Serious game principles	Criteria	Classification (points)	Scoring
Storyline	Storyline is focused on achieving a long term goal or an end purpose, is embedded throughout the entire game and the player is able to engage in purposeful interactions with game characters.	Yes (2)	
	The game has some storyline OR the player is able to purposefully interact with characters within the game. However, the game does not have a clear long term goal.	Some (1)	
	Game includes themed based content without a storyline.	Themed (1)	
	No storyline included in the game. Game focuses on explicit training of skills.	None (0)	
Goals	Game has a clear long term goal and achieved through short and medium term goals.	Long (2)	
	Game has a medium term goal that is achieved through completing short term goals/tasks but does not include a clear long term goal.	Medium (1)	
	Game includes smaller, targeted and isolated individual learning goals/ tasks.	Short (1)	
Rewards and feedback Note: Rewards are defined as features supporting immediate reinforcement. Feedback provides information on player's progress throughout the game.	Game provides a combination of reward and feedback features.	Both (2)	
	Game includes reward systems to provide immediate reinforcement such as visual and auditory stimuli or collectables.	Reward (1)	
	Game provides feedback on the player's performance to guide the player towards achieving learning goals.	Feedback (1)	
	Game does not include reward and/or feedback features.	None (0)	
Increasing levels of difficulty	Game occurs in a variety of context/stimuli to achieve learning goals and increases in level of difficulty (e.g. adding more distractors, amount of stimuli, increasing the speed) in a step-wise manner throughout the game and/or has individualised starting points.	Yes (2)	
	Game either uses different contexts/stimuli to train skills or increases in level of difficulty.	Some (1)	
	Game has a consistent level of difficulty throughout the program.	None (0)	
Individualised (individualising the level of difficulty for the player)	Individualisation is auto-generated by the program. The game automatically generates individualised starting points and adapts the level of difficulty based on the player's game performance.	Yes (2)	
	Individualisation is external. The player is autonomous in choosing the level of difficulty of game tasks.	Choice (1)	
	Individualisation is external. Level of difficulty is adapted based on the facilitator's judgment of the player's progress.	Facilitator (1)	
	No evidence of individualisation.	None (0)	
Total Serious Game Score (Maximum Score: 10)			

Table 4
Social emotional outcomes.

Social emotional outcome	Definition
Close generalisation	Outcomes evaluating the skills targeted by a given CBI under similar conditions to the intervention.
Distant generalisation	Measurements of targeted skills in a different assessment context.
Transferability	Social emotional skills which were not specifically targeted by the intervention itself.
Maintenance	The effect of the CBI to maintain social emotional skills after a certain period of time post-intervention, as measured via follow-up reports.
Engagement	Descriptions of the participant's motivation or interest during CBI, including measurements of attrition rates and qualitative reports.

2.5. Assessment of methodological quality and risk of bias

The methodological quality of the included studies was assessed according to the Standard Quality Assessment Criteria for Quantitative Studies developed by Kmet, Lee, and Cook (2011). The checklist contains 14 questions linked with a scoring system quantifying if the criteria was met (two points), partially met (one point) or not achieved (zero points). The strength of the methodological quality was ranked with > 80% summary scores were ranked as strong, 70–80% as good, 50–70% as adequate and < 50% as limited (Lee, Packer, Tang, & Girdler, 2008).

2.6. Meta-analysis procedures

Studies evaluating social emotional CBI with a controlled group comparison were retrieved for meta-analysis. Social emotional

outcomes were analysed with RStudio Version 1.0.143 (RStudio Team, 2016) and existing RStudio meta-analyses packages, metafor, compute.es and MAd (Del Re & Hoyt, 2014; Del Re, 2013; Viechtbauer, 2010). Estimates of effect sizes with a bias correction (Hedges g) was calculated by dividing the mean difference of pre-post assessment outcomes between intervention and control groups with the pooled pre-standard deviation (Morris, 2008). When means and standard deviations were not explicitly reported, F values and t values were extracted to compute Hedges g treatment effects of CBI from outcome measures in treatment and control groups (Borenstein, 2009).

Effect size data were analysed in four separate random effects meta-analyses to investigate the effect of social emotional CBI in improving close and distant generalisation, transferability and maintenance outcomes. Random effects model was selected to account for possible variance within and between studies (Borenstein, Hedges, Higgins, & Rothstein, 2010). Under each outcome of interest, effect sizes and variances within each individual study were aggregated to reduce the number of data entries (Borenstein, Hedges, Higgins, & Rothstein, 2009). Given the correlation between outcomes were not readily available, a correlation coefficient value of 0.5 was set for all studies (Borenstein et al., 2009). Subgroup analyses investigated the effect of CBI in supporting long term close and distant generalisation as well as transferability outcomes.

If there was sufficient range, i.e. the moderator values within a model were varied, meta-regressions were implemented with Serious Game scores for each individual study applied as moderator variable to determine the influence of Serious Game principles on the effect of social emotional CBI. Total intervention duration converted to minutes and mean age of participants were analysed as potential moderator variables. To obtain a more interpretable intercept value, the moderator variables were centred to the average moderator value across the outcomes included in the meta-regression (Del Re, 2015).

Assessment of heterogeneity was assessed using Chi^2 statistics. A value of $p < 0.05$ was applied to determine statistical significance (Higgins & Green, 2011). The degree of inconsistency was described with I^2 statistics, values of 25%, 50%, and 75% representing low, moderate and high heterogeneity, respectively (Higgins, Thompson, Deeks, & Altman, 2003). In order to assess publication bias, funnel plots followed by an Egger's test were undertaken to determine funnel plot asymmetry (Egger, Smith, Schneider, & Minder, 1997). Each funnel plot illustrates the observed effect size against the standard errors on the y-axis. Visual inspection of the funnel plots were completed for outcomes with a minimum of ten studies, due to the increased likelihood of Type I error in distinguishing real asymmetry for meta-analysis containing less than ten studies (Sterne et al., 2011). If evidence for publication bias emerged, secondary analysis using the trim and fill method outlined by Duval and Tweedie (2000) was completed.

3. Results

3.1. Study selection

Following removal of duplicates, a total of 5283 articles were identified through searches of electronic databases. A *priori* screening of the title and abstracts retrieved 114 articles for full text evaluation of eligibility. Overall, a total of 81 articles were excluded, resulting in 32 eligible articles. A manual screening of the references of eligible articles and the systematic review (Ramdoss et al., 2012) retrieved one article. A total of 33 articles were selected for the final inclusion in the review. One published article reported two separate trials (Golan & Baron-Cohen, 2006), resulting in a total of 34 experimental trials meeting the inclusion criteria of this review.

Amongst the 34 studies retrieved from the electronic search, 20 studies adopted a control group based design with autistic participants and were evaluated for meta-analysis inclusion. Two studies were excluded from the meta-analyses, one due to incomplete reporting of data (Bölte et al., 2002) and the other evaluated an intervention other than CBI, i.e. computer game with neurofeedback training (Friedrich et al., 2015). Following visual inspection of funnel plots, one study with a small sample size was removed from the meta-analysis due to the potential for publication bias resulting from an overestimation of effect size (Cheng, Luo, Lin, & Yang, 2018). This resulted in a total of 17 studies being eligible for inclusion in the meta-analyses. Fig. 1 presents a flow diagram of the process for selecting articles.

3.2. Study characteristics

In total, 868 participants were included in the 34 experimental trials. Of these 34 trials, 15 evaluated the CBI via a randomised controlled design (Beaumont & Sofronoff, 2008; Bölte et al., 2002, 2006; Cheng et al., 2018; Faja et al., 2012; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Friedrich et al., 2015; Hopkins et al., 2011; Lopata, Thomeer, Rodgers, Donnelly, & McDonald, 2016; Rice, Wall, Fogel, & Shic, 2015; Russo-Ponsaran, Evans-Smith, Johnson, Russo, & McKown, 2016; Silver & Oakes, 2001; Tanaka et al., 2010; Thomeer et al., 2015). Fifteen experimental studies employed a pre-test-post-test design (Bauminger-Zviely, Eden, Zancanaro, Weiss, & Gal, 2013; Bernardini, Porayska-Pomsta, & Smith, 2014; Bernard-Opitz, Sriram, & Nakhoda-Sapuan, 2001; Bölte et al., 2015; Faja, Aylward, Bernier, & Dawson, 2008; Gordon, Pierce, Bartlett, & Tanaka, 2014; Golan & Baron-Cohen, 2006; Jouen et al., 2017; LaCava, Golan, Baron-Cohen, & Myles, 2007; Malinverni et al., 2017; Serret et al., 2014; Swettenham, 1996; Thomeer et al., 2011; White et al., 2018), and four studies were described as case studies (Jeffries, Crosland, & Miltenberger, 2016; Lacava, Rankin, Mahlios, Cook, & Simpson, 2010; Miller, Wyatt, Casey, & Smith, 2018; Russo-Ponsaran, Evans-Smith, Johnson, & McKown, 2014).

A total of 20 experiments adopted a group controlled design with autistic individuals, with 13 control groups wait-listed or receiving standard care (Beaumont & Sofronoff, 2008; Bölte et al., 2002, 2006; Bölte et al., 2015; Faja et al., 2008; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Jouen et al., 2017; Russo-Ponsaran et al., 2016; Silver & Oakes,

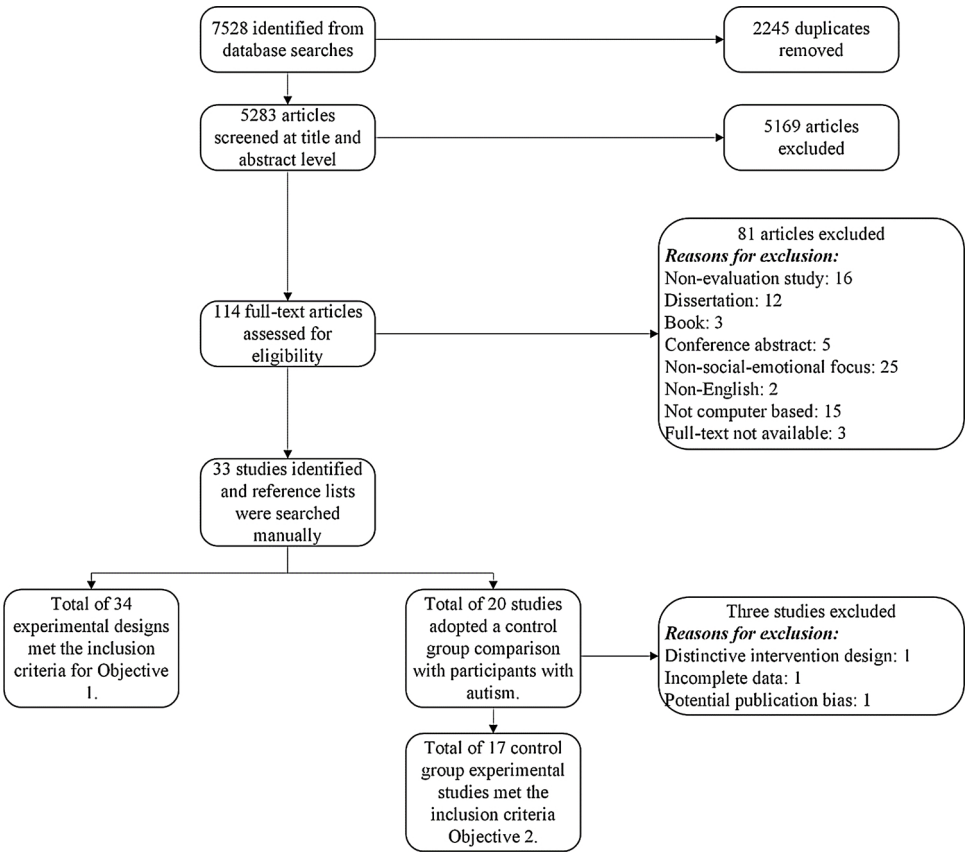

Fig. 1. Flow diagram for the selection of articles.

2001; Tanaka et al., 2010; Thomeer et al., 2015), three receiving face-to-face social skills training (Golan & Baron-Cohen, 2006; Lopata et al., 2016; Rice et al., 2015), two groups receiving an intervention without social emotional content (Faja et al., 2012; Hopkins et al., 2011), one study comparing the relative effect of bi-directional versus uni-directional neurofeedback training (Friedrich et al., 2015) and one study using a paper-based emotion recognition training as a control comparison (Cheng et al., 2018). The majority of designs exposed the intervention group to the CBI only, with the exception of three experimental studies which combined CBI with a group therapy program (Beaumont & Sofronoff, 2008; Golan & Baron-Cohen, 2006; Lopata et al., 2016). Details on study designs and group comparison are presented in Table 5.

3.3. Participant characteristics

3.3.1. Age and gender

Participants ranged in chronological age from pre-school aged children to middle-aged adults (ranged 3–52 years), with majority of participants in studies being school-aged children and a smaller percentage (k = 7; 21%) of studies targeting adults. Males comprised the majority of participants, making up 89% of the total sample across studies.

3.3.2. Sample size

Sample sizes ranged from 3 to 79 (Jeffries et al., 2016; Russo-Ponsaran et al., 2014; Tanaka et al., 2010). Studies with very small sample sizes typically adopted multiple single-case study designs (Jeffries et al., 2016; Lacava et al., 2010; Miller et al., 2018; Russo-Ponsaran et al., 2014). Larger sample sizes were recruited for studies employing randomised controlled trial designs, with the largest study including 42 participants in the CBI group (Tanaka et al., 2010).

3.3.3. Diagnosis

Overall, participants were mainly diagnosed with ASD without an additional intellectual or language impairment, Asperger's syndrome and Pervasive Developmental Disorder Otherwise Not Specified under the DSM-IV, while the remaining were classified under autism or ASD. Twenty studies confirmed diagnosis via validated diagnostic tools such as the Autism Diagnostic Observation Schedule- Generic (Lord et al., 2000), Autism Diagnostic Interview- Revised (Lord, Rutter, & Le Couteur, 1994) or the Childhood Autism Rating Scale (Schopler, Reichier, & Renner, 1986). Nine reported a priori confirmed diagnosis by a clinician supplemented

Table 5
Intervention characteristics.

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
Bauminger-Zviely et al. (2013)	Pre-post test	Social collaboration and conversation	Join-In: Participants engaged in themed cooperative activities and collaborated to discuss solutions. No-Problem: Participants were presented with short social conversation vignettes in different social environments.	Hardware: DiamondTouch device and laptop computer Setting: School environment	x12 45 minutes lessons	Guided by facilitators with special education or occupational therapy background. Participants paired with another peer.	NA	Adequate quality (Score = 17/28). Study was limited due to study design, small sample size, limited reporting of analytical methods. Participants differed in ages. Unclear selection strategy. Limited reporting of randomisation method and blinding protocol (blinding of evaluators and participants to intervention condition).
Beaumont and Sofronoff (2008)	RCT	Emotion recognition, emotion regulation and social interaction	Secret Agent Society: Game is themed in a futuristic setting. Participants play as a junior detective and undergo three level training program, Level 1: decoding clues from facial expressions, body posture and voice prosody, Level 2: self-regulation and detecting emotions based on situational clues, Level 3: completing missions based on various social conflicts.	Hardware: DVD based application with USB activation. Setting: University laboratory	x4 sessions for computer game (two one hour and two 45 minutes session).	Delivery: Guided by therapists. Sessions are conducted concurrently with parent training. Includes group therapy sessions and 'home missions' with parents. Activities include role-play games, positive reward system with 'tokens', group discussions and feedback Overall, x7 2 hour weekly sessions (time allocated between computer game, parent training and group sessions).	Waitlist	Strong quality (Score = 23/28). Randomisation method was not described. Limited reporting to account for performance bias (blinding of evaluators and participants to intervention condition).
Bernardini et al. (2014)	Pre-post test	Joint attention and symbol use (understanding gestures, words, use objects and non-verbal means to in social exchange)	ECHOES: Participants interacts with virtual character named 'Andy' and engages in different activities with 'magic' objects in the sensory garden.	Hardware: Multitouch LCD display with eye tracking. Setting: NA	10-20 minutes several times per week over a 6 week period.	Participants independently played with the interface with a practitioner supervising in the room (but out of the child's sight).	NA	Poor quality (Score = 9/28). Aim insufficiently described. Limited description of sampling strategy and results, has small sample size and non-randomisation protocol not reported. Limited

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
								reporting to account for performance bias (blinding of evaluators and participants). Heterogeneous study sample.
Bernard-Opitz et al. (2001)	Case-control	Social problem solving	Contains four easy and four difficult social vignettes of conflicts. Participants are prompted by the program to generate solutions to each social conflict.	Hardware: CD-ROM on Windows 95 PCs. Setting: NA	Ten sessions, duration not reported	NA	Typically developing peers	Poor quality (Score = 8/28). Limited in study design. Small sample size. Limited reporting to account for performance bias (blinding of evaluators and participants) and sampling strategy not reported. Participants differed in age and verbal ability. Limited description of results.
Bölte et al. (2002)	Pilot RCT	Emotion recognition	Frankfurt Test and Training: Photographs of adult faces and eyes displaying basic emotions. Contains textual information of the emotion and comic strips.	NA	x5 2 hours weekly sessions	NA	No intervention	Poor quality (Score = 8/28). Minimal description of sampling strategy, group characteristics, outcome measurement and results. Randomisation method was not described. Limited reporting to account for performance bias (blinding of evaluators and participants). Small sample size. Inadequate reporting of control of participants' baseline characteristics and results.
Bölte et al. (2006)	RCT	Emotion recognition	Frankfurt Test and Training: Same as Bolte et al. (2002).	NA	x5 2 hours weekly sessions	NA	No intervention	Adequate quality (Score = 16/28). Minimal description in inclusion and exclusion criteria. Participants' diagnostic and social skills functioning and matching of

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
								baseline characteristics were not reported. Limited reporting to account for performance bias (blinding of evaluators and participants). Small sample size.
Bölte et al. (2015)	Quasi	Emotion recognition	Frankfurt Test and Training: Same as Bolte et al. (2002)	NA	x8 1 hour weekly sessions	Assisted by a clinical psychologist.	ASD group received no intervention and typically developing peers.	Adequate quality (Score = 19/28). Randomisation protocol not described. Limited reporting to account for performance bias (blinding of evaluators and participants).
Cheng et al. (2018)	RCT	Emotion recognition	3D Complex Facial Expression Recognition: Contains 3D animated characters and social scenarios, displaying surprise, shy, nervous and embarrassed. Positive textual feedback provided for correct responses.	Hardware: Tablet (Android operating system). Setting: School or home.	x3 40 minutes sessions over 5 weeks.	Educators monitored participant's answers but provided minimal supervision.	Control group received paper-based emotion recognition training.	Adequate quality (Score = 18/28). Randomisation protocol not described. Limited reporting to account for performance bias (blinding of evaluators and participants). Unclear validity of outcome measurement.
Faja et al. (2008)	Quasi	Face processing	24 black and white pictures of male and female faces, either cropped (outer facial features removed) or filtered (distinct details removed). Explicit instruction of configural processing provided. Sessions focused on gender, age and identity matching. Reward image displayed for correct responses.	Hardware: Microsoft PowerPoint and E-Prime software, laptop computer with screen resolution of 1024 × 748 and image resolution of 72 pixels. Setting: NA	x8 30 minutes to 1 hour individual sessions over 3 weeks.	NA	No intervention	Adequate quality (Score = 14/28). Randomisation protocol not reported. Sampling strategy not described. Limited reporting to account for performance bias (blinding of evaluators and participants). Validity and reliability of selected outcome measurement not described. Small sample size. Minimal reporting of results.

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
			Presented face picture with the correct label for incorrect responses. Additional pictures are added after each session.					
Faja et al. (2012)	RCT	Face processing	Same as Faja et al. (2008)	Same as Faja et al. (2008)	x8 training sessions	NA	House training: Involved images of houses, structured in a similar format to the intervention.	Good quality (Score = 17/28). Randomisation method was not reported. Limited reporting of blinding protocol (participants and evaluators). Unclear validity and reliability of outcome measurements. Minimal reporting of statistical methods and results.
Fletcher-Watson et al. (2015)	RCT	Social communication	FindMe: Joint attention games involving finding a person in various outdoor environments (Part 1) or choosing objects in different shops based on eye or gestural cues (Part 2). Token system with a display of an animation after five tokens were collected. Difficulty increase through increasing the amount of distractors or from displaying gestural to eye cues.	Hardware: iPad application. Setting: Home setting with parents.	Varied. 2 months access to game was provided. Recommended usage was at least 5-10 minutes per day.	All participants received treatment as usual along with computer based intervention.	Waitlist	Strong quality (Score = 24/28). Reporting of participant blinded to intervention condition not described. One assessment was conducted by a blinded assessor. Differences in participant's verbal and non-verbal ability reported.
Friedrich et al. (2015)	RCT	Social interactions, imitation and emotional responsiveness	Social Mirroring Game: Participant plays as an avatar completing a treasure hunt mission with another game character.	Hardware: EEG set up recorded using Thought Technology Ltd. Bioamplifiers and proprietary software. Setting: NA	x16 1 hour sessions, 2-3 times per week over 6-10 weeks.	Neurofeedback training to increase mu rhythm during social and non-social episodes. Game includes EEG thresholds which subsequently increases as the	Neurofeedback training to decrease mu rhythm during social and non social episodes.	Adequate quality (Score = 18/28). Sampling strategy and randomisation protocol not reported. Limited reporting of blinding protocol. Small sample size. Minimal

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
			Includes social (imitates facial expressions) and non-social gaming tasks. Positive feedback involves avatar successfully imitating the expression of the other character (social tasks) and objects hitting the target (non-social tasks).			participant progresses in the game.		description in results.
Friedenson-Hayo et al. (2017)	RCT	Emotion recognition	Emotiplay: Participant plays the role of an explorer in a center researching human behaviour and emotion recognition. Game provided options to personalise avatar and collect items to purchase objects. Participants progress through educational materials from a professor and complete themed based games.	Hardware: Online computer game. Setting: Home	x8 weeks intervention period, at least 2 hours per week.	Parents were provided with an activity guide.	Waitlist	Good quality (Score = 21/28). Randomisation protocol not reported. Limited blinding of participants and evaluators to intervention condition. Insufficient description of the validity and reliability of close generalisation outcome measurements.
Golan and Baron-Cohen (2006)-1	Quasi	Emotion recognition	Mindreading: Contains three gaming areas, (1) Emotion library: A library of videos of facial expressions, voice recordings and situational written examples of the emotion. (2) Learning centre: Quizzes and lessons about emotions. Collectables used as rewards for correct responses. (3) Game zone.	Hardware: CD-ROM or DVD-ROM on an IBM compatible computer. Setting: Home	Total of 20 hours (2 hours/week over 10-14 weeks).	NA	No intervention ASD group and typically developing peers group	Good quality (Score = 21/28). Randomisation protocol not reported. Some (2/3) evaluators were blinded. Blinding of participants to treatment condition was not reported. High attrition rate.

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
Golan and Baron-Cohen (2006)-2	Quasi	Emotion recognition	Same as Golan & Baron-Cohen (2006) above	Same as Golan and Baron-Cohen (2006) above	x10 sessions (2 hour per week)	Participants attended additional group sessions facilitated by a tutor. Activities include analysis of emotions from real life examples, films or pictures. Group sessions were associated with the computer program.	Social skills group with explicit teaching, group discussions, role plays and analysing emotions from pictures.	Good quality (Score = 20/28). Randomisation not reported. Half of the evaluators were blinded (2/3). Blinding of participants to intervention group not reported. Small sample size.
Gordon et al. (2014)	Pre-post test	Emotion imitation	FaceMaze: Participants control a neutral face through a maze and mimics expressions to remove obstacles. Progress bars are filled after each correct responses and participants collects tokens along the maze.	Hardware: Computer Emotion Recognition Toolbox- real life facial expressions analysis. Setting: NA	NA	Real time feedback of emotion expressions based on video webcam analysis.	Typically developing peers	Adequate quality (Score = 15/28). Limited reporting of autism severity and blinding protocol. Small sample size. Heterogeneous study population.
Hopkins et al. (2011)	RCT	Joint attention, face and emotion recognition	FaceSay: Animated faces used in three themed games, 'Amazing Gazing', 'Band-Aid Clinic' and 'Follow the Leader'. Response options progressively increases. Player gains points for correct responses.	Hardware: Windows or Apple computers. Touch screen applications available. Setting: School or after school care	x12 10-25 minutes session, twice a week over 6 weeks.	One or two experimenters available to provide assistance. Positive reinforcement and reward was provided if participants demonstrated appropriate behaviour.	TuxPaint, computer based drawing software	Good quality (Score = 22/28). Randomisation method not described. Limited reporting of matching of participants' baseline characteristics. Small sample size. Limited reporting of blinding protocol (participants). Blinding of evaluators not reported, however achieved strong inter-rater reliability scores.
Jeffries et al. (2016)	Case studies	Joint attention (eye contact)	Look in My Eyes Steam Train: Participants play as a train engineer. In the training phase, coals are rewarded for correct identification of the number	Hardware: iPad application. Setting: Therapy centre.	NA	Therapist available during session but no prompts were provided.	NA	Poor quality (Score = 7/28). Study design without random allocation. Limited description of sampling strategy, participants' characteristics, analytical method and results.

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
			shown on the person's eyes. After four coals are collected, the reward phase is activated (train delivers package to the designated location).					Blinding of evaluators and participants not reported. Small sample size. Validity and reliability in outcome measurement not reported.
Jouen et al. (2017)	Quasi	Imitation and joint attention	GOLIAH: Contains 11 games targeting imitation and joint attention. Activities adapted from the Early Start Denver Model program. Levels of difficulty or goals are allocated by the therapist or parent. Smiley faces and auditory feedback are provided at the end of each game.	Hardware: Microsoft Visual Studio 10 Platform in C# language. Setting: Therapy centre and home.	x5 sessions per week (each session lasts approximately 30 minutes to 1 hour), over 6 months.	Therapist meet with child and parent once per week for one hour for review, planning and intervention (face to face and computer game).	Waitlist	Good quality (Score = 21/28). Limited in study design. Blinding of assessors and participants not reported. Small sample size.
LaCava et al. (2007)	Pre-post test	Emotion recognition	Mindreading: Details available in Golan and Baron-Cohen (2006).	Hardware: CD-ROM or DVD-ROM on an IBM compatible computer. Setting: Home or school.	x10 weeks	NA	NA	Adequate quality (Score = 15/28). Minimal description in results. Heterogeneous study population. Small sample size. Absence of control group and limited reporting of blinding protocol.
Lacava et al. (2010)	Case studies	Emotion recognition	Mindreading: Details available in Golan and Baron-Cohen (2006).	Hardware: CD-ROM or DVD-ROM on an IBM compatible computer. Setting: School	x7-10 weeks	NA	NA	Poor quality (Score = 10/28). Study design without random allocation. Insufficient reporting of participants' characteristics and results. Autism diagnosis confirmation not reported. Limited reporting to account for performance and detection bias. Small sample size. Heterogeneous study sample.

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
Lopata et al. (2016)	RCT	Emotion recognition	Mindreading: Details available in Golan and Baron-Cohen (2006).	Hardware: CD-ROM or DVD-ROM on an IBM compatible computer. Setting: Summer camp	x3 70 minutes treatment cycles per week over 5 weeks	Participants mainly played the game independently. Each week begins with an introduction of two new emotion groups and the third session focuses on a review of the emotions learnt. Includes a social skills group component with explicit teaching of other social skills.	SummerMAX social skills group (targeted social skills with role-plays, modelling, and feedback).	Strong quality (Score = 26/28). Limited reporting of blinding of participants to the intervention. Small sample size.
Malinverni et al. (2017)	Pre-post test	Social initiation	Pico's adventures: Participants engage in missions to initiate a friendship and go on adventures with a virtual character (Pico).	Hardware: Kinect based video game. Setting: Research facility.	One 45 minutes session	Participant played the game independently in the initial level. Involvement of parents and other children is introduced in the subsequent levels.	NA	Poor quality (Score = 7/28). Observational study design. Minimal description of sample characteristics.
Miller et al. (2017)	Case studies	Joint attention	Game contains static facial pictures accompanied by a voice recording prompt, 'Look at me'. Audio feedback was provided for eye contact made for five seconds.	Hardware: Computer monitor attached to an infrared camera, tracking the player's eye gaze. Setting: Research centre	Not specified	Participants were accompanied by a researcher who provided additional prompts or reinforcement when required.	NA	Poor quality (Score = 5/24). Weak study design, sample strategy and participant characteristics minimally described. Limited description of outcome measures and analytic methods. Small sample size.
Rice et al. (2015)	RCT	Joint attention and face processing	FaceSay: Details available in Hopkins et al. (2011).	Hardware: Windows or Apple computers. Touch screen applications available. Setting: School or after school care	x10 25 minutes session, once per week.	NA	SuccessMaker computer-based courses on language and analytical skills.	Good quality (Score = 21/28). Limited description of autism diagnosis, randomisation method and blinding protocol. Small sample size and participants' baseline characteristics not matched.

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
Russo-Ponsaran et al. (2014)	Case studies	Emotion recognition and self-expression	MiX: Contains videos of seven targeted emotions (happy, sad, angry, fear, surprised, contempt). The speed of the videos can be manually adjusted. Explicit instructions of each emotion is available in the program.	Hardware: Available online via internet browser. Setting: Research Centre.	X16 sessions over 8 weeks (1 hour sessions twice per week)	Facilitator available to provide explicit instructions and feedback of the facial features in each emotions and provide participants with prompt to imitate an emotion. Participants view their facial expressions via webcam.	NA	Poor quality (Score = 13/28). Study design without random allocation. Inadequate description of blinding protocol. Small sample size. Minimal reporting of statistical methods and control for possible confounding factors.
Russo-Ponsaran et al. (2016)	RCT	Emotion recognition and self-expression	MiX: Details available in Russo-Ponsaran et al. (2014).	Details available in Russo-Ponsaran et al. (2014).	Details available in Russo-Ponsaran et al. (2014).	Details available in Russo-Ponsaran et al. (2014).	Waitlist	Good quality (Score = 20/28). Independent evaluator was used for one outcome. Possible bias in randomisation method (coin flip), small sample size, incomplete reporting of results, age differences is a possible confounding factor and insufficient reporting of blinding protocol.
Serret et al. (2014)	Pre-post test	Emotion recognition	JeStiMuIE: Game has a learning phase and a training phase. The learning phase has three levels involving recognition of emotions from avatars based on face only, face with gestures and faces combined with both gestures and verbal content. Training phase is based in a virtual city, embedded with social scenarios with a task to recognise the emotion displayed in each scenario.	NA	Twice a week over 4 weeks (one hour per session)	Includes tactile stimulation on gamepad.	NA	Adequate quality (Score = 18/28). Study design did not include a controlled group. Insufficient reporting of blinding protocol. Heterogeneous study population. Unclear reliability in measurements as it is designed by investigators. Small sample size.

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
Silver and Oakes (2001)	RCT	Emotion recognition and theory of mind	Emotion Trainer: Includes five game tasks each containing photographs of facial expressions, an emotion triggering situation, or mental states tasks. Hints are provided for incorrect responses. Textual positive message and animation followed a correct response.	Hardware: NA Setting: School.	X10 daily 30 minutes sessions over 2-3 weeks	NA	Treatment as usual	Adequate quality (Score = 14/28). Sampling strategy insufficiently described. Autism diagnosis not described and confirmed. Minimal reporting of randomisation and blinding protocol. Small sample size. Unclear validity and reliability of outcome measurements. Insufficient reporting of results. Participants' baseline characteristic not matched.
Swettenham (1996)	Case control	Theory of mind	Computerised version of Sally-Anne false belief task. Program provides textual prompts, and immediate positive reinforcement using music and animation.	NA	Two sessions per day over four days. Each session consists of six trials.	NA	Typically developing peers and peers with Down Syndrome	Poor quality (Score = 8/28). Insufficient description of aim, sampling strategy, participants' autism severity, analytical methods and results. Non-random allocation. Limited reporting of blinding protocol. Small sample size.
Tanaka et al. (2010)	RCT	Face processing	Let's Face It!: Contains seven themed games. Participants have the option to select the mode and level of difficulty. They had the option to input their scores on the high score website.	Hardware: Program downloadable online and compatible with IBM computers. Setting: Home.	20 hours (instructed to play at least 100 minutes per week)	Parents are provided with tokens as an incentive to play the game.	Waitlist	Adequate quality (Score = 18/28). Minimal description in randomisation protocol and estimate of variance. Limited reporting of blinding protocol and results. Unclear validity and reliability in selected outcome measurement.
Thomeer et al. (2011)	Pre-post test	Emotion recognition	Mindreading: Details available in Golan and Baron-Cohen (2006).	Hardware: CD-ROM or DVD-ROM on an IBM compatible computer. Setting: University computer laboratory.	X12 sessions, two 90 minutes session per week over 6 weeks	Participants practised recognising and displaying emotions with a facilitator. Behaviour reinforcement system was used	NA	Adequate quality (Score = 18/28). Limited in study design as a controlled comparison group not available. Small sample size. Minimal

(*continued on next page*)

Table 5 (*continued*)

Author (Year)	Study design	Targeted domain	Intervention				Control group	Methodological quality
			Computer game content	Hardware/ Setting	Computer game duration/ Intensity	Additional components		
						outside of the computer program whereby participants earn points for appropriate social behaviours and correct identification/ expression of emotions.		description in analytical method.
Thomeer et al. (2015)	RCT	Emotion recognition	Mindreading: Details available in Golan and Baron-Cohen (2006).	Hardware: CD-ROM or DVD-ROM on an IBM compatible computer. Setting: University computer laboratory.	X24 sessions, two 90 minutes session per week over 12 weeks	Details available in Thomeer et al. (2011).	Waitlist	Strong quality (Score = 24/28). Limited reporting of blinding protocol.
White et al. (2018)	Case control	Emotion recognition	FEET: Contains four levels, consisting of cartoon faces, dynamic videos of real faces, audio recordings and avatars. Prompts are more subtle as they progress through the levels. Visual and audio feedback are provided for correct responses. Players are given the option to reattempt the question after an incorrect response.	Hardware: Kinect sensor attached to a computer. Setting: Research centre.	x1 session 60-90 minutes	Kinect senor capture player's facial expression and provide real time feedback.	NA	Adequate quality (Score = 12/24). Sample strategy not described. Limited description of analytical methods.

Note: ASD = Autism Spectrum Disorders; EEG = Electroencephalography; LCD = Liquid-crystal display; NA = Not available/applicable; PC = Personal computer; Quasi = Quasi-experimental study; RCT = Randomised Controlled Trial.

with confirmation of autism symptomatology via screening instruments, such as the Childhood Asperger Syndrome Test (Scott, Baron-Cohen, Bolton, & Brayne, 2002), Autism Behaviour Checklist (Krug, Arick, & Almond, 2008), Social Communication Questionnaire (Rutter, Bailey, & Lord, 2003), Social Responsiveness Scale (Constantino & Gruber, 2007) or the Autism Spectrum Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). In five of the included studies, procedures for verifying ASD diagnosis were not described (Jeffries et al., 2016; Lacava et al., 2010; Miller et al., 2018; Silver & Oakes, 2001; Swettenham, 1996).

3.3.4. Level of functioning

The majority of participants were reported to have a verbal and nonverbal intellectual quotient of 70 and above, with a subset of studies (k = 4) including individuals with an intellectual disability (Hopkins et al., 2011; Jouen et al., 2017; Miller et al., 2018; Serret et al., 2014). Table 6 presented the characteristics of participants in the included studies.

3.4. Intervention targets

As per the inclusion criteria of this review the intervention targets of studies were social skills relating to social emotional outcomes. Data synthesis categorised targeted skills according to social cognitive skills and social skills. Twenty nine studies focused on social cognition skills associated with theory of mind (Swettenham, 1996), face processing (Faja et al., 2008, 2012; Hopkins et al., 2011; Rice et al., 2015; Tanaka et al., 2010), emotion recognition (Beaumont & Sofronoff, 2008; Bölte et al., 2002, 2006; Bölte et al., 2015; Cheng et al., 2018; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Hopkins et al., 2011; LaCava et al., 2007; Lacava et al., 2010; Lopata et al., 2016; Russo-Ponsaran et al., 2016, 2014; Serret et al., 2014; Silver & Oakes, 2001; Thomeer et al., 2011, 2015; White et al., 2007) and joint attention (Bernardini et al., 2014; Fletcher-Watson et al., 2015; Hopkins et al., 2011; Jeffries et al., 2016; Jouen et al., 2017; Malinverni et al., 2017; Miller et al., 2018; Rice et al., 2015). Five experimental studies examined social skills involving collaboration with a partner to generate solutions (Bauminger-Zviely et al., 2013; Malinverni et al., 2017), emotion management (Beaumont & Sofronoff, 2008), problem-solving social conflicts (Bernard-Opitz et al., 2001), imitating and responding to emotions (Friedrich et al., 2015; Gordon et al., 2014), social conversations (Bauminger-Zviely et al., 2013), and social initiation skills (Malinverni et al., 2017).

Across the 34 included articles, a total of 24 different social emotional CBI were evaluated. Eight computer programs were commercially available for purchase; Emotion Trainer (Silver & Oakes, 2001), FaceSay (Hopkins et al., 2011), FindMe (Fletcher-Watson et al., 2015), Let's Face It! (Tanaka et al., 2010), Look in eyes: Steam Train (Jeffries et al., 2016), Mindreading (Golan & Baron-Cohen, 2006), MiX (Russo-Ponsaran et al., 2016), and Secret Agent Society (formerly known as the Junior Detective Training Program) (Beaumont & Sofronoff, 2008). Six computer program, ECHOES (Bernardini et al., 2014), Gaming Open Library for Intervention in Autism at Home (GOLIAH) (Jouen et al., 2017), Join In and No Problem (Bauminger-Zviely et al., 2013), JeStiMule (Serret et al., 2014) and Pico's Adventures (Malinverni et al., 2017) were at the piloting and/or development phase. Several computer programs were integrated with automated detection software enabling participants to manipulate the game using facial movements (Gordon et al., 2014; White et al., 2018), eye movements (Miller et al., 2018) or brain activity (Friedrich et al., 2015). One program integrated the player's image into the gaming environment through Kinect technology (Malinverni et al., 2017). Other programs presented social emotional stimuli such as animated characters (Cheng et al., 2018), photographs (Bölte et al., 2002; Faja et al., 2008) or a series of social vignettes (Bernard-Opitz et al., 2001; Swettenham, 1996). Additional details on the computer based interventions are outlined in Table 5.

3.5. Social emotional assessments

Raw data from social emotional outcome measurements is available in Appendix D.

3.5.1. Close generalisation

A total of 16 studies evaluated close generalisation outcomes using multiple assessment measures relating to social cognitive and social skills outcomes. Social cognitive outcomes targeting face and affect processing were measured in response to a variety of stimuli, including static images, dynamic videos of real life faces or voice recordings in a similar format to the intervention (Beaumont & Sofronoff, 2008; Bölte et al., 2006, 2015; Cheng et al., 2018; Faja et al., 2008, 2012; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Hopkins et al., 2011; Lopata et al., 2016; Rice et al., 2015; Russo-Ponsaran et al., 2016; Silver & Oakes, 2001; Tanaka et al., 2010; Thomeer et al., 2015). Emotion imitation skills were measured in one study via ratings of the emotion expressed by the participant (Russo-Ponsaran et al., 2016). One study evaluated emotion management using an interviewer administered questionnaire eliciting participants' response to a social scenario (Beaumont & Sofronoff, 2008).

3.5.2. Distant generalisation

Amongst the 17 included controlled trials, eight studies reported pre and post distant generalisation outcomes (Beaumont & Sofronoff, 2008; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Lopata et al., 2016; Russo-Ponsaran et al., 2016; Thomeer et al., 2015), assessing emotion recognition, emotion awareness, emotion regulation and social communication skills. Distant generalisation assessments were mainly computer-based (Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Russo-Ponsaran et al., 2016) or researcher administered assessments (Lopata et al., 2016; Russo-Ponsaran et al., 2016) scoring participants' accuracy in identifying or interpreting static images, audio recordings or naturalistic social interaction videos. Four studies measured distant generalisation via parental reports (Beaumont & Sofronoff, 2008; Fletcher-Watson et al., 2015; Lopata

Table 6
Participants' characteristics.

Author (Year)	N (male/female)	Mean age* (SD)		Diagnosis	Comorbidities	Inclusion and exclusion criteria
		Intervention	Autism control group			
Bauminger-Zviely et al. (2013)	22 (18/4)	9.83 (10.72)	NA	Autistic disorders and ASD, DSM-IV-TR, ADOS & SCQ	NA	Inclusion: WISC-IV performance IQ and PPVT verbal IQ ≥ 70 Exclusion: Behaviour difficulties and/or ADHD
Beaumont and Sofronoff (2008)	49 (44/5)	9.64 (1.21)	9.81 (1.26)	AS, DSM-IV-TR, CAST	NA	Inclusion: WISC-III IQ ≥ 85
Bernardini et al. (2014)	19	8.42	NA	ASD, AS, SCQ	NA	Inclusion criteria: Verbal language ability- BPVS M = 36.62. 7/9 participants scored low on BPVS.
Bernard-Opitz et al. (2001)	16	7.1	NA	Autism, ABC	NA	Inclusion: Autism score > 65 and IQ in the normal range (Cognitive functioning-KBIT and verbal IQ- BPVS)
Bölte et al. (2002)	10 (10/0)	27.2 (7.0)	NA	HFA, AS, ADI-R, ADOS-G	NA	Inclusion: IQ (Mean- 104.2, SD-17.1, range-58-126)
Bölte et al. (2006)	10 (10/0)	29.4 (5.9)	25.8 (8.0)	Idiopathic autism, ICD-10	NA	Inclusion: non verbal IQ (M = 104.2, SD = 17.1, range-58-126)
Bölte et al. (2015)	32 (30/2)	19.3	NA	Autism, AS, Atypical autism/PDD-NOS, ICD-10, ADI-R, ADOS		Inclusion: RSPM IQ ≥ 79, PBVT ≥ 8 Exclusion: Neurological disorders, mental disorders and genetic syndromes.
Cheng et al. (2018)	24 (16/8)	11.3	10.9	ASD, DSM-5	NA	Inclusion: WASI-IV > 70
Faja et al. (2008)	10 (10/0)	19.6 (7.37)	19.8 (3.42)	Autistic Disorder, AS, PDD-NOS, DSM-IV, ADI-R, ADOS	NA	Inclusion: WISC-III or WAIS-III IQ ≥ 85
Faja et al. (2012)	18	22.4 (4.4)	21.5 (5.6)	HFA, ADI-R, DSM-IV	NA	Inclusion: Impaired face recognition according to the Wechsler Memory Scale (WAIS IQ) and Benton Test of Facial Recognition.
Fletcher-Watson et al. (2015)	54 (43/11)	49.3 (10.9) months	49.96 (13.2) months	ASD, ADOS	Down syndrome (n = 2)	Inclusion: English speaking Exclusion: Neurological disorders
Friedenson-Hayo et al. (2017)-1	15 (11/4)	8.52 (1.11)	NA	ASD, ADOS-2, DSM-IV, ICD-10	NA	Inclusion: Met the ADOS-2 cut-off criteria for ASD
Friedenson-Hayo et al. (2017)-2	74 (66/8)	Israel: 7.68 (1.2), Sweden: 6.95 (0.96)	Israel: 7.28 (1.3), Sweden: 7.24 (0.99)	ASD, ADOS-2, DSM-IV, ICD-10	NA	Inclusion: Met the ADOS-2 cut-off criteria for ASD
Friedrich et al. (2015)	13 (12/1)	11.1 (2.3)	11.8 (3.9)	ASD, ADOS, ADI-R	NA	Inclusion: WASI, WISC, M = 100, SD = 15
Golan and Baron-Cohen (2006)-1	41 (31/10)	30.5 (10.3),	30.9 (11.2)	HFA, AS, DSM-IV, ICD-10, AQ	Depression, ADHD (n = 10)	Inclusion: WASI ≥ 70
Golan and Baron-Cohen (2006)-2	26 (22/4)	25.5 (9.3)	24.4 (6.4)	HFA, AS, DSM-IV, ICD-10, AQ	NA	Inclusion: WASI ≥ 70
Gordon et al. (2014)	17	10.76 (3.59),	NA	ASD, ADOS, ADI	NA	Inclusion: KBIT-2 > 100
Hopkins et al. (2011)	49 (44/5)	LFA training: 10.31 (3.31), HFA training: 10.05 (2.30)	LFA control: 10.57 (3.2), HFA control: 9.85 (2.87)	LFA, HFA, DSM-IV, CARS	NA	Inclusion: KBIT-II > 70 for HFA group and KBIT-II < 70 in LFA group
Jeffries et al. (2016)	3 (3/0)	Case 1: 3.67, Case 2: 3.58, Case 3: 5.92	NA	Autism	NA	NA
Jouen et al. (2017)	24 (24/0)	6.85 (1.34)	7.17 (1.62)	Autism, ASD, AS, ADI-R	NA	Inclusion: IQ ≥ 60, aged between 5-8 years, parents' motivation to complete study protocol. Participants matched for gender, age, IQ and study sites. Exclusion: Organic syndrome and/or non-stabilised neuropediatric (e.g. seizures) or medical (e.g. diabetes) comorbidities
LaCava et al. (2007)	8 (6/2)	10.27 (1.24)	NA	AS, DSM-IV. ASDS, ASQ	NA	NA
Lacava et al. (2010)	4 (4/0)	8.6 (0.8)	NA	Autism, PDD-NOS	NA	NA

(*continued on next page*)

Table 6 (*continued*)

Author (Year)	N (male/ female)	Mean age* (SD)		Diagnosis	Comorbidities	Inclusion and exclusion criteria
		Intervention	Autism control group			
Lopata et al. (2016)	36 (34/2)	8.83 (1.47)	8.83 (1.5)	Autism, AS, PDD-NOS, ADI-R	NA	Inclusion: WISC-IV > 70 (Verbal and Perceptual Reasoning of ≥ 80). Exclusion: Behavioural difficulties
Malinverni et al. (2017)	10 (10/0)	5.3 (0.94)	NA	ASD, DSM-IV, ADOS, ADI-R	NA	Inclusion: WISC > 70 Exclusion: Participation in other psychological treatment and presence of changes in medication during the trial.
Miller et al. (2017)	3 (3/0)	Case 1: 9 Case 2: 10 Case 3: 10	NA	ASD	Limited verbal abilities and intellectual disability.	Inclusion: Able to understand the instruction 'look at me' and have an interest in computers.
Rice et al. (2015)	31 (28/3)	7.68 (1.45)	7.87 (1.6)	ASD	NA	Inclusion: FSIQ > 70
Russo-Ponsaran et al. (2014)	3	NA, Range: 8-14 years	NA	AS, PDD-NOS, ASD, SCQ, ADOS, ADI-R	NA	Inclusion: WASI IQ ≥ 80, participants with difficulties in recognising emotions- scoring at or below 80% on a MiX pretest.
Russo-Ponsaran et al. (2016)	31 (26/5)	10.6 (1.7)	12.4 (1.9)	AS, PDD-NOS, ASD, SCQ, ADOS, ADI-R	NA	Inclusion: WASI IQ ≥ 80, SCQ ≥ 12, difficulties in emotion recognition
Serret et al. (2014)	33 (31/2)	11.4 (3.16)	NA	LFA, HFA, AS, PDD-NOS, DSM-IV-TR, ADI-R, ADOS	NA	Inclusion: WASI range 35-129, able to discriminate colours and use a computer
Silver and Oakes (2001)	22	13.92 (0.92)	14.75 (2.00)	Autism, AS	NA	Inclusion: Age equivalent of 7 years on BPVS.
Swettenham (1996)	8	10.9 (3.8)	NA	Autism, DSM-III-R	NA	Inclusion: Matched verbal ability using the BPVS and non-verbal ability using the LIPS
Tanaka et al. (2010)	79 (62/17)	10.5 (3.8)	11.4 (3.7)	Autistic disorder, AS, PDD-NOS, DSM-IV, ADI-R, ADOS-G	NA	Exclusion: Vision less than 20-100 in both eyes, difficulty in comprehending instructions Inclusion: Face processing abilities above 2 or more SD below age expected range.
Thomeer et al. (2011)	11 (8/3)	9.09 (1.76)	NA	Autism, AS, PDD-NOS, ADI-R	NA	Inclusion: WISC-IV IQ ≥ 70, receptive or expressive language score ≥ 80 on CASL.
Thomeer et al. (2015)	43	8.86 (1.39)	8.57 (1.16)	Autism, AS, PDD-NOS, ADI-R	NA	Inclusion: WISC-IV IQ ≥ 70, receptive or expressive language score ≥ 80 on CASL. Exclusion: psychiatric symptoms
White et al. (2018)	20 (18/2)	122.50 months	NA	ASD, ADOS-2	NA	Inclusion: Age between 9 and 12 years, free from any co-occurring intellectual disability. WASI-II was administered.

Note. ABC = Autism Behaviour Checklist; ADHD = Attention Deficit Hyperactivity Disorder; ADI-R = Autism Diagnostic Interview- Revised; ADOS = Autism Diagnostic Observation Schedule; ADOS-G = Autism Diagnostic Observation Schedule- Generic; AS = Asperger's syndrome; ASD = Autism Spectrum Disorders; ASQ = Autism Spectrum Quotient; ASDS = Asperger Syndrome Diagnostic Scale; BPVS = British Picture Vocabulary Scale; CARS = Childhood Autism Rating Scale; CASL = Comprehensive Assessment of Spoken Language; CAST = Childhood Asperger Syndrome Test; DSM = Diagnostic and Statistical Manual of Mental Disorders; FSIQ = Full Scale Intelligence Quotient; HFA = High functioning autism; ICD-10 = International Classification of Diseases 10th edition; IQ = Intelligence quotient; KBIT = Kaufman Brief Intelligence Test; LFA = Low functioning autism; LIPS = Leiter International Performance Scale; M = Mean; N = Total number; NA = Not available/ applicable; PDD-NOS = Pervasive Developmental Disorder- Not Otherwise Specified; PPVT = Peabody Picture Vocabulary Test; RSPM = Raven's Standard Progressive Matrices; SCQ = Social Communication Questionnaire; SD = Standard deviation; WAIS = Wechsler Adult Intelligence Scale; WASI = Wechsler Abbreviated Scale of Intelligence; WISC = Wechsler Intelligence Scale for Children.

et al., 2016; Thomeer et al., 2015).

3.5.3. Transferability

Measurements of transferability outcomes, or performance in areas not targeted by the CBI included social skills based on parental or teacher reports or naturalistic observations of social interactions (Beaumont & Sofronoff, 2008; Hopkins et al., 2011; Jouen et al., 2017; Lopata et al., 2016; Rice et al., 2015; Russo-Ponsaran et al., 2014; Thomeer et al., 2015), ASD symptomatology (Jouen et al., 2017; Lopata et al., 2016; Rice et al., 2015; Thomeer et al., 2015), brain activation using functional magnetic resonance imaging or electroencephalography (Bölte et al., 2006, 2015; Faja et al., 2012), language skills (Fletcher-Watson et al., 2015), emotion awareness (Russo-Ponsaran et al., 2014) and theory of mind (Rice et al., 2015; Silver & Oakes, 2001).

3.5.4. Maintenance

Four studies conducted a follow-up session ranging from four weeks to six months following the intervention to evaluate the maintenance of skills post intervention (Beaumont & Sofronoff, 2008; Fletcher-Watson et al., 2015; Russo-Ponsaran et al., 2016; Thomeer et al., 2015).

3.5.5. Engagement

Narrative synthesis of user engagement during CBI was reported in 12 experimental studies. Participants' level of engagement was captured via feedback from participants or parents and by examining attrition rates (Bernardini et al., 2014; Cheng et al., 2018; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Jouen et al., 2017; LaCava et al., 2007; Lacava et al., 2010; Silver & Oakes, 2001; Thomeer et al., 2011, 2015).

3.6. Quality assessment of studies

Assessment of the methodological quality of the included studies, evaluated via the Standard Quality Assessment of Quantitative studies by Kmet et al. (2011) revealed few studies achieving a strong (12%, k = 4) (Beaumont & Sofronoff, 2008; Fletcher-Watson et al., 2015; Lopata et al., 2016; Thomeer et al., 2015) or good (24%, k = 8) methodological quality (Faja et al., 2012; Fridenson-Hayo et al., 2017; Golan et al., 2010; Hopkins et al., 2011; Jouen et al., 2017; Rice et al., 2015; Russo-Ponsaran et al., 2016). The majority achieved adequate methodological quality (38%, k = 13) (Bauminger-Zviely et al., 2013; Bölte et al., 2006, 2015; Faja et al., 2008; Friedrich et al., 2015; Gordon et al., 2014; LaCava et al., 2007; Serret et al., 2014; Silver & Oakes, 2001; Tanaka et al., 2010; Thomeer et al., 2011) and 26% (k = 9) of studies were judged to be of poor quality (Bernardini et al., 2014; Bernard-Opitz et al., 2001; Bölte et al., 2002; Lacava et al., 2010; Russo-Ponsaran et al., 2014; Swettenham, 1996).

Study quality was mainly limited due to high performance and detection bias, with blinding of participants and assessors rarely evaluated. Although 15 studies employed a randomised controlled trial design (Beaumont & Sofronoff, 2008; Bölte et al., 2002, 2006; Cheng et al., 2018; Faja et al., 2012; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Friedrich et al., 2015; Hopkins et al., 2011; Lopata et al., 2016; Rice et al., 2015; Russo-Ponsaran et al., 2016; Silver & Oakes, 2001; Tanaka et al., 2010; Thomeer et al., 2015), there was often limited reporting of randomisation protocols. Other factors contributing to low methodological quality included small sample sizes and heterogeneity between groups in baseline characteristics. Table 5 and Appendix E outline the methodological scoring for all studies.

3.7. Serious game outcomes

The 34 included studies described 24 social emotional CBI designed for use with autistic individuals. Five computer programs were evaluated in more than one study; the Frankfurt Test and Training of Facial Affect Recognition (Bölte et al., 2002, 2006; Bölte et al., 2015), MiX (Russo-Ponsaran et al., 2014, 2016), FaceSay (Hopkins et al., 2011; Rice et al., 2015), Mindreading (Golan & Baron-Cohen, 2006; LaCava et al., 2007; Lacava et al., 2010; Lopata et al., 2016; Thomeer et al., 2011, 2015) and a face processing program developed by Faja et al. (2012) and Faja et al. (2008). Bauminger-Zviely et al. (2013) evaluated two computer programs, No Problem and Join In, with findings presented separately given the differences in game features and targeted skills.

3.7.1. Storyline

Three computer programs integrated learning goals within a storyline running throughout the game (Beaumont & Sofronoff, 2008; Jeffries et al., 2016; Malinverni et al., 2017), with the remaining CBI having limited or no narrative components. The *'Secret Agent Society'* is set in the future with players informed they have been selected to undergo training as a secret agent completing graded training missions focusing on detecting the emotions and thoughts of potential suspects (Beaumont & Sofronoff, 2008). Towards the end of the program, players are promoted to an official 'Secret Agent'. The *'Look in My Eyes: Steam Train'* game assigns players the role of a train engineer, responsible for delivering a package to an assigned destination (Jeffries et al., 2016), achieved by completing a series of joint attention tasks and rewarded by collecting enough coal to power the delivery train. In *Pico's Adventure'* game, the player is required to initiate a friendship and cooperate with an alien character named Pico in a series of adventure tasks including repairing a spaceship and travelling to another planet (Malinverni et al., 2017).

While five games incorporated narrative aspects focused on players interacting reciprocally with game characters, their narratives were not integrated with learning objectives (Bernardini et al., 2014; Fridenson-Hayo et al., 2017; Friedrich et al., 2015; Serret et al., 2014; Swettenham, 1996). Within these games, characters typically adopt the role of a peer mentor engaging in dialogue with the

goal of guiding players through the game. Several games incorporated artificial intelligence technology enabling game characters to adapt to player behaviour (Bernardini et al., 2014; Friedrich et al., 2015).

Five programs adopted a themed approach without narrative content (Bauminger-Zviely et al., 2013; Fletcher-Watson et al., 2015; Gordon et al., 2014; Hopkins et al., 2011; Tanaka et al., 2010). In some games, various themes within the game associated with individual sub-skills (Bauminger-Zviely et al., 2013; Hopkins et al., 2011; Tanaka et al., 2010). For example, *Let's Face It!* targets face recognition skills via a shooting theme called *'Zap It'* and joint attention goals in the another game called *'Eye Spy'* (Tanaka et al., 2010). Other games such as the *FaceMaze* and *FindMe* employed a consistent theme throughout the game (Fletcher-Watson et al., 2015; Gordon et al., 2014). In *FindMe*, players were required to find a person or the character's desired object hidden within different locations (Fletcher-Watson et al., 2015).

The 11 remaining computer programs focused on targeted goals via a question and answer format, and did not employ a narrative approach to target learning goals (Bauminger-Zviely et al., 2013; Bernard-Opitz et al., 2001; Bölte et al., 2002, 2006; Bölte et al., 2015; Cheng et al., 2018; Faja et al., 2008, 2012; Golan & Baron-Cohen, 2006; Jouen et al., 2017; LaCava et al., 2007; Lacava et al., 2010; Lopata et al., 2016; Miller et al., 2018; Russo-Ponsaran et al., 2016, 2014; Silver & Oakes, 2001; Thomeer et al., 2011, 2015; White et al., 2018).

3.7.2. Goal-Directed learning

Only four programs incorporated long term objectives (Beaumont & Sofronoff, 2008; Jeffries et al., 2016; Malinverni et al., 2017; Serret et al., 2014). *'Secret Agent Society'* and *'Look in My Eyes: Steam Train'* required players to graduate as a 'Secret Agent' or deliver a package, respectively (Beaumont & Sofronoff, 2008; Jeffries et al., 2016). The final aim of *'Pico's Adventures'* was to assist the alien character safely back to its planet (Malinverni et al., 2017). Although *JeStiMule* did not include a narrative component, the long term objective of the game was to collect puzzle pieces with the goal of completing the 30-piece puzzle by the end of the game (Serret et al., 2014).

Eight CBI were classified as having medium term goals, requiring players to reach a threshold of performance before progressing to the next stage of the game (Bauminger-Zviely et al., 2013; Bernardini et al., 2014; Fletcher-Watson et al., 2015; Friedrich et al., 2015; Hopkins et al., 2011; Jouen et al., 2017; Rice et al., 2015; Tanaka et al., 2010). Some games employed a point system to implement medium term goals, with players' performance rewarded with a set number of points allowing them to receive an award or progress within the game (Fletcher-Watson et al., 2015; Friedrich et al., 2015; Hopkins et al., 2011; Rice et al., 2015; Tanaka et al., 2010). One game required achievement of prerequisite learning objectives before allowing players to progress to the next objective (Bauminger-Zviely et al., 2013). In *ECHOES,* the multi-sensory garden was divided into zones representing various learning activities, with some requiring the players to complete a series of tasks before achieving the end-goal of the activity (Bernardini et al., 2014).

Twelve CBI incorporated short term learning objectives largely delivered via repeated presentation of differing social stimuli (Bernard-Opitz et al., 2001; Bölte et al., 2002, 2006; Bölte et al., 2015; Cheng et al., 2018; Faja et al., 2008, 2012; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Gordon et al., 2014; LaCava et al., 2007; Lacava et al., 2010; Lopata et al., 2016; Miller et al., 2018; Russo-Ponsaran et al., 2016, 2014; Silver & Oakes, 2001; Swettenham, 1996; Thomeer et al., 2011, 2015; White et al., 2018).

3.7.3. Rewards and feedback

All CBI included some reward system, with the exception of *Join In!,* which only provided feedback (Bauminger-Zviely et al., 2013). Four programs incorporated both reward and feedback systems within the game (Gordon et al., 2014; Silver & Oakes, 2001; Swettenham, 1996). Rewards system typically took the form of auditory, visual or textual positive reinforcement, with some games rewarding players with collectables such as points or objects (Beaumont & Sofronoff, 2008; Bernard-Opitz et al., 2001; Cheng et al., 2018; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Gordon et al., 2014; Hopkins et al., 2011; Jeffries et al., 2016; Jouen et al., 2017; LaCava et al., 2007; Lacava et al., 2010; Lopata et al., 2016; Malinverni et al., 2017; Miller et al., 2018; Rice et al., 2015; Serret et al., 2014; Tanaka et al., 2010; Thomeer et al., 2011, 2015; White et al., 2018). Few CBI incorporated feedback systems which provided players with information on their progress throughout the intervention. Typically, feedback systems relied on visual progress bars (Gordon et al., 2014; Serret et al., 2014) or provided players with hints or prompts (Bauminger-Zviely et al., 2013; Silver & Oakes, 2001; Swettenham, 1996).

3.7.4. Increasing levels of difficulty

Four CBI employed a variety of contexts, increasing the level of the difficulty within the game by introducing increasingly complex stimuli with additional emotional or environmental clues, presenting more trials and distractors (Beaumont & Sofronoff, 2008; Faja et al., 2008, 2012; Fletcher-Watson et al., 2015; Serret et al., 2014). Several programs demonstrated some attempt to adjust the level of difficulty of the game by increasing the speed of presentation, number of completed trials or distractors (Bauminger-Zviely et al., 2013; Friedrich et al., 2015; Jeffries et al., 2016; Jouen et al., 2017; Russo-Ponsaran et al., 2014, 2016). Other programs addressed complexity through addressing skills in different stages and contexts (Bauminger-Zviely et al., 2013; Bernard-Opitz et al., 2001; Bölte et al., 2002, 2006; Bölte et al., 2015; Cheng et al., 2018; Fridenson-Hayo et al., 2017; Malinverni et al., 2017; Silver & Oakes, 2001; White et al., 2018) or adjusting the complexity of emotions, without arranging the tasks in a stepwise level of difficulty (Golan & Baron-Cohen, 2006; Hopkins et al., 2011; Rice et al., 2015). Overall, four CBI maintained a consistent level of difficulty throughout the game (Bernardini et al., 2014; Gordon et al., 2014; Swettenham, 1996).

3.7.5. Individualisation

Individualisation features were either implemented using an in-built computer system or manually customised by a facilitator or the player. Among the included CBI, individualisation features were not reported in 14 programs (Beaumont & Sofronoff, 2008; Bernardini et al., 2014; Bernard-Opitz et al., 2001; Cheng et al., 2018; Faja et al., 2008, 2012; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Friedrich et al., 2015; Gordon et al., 2014; Malinverni et al., 2017; Miller et al., 2018; Silver & Oakes, 2001; Swettenham, 1996; White et al., 2018). Only three programs reported in-built individualisation features, using starting points individualised to the player or automatic functions adapted according to the players' progress (Bölte et al., 2002, 2006; Bölte et al., 2015; Friedrich et al., 2015; Serret et al., 2014). For example, in a neurofeedback game, players were required to achieve a threshold of 80% before progressing to the next level (Friedrich et al., 2015) or in *JeStiMule,* a failed task was automatically presented in subsequent trials (Serret et al., 2014).

Some interventions individualised the program according to facilitator's judgement (Bauminger-Zviely et al., 2013; Jouen et al., 2017; Russo-Ponsaran et al., 2014, 2016) or by allowing players to autonomously customise their own game pathways (Golan et al., 2010; Golan & Baron-Cohen, 2006; Hopkins et al., 2011; Jeffries et al., 2016; LaCava et al., 2007; Lacava et al., 2010; Lopata et al., 2016; Rice et al., 2015; Tanaka et al., 2010; Thomeer et al., 2011, 2015). These external individualisation strategies were typically used to adjust the level of difficulty of the game tasks.

3.7.6. Summary of design quality

Overall, the minority of social emotional CBI aimed at autistic individuals integrated Serious Game principles, with interventions included in this review obtaining on average a Serious Game score of 45% (limited). Table 7 provides an overview of the classification of the included studies in relation to the Serious Game principles (Whyte et al., 2015).

3.8. Effects of interventions: Social emotional outcomes

3.8.1. Close generalisation outcomes

Fifteen of the 17 included studies evaluated close generalisation outcomes. The overall random effects meta-regressions for close generalisation outcomes was significant ($g = 0.68$. 95% CI [0.50-0.86], $p < 0.01$), indicating that CBI had a medium effect in remediating social emotional close generalisation outcomes for autistic individuals (Fig. 2). The overall heterogeneity for close generalisation social emotional outcomes was not significant ($Q = 17.55$, $p = 0.23$, $I^2 = 27.33\%$). Moderation analysis, examining the influence of Serious Game features on the effect of CBI on close generalisation outcomes was not significant ($g = 0.06$, $p = 0.61$). Intervention duration and age did not reveal any significant moderating effects. The Egger's regression test was significant ($p < 0.02$), suggesting that publication bias may have impacted these results. The trim and fill method accounted for three missing studies due to publication bias, resulting in an adjusted effect size of $g = 0.61$ (Appendix F) (Fig. 3).

3.8.2. Distant generalisation outcomes

Eight studies evaluated distant generalisation outcomes relating to social communication, emotion regulation, emotion imitation, emotion recognition and emotion awareness. Social emotional CBI revealed a medium effect size for distant generalisation outcomes ($g = 0.46$, 95% CI [0.14–0.78], $p = 0.01$). The heterogeneity of effect sizes in distant generalisation outcomes between social emotional CBI was moderate ($Q = 17.38$, $p = 0.02$, $I^2 = 60.53\%$). Importantly, a significant moderating effect of Serious Game features was evident, suggesting that increased Serious Game feature implementation was associated with greater CBI-linked improvements in distant generalisation outcomes ($g = 0.31$, $p = 0.03$). Intervention duration and age did not reveal any significant moderating effects.

3.8.3. Transferability outcomes

Eleven studies measured six transferability outcomes, behaviour, social skills, ASD symptomatology, brain activation, language and theory of mind scores. The potential for social emotional CBI to improve skills outside of the training context was marginally significant and associated with a small effect size of 0.32 (95% CI [-0.01- 0.65], $p = 0.06$). Effect sizes for transferability outcomes across studies were moderately heterogeneous ($Q = 25.72$, $p < 0.01$, $I^2 = 63.76\%$). Serious Game features demonstrated a significant moderating effect on CBI-linked improvements in transferability outcomes ($g = 0.34$, $p = 0.03$), indicating that higher integration of Serious Game principles was associated with improved transferability outcomes. Intervention duration and age did not reveal any significant moderating effects. Non-significant results were observed for the Egger's test ($p = 0.21$) and visual inspection of funnel plots showed a symmetrical distribution of effect sizes, suggesting a low likelihood that publication bias impacted these results (Appendix F) (Fig. 4).

3.8.4. Maintenance outcomes

Meta-analysis of maintenance outcomes for social emotional CBIs revealed an overall small effect ($g = 0.31$, 95% CI [0.03-0.59], $p = 0.03$) across all outcomes post-intervention. Maintenance outcomes were significantly heterogeneous ($Q = 29.24$, $p < 0.01$, $I^2 = 67.39\%$). Incorporating Serious Game principles in CBIs demonstrated no significant impact in improving social emotional outcomes post intervention ($g = -0.12$, $p = 0.26$). Intervention duration and age did not reveal any significant moderating effects. Egger's test revealed no significant effects in funnel plot asymmetry ($p = 0.88$), suggesting a low likelihood that publication bias impacted on the overall effect size (Appendix F).

Sub-group analyses of close ($g = 0.53$, 95% CI [-0.12-1.18], $p = 0.11$) and distant generalisation ($g = 0.46$, 95% CI [-0.02-0.93],

Table 7
Serious game principles- Classifications of studies.

Author (Year)	Targeted Domain	Serious Game principles					Serious Game Score
		Storyline	Goals	Reward	Difficulty Increases	Individualised	
Bauminger-Zviely et al. (2013)	Social collaboration	Themed	Medium	None	Some	Facilitator	4
Bauminger-Zviely et al. (2013)	Social conversation	None	Medium	Reward	Some	Choice/ Facilitator	4
Beaumont and Sofronoff (2008)	Social interaction (emotion management, initiating and maintaining interactions, managing bullies)	Yes	Long	Reward	Yes	None	7
Bernardini et al. (2013)	Joint attention and symbol use	Some	Medium	Reward	None	None	3
Bernard-Opitz et al. (2001)	Social problem solving	None	Short	Reward	Some	None	3
Bölte et al. (2002), 2006; Bölte et al. (2015)	Emotion recognition	None	Short	Reward	Some	Yes	5
Cheng et al. (2018)	Emotion recognition	None	Short	Reward	Some	None	3
Faja et al. (2007); Faja et al. (2012)	Face processing	None	Short	Reward	Yes	None	4
Friedenson-Hayo et al. (2017)	Emotion recognition	Some	Short	Reward	Some	None	4
Fletcher-Watson et al. (2015)	Joint attention	Themed	Medium	Reward	Yes	None	5
Friedrich et al. (2015)	Social imitation	Some	Medium	Reward	Some	Yes	6
Golan and Baron-Cohen (2006); LaCava et al. (2007), 2010; Lopata et al. (2016); Thomeer et al. (2011), 2015	Emotion recognition	None	Short	Reward	Some	Choice	4
Gordan et al. (2014)	Emotion imitation	Themed	Short	Both	None	None	4
Hopkins et al. (2011); Rice et al. (2015)	Face processing	Themed	Medium	Reward	Some	Choice	5
Jeffries et al. (2016)	Joint attention	Yes	Long	Reward	Some	Choice	7
Jouen et al. (2017)	Joint attention	None	Medium	Reward	Some	Facilitator	4
Malinverni et al. (2017)	Social initiation	Yes	Long	Reward	Some	None	6
Miller et al. (2017)	Joint attention	None	Short	Reward	None	None	2
Russo-Ponsaran et al. (2014), 2016	Emotion recognition	None	Short	Reward	Some	Facilitator	4
Serret et al. (2014)	Emotion recognition	Some	Long	Both	Yes	Yes	9
Silver and Oakes (2001)	Emotion recognition and theory of mind	None	Short	Both	Some	None	4
Swettenham (1996)	Theory of mind	Some	Short	Both	None	None	4
Tanaka et al. (2010)	Face processing	Themed	Medium	Reward	Some	Choice	5
White et al. (2018)	Emotion recognition and expression	None	Short	Reward	Some	None	2

$p = 0.06$) maintenance outcomes indicated a non-significant effect for autistic individuals participating in a social emotional CBI maintaining these skills post-intervention. There was a negligible effect for CBI in maintaining transferable outcomes post intervention ($g = 0.03$, 95% CI [-0.30-0.37], $p = 0.84$) (Fig. 5).

3.8.5. *Engagement outcomes*

While few experimental studies reported attrition rates (Bernardini et al., 2014; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Jouen et al., 2017), available rates revealed that on average 19.17% of participants dropout from CBI. Reported attrition rates varied across studies, ranging from low (0–11%) (Fletcher-Watson et al., 2015; Jouen et al., 2017) to high (34%) (Bernardini et al., 2014). Qualitative reports from parents and participants revealed generally high levels of satisfaction with CBI (Cheng et al., 2018; Fletcher-Watson et al., 2015; LaCava et al., 2007; Lacava et al., 2010; Thomeer et al., 2011, 2015). Two feasibility studies of the *Mindreading* game reported participants preferentially spent the majority of their time in the 'games' and 'reward' zones, finding the repetitive nature of the game uninteresting (LaCava et al., 2007; Lacava et al., 2010). Engagement outcomes are summarised in Table 8.

4. Discussion

Using a Serious Game Assessment Tool to quantify the implementation of Serious Game principles in CBI for autistic individuals, the current review investigated the influence of Serious Game design on the effect of CBI in remediating the social emotional skills of autistic individuals. While CBI was associated with an overall improvement in social emotional outcomes, interestingly, this effect was further moderated by the inclusion of Serious Game design principles. Specifically, a greater implementation of Serious Game design principles in designing CBI was associated with greater improvements in distant generalisation and transferability outcomes. These findings support the hypothesis that integrating Serious Game principles in social emotional CBI will result in improved outcomes for autistic participants (Whyte et al., 2015).

The present review examined for the first time the effect of CBI for autistic individuals across a spectrum of scaffolded social

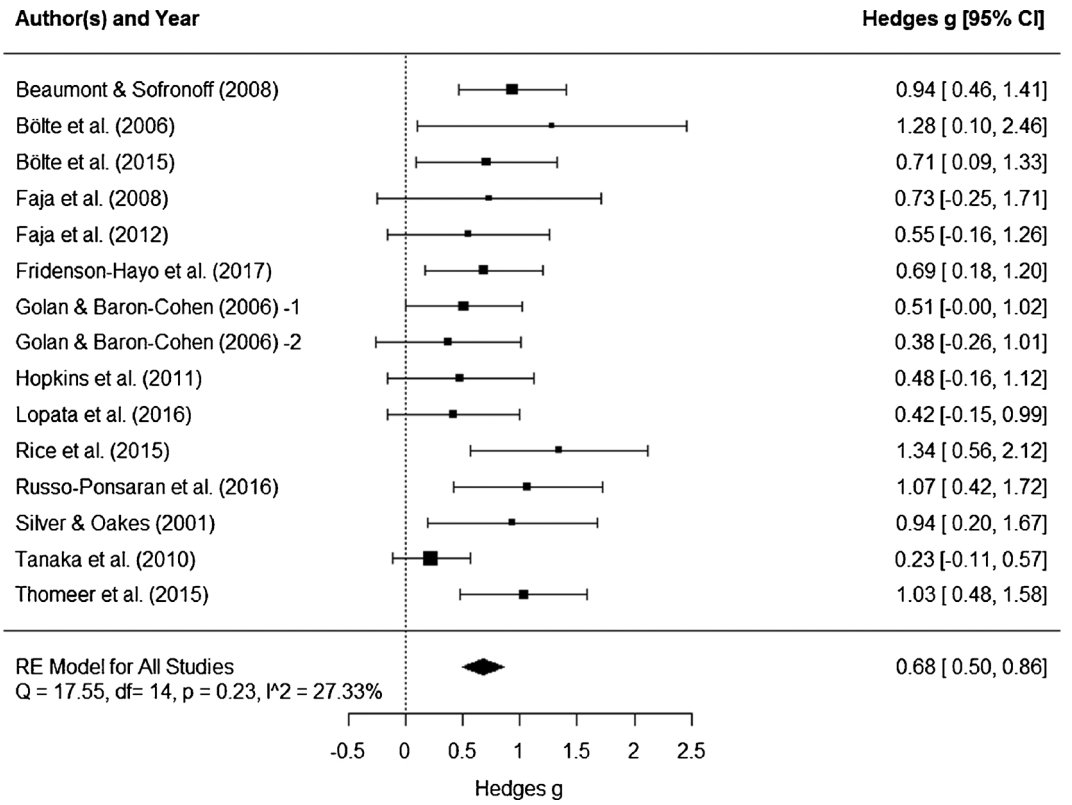

Fig. 2. Forrest plot comparison for pre-post-test close generalisation outcomes between social emotional computer based interventions and control groups.

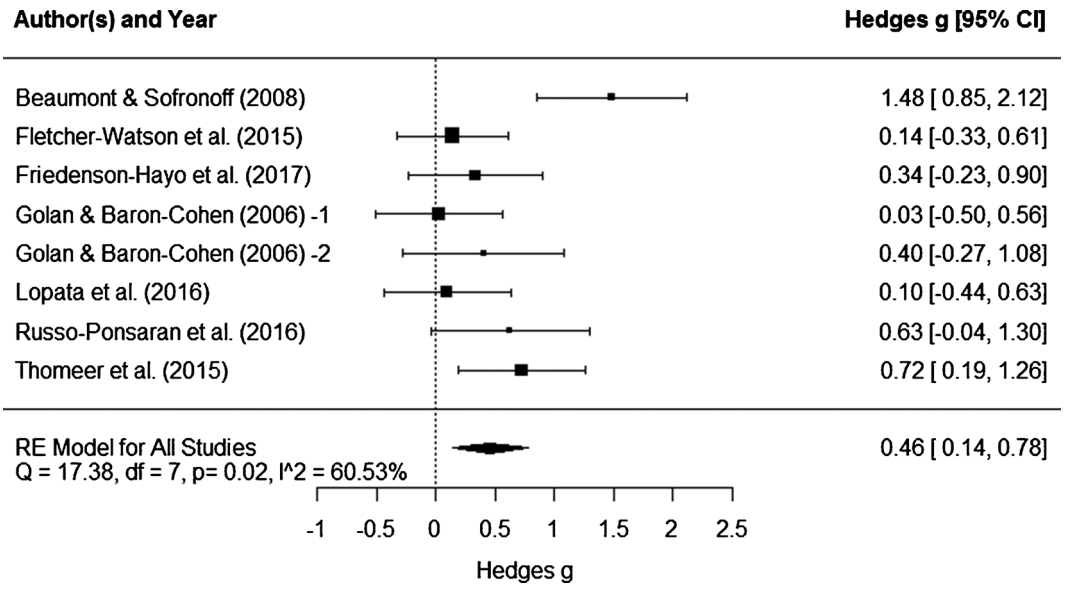

Fig. 3. Forrest plot comparison for pre-post-test distant generalisation outcomes between social emotional computer based interventions and control groups.

emotional outcomes. Meta-analyses revealed an overall positive effect of CBI in targeting close generalisation, distant generalisation, transferability to other skills and maintenance of intervention gains for autistic individuals, with the largest effect size observed for close generalisation outcomes ($g = 0.68$). Medium effect size was found for distant generalisation outcomes ($g = 0.46$), which is comparable to a previous meta-analysis evaluating the overall effect of innovative technologies for autistic individuals ($d = 0.47$)

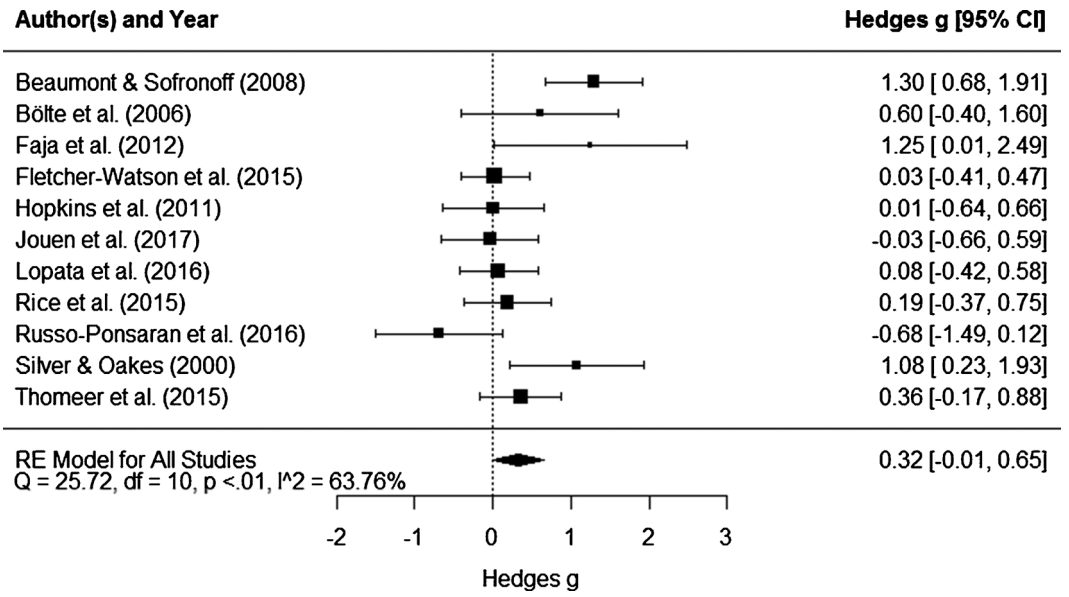

Fig. 4. Forrest plot comparison for pre-post-test transferability outcomes between social emotional computer based interventions and control groups.

(Grynszpan et al., 2014). A medium effect was similarly observed for social skills groups for autistic individuals (Gates et al., 2017; Reichow et al., 2012). Given the comparable results, it is unclear whether social emotional skills are better targeted in group based settings or through CBI. CBI may be a promising avenue for future research to explore, given their key features leverage the recognised strengths of autistic individuals as visual learners and are aligned with their interest in computers (Heimann et al., 1995; Shane & Albert, 2008).

It is interesting to note the influence of Serious Game design principles on distant generalisation and transferability to skills, extending those skills specifically targeted within the interventions. It has been argued that impacting the distant generalisation of skills is a criteria of critical importance when evaluating the effect of a given intervention (Bölte, Golan, Goodwin, & Zwaigenbaum, 2010). The ultimate effect of interventions in targeting aspects of impairment such as emotion recognition must be judged in relation to their transfer to everyday functioning (Bölte et al., 2017). Specifically, we posit that improvements in distant generalisation are likely to be more indicative of genuine improvement in a given social emotional skill than improvements in close generalisation, given these skills are assessed in different contexts. Measures of distant generalisation likely reflect the generalisation of targeted skills to everyday contexts supporting improvements in social abilities not explicitly addressed in the intervention itself. Paralleling this argument, we suggest that the observed increase in transferability to other skills associated with the implementation of Serious Game features in CBI provides some evidence that these features facilitate improvement across a broader range of social emotional outcomes. Collectively, the finding that Serious Game features may specifically enhance distant generalisation and transferability outcomes further bolsters its relevance for the development of social emotional CBI.

4.1. Limitations

In addition to examining the influence of Serious Game design on CBI social emotional outcomes, we sought to understand if these game designs increased participant engagement. While previous reports propose that CBI may be a motivating platform for autistic individuals to learn complex skills (Heimann et al., 1995), understanding of the motivational value of such interventions is limited by reporting of engagement outcomes across the included studies. Only five available studies (Bernardini et al., 2014; Fletcher-Watson et al., 2015; Fridenson-Hayo et al., 2017; Golan & Baron-Cohen, 2006; Jouen et al., 2017) reported attrition rates, despite it being specified as one of the reporting requirements under the Consolidated Standards of Reporting Trials (CONSORT) guidelines. Future research evaluating CBI may benefit from the reporting of attrition rates or other potential outcome measurements for engagement, such as gameplay statistics or participants' satisfaction during the game (Fletcher-Watson, 2014).

A number of other inherent limitations of the present review must also be acknowledged. It was evident that a number of studies evaluating CBI did not implement a formal randomised controlled trial design. While we included these studies in the meta-analysis in order to best represent the existing literature, the inclusion of non-randomised controlled trials in the analysis may have introduced performance and detection bias due to a lack of blinding of participants and assessors. Similarly, few studies provided a description of allocation concealment and sequence generalisation, suggesting the need for greater transparency of reporting.

Small sample size and insufficient reporting of participants' characteristics may further limit the generalisability of the results. Participants included in current social emotional CBI in ASD research were largely younger males with high cognitive functioning abilities. The inclusion of larger more representative samples and improving the reporting of participants' characteristics will assist

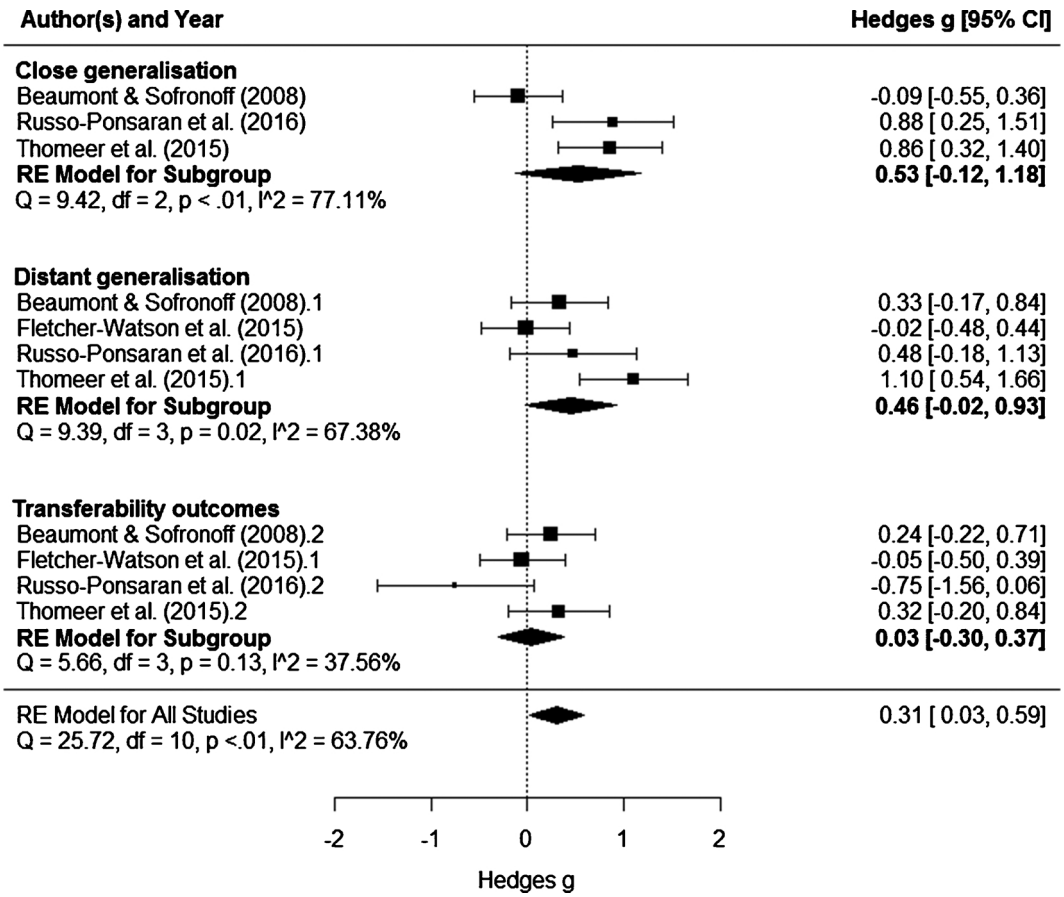

Fig. 5. Forrest plot comparison for pre-post-test maintenance outcomes between social emotional computer based interventions and control groups.

Table 8
User engagement of participants with ASD in social emotional computer based interventions.

Author (Year)	Engagement outcomes	
	Attrition rate %	Qualitative reports
Bernardini et al. (2014)	34	–
Cheng et al. (2018)	–	Most participants reported that they enjoyed the game and found it beneficial in learning facial expressions.
Fletcher-Watson et al. (2015)	11	High percentage of enjoyment reported by parents (92%) and children (96%)
Fridenson-Hayo et al. (2017)	21	–
Golan and Baron-Cohen (2006)-1	21	–
Golan and Baron-Cohen (2006)-2	28	–
Jouen et al. (2017)	0 (Three children had lower participation rates. Total participation rates = 39.9%)	–
LaCava et al. (2007)	–	Participants preferred the “game zone” but found the repetition uninteresting.
Lacava et al. (2010)	–	Participants enjoyed “games” and “reward” zones in the game.
Silver and Oakes (2001)	–	–
Thomeer et al. (2011)	–	High satisfaction rates reported.
Thomeer et al. (2015)	–	Parents and children reported high satisfaction rates

future researchers and clinicians to make informed decisions on the suitability of the CBI for their clients (Berggren et al., 2017).

Heterogeneity in social emotional outcome measures was also observed, making it challenging to draw definite conclusions on the efficacy of CBI on distant generalisation outcomes. Limited availability of standardised assessments evaluating distant generalisation outcomes may in part contribute to the heterogeneity in the findings. Most researchers relied on self-developed measures to capture

the potential efficacy of Serious Game design in improving distant generalisation outcomes, mainly utilising parent or self-reports measures of social emotional improvements in everyday activities. Previous social skills group meta-analyses reported effect sizes differences in parent, self-reported and observer and self-reported outcome measures (Gates et al., 2017; Wolstencroft et al., 2018). Future research should consider developing and testing the reliability and validity of standardised assessments evaluating social emotional skills outside of the training context.

The absence of a moderating effect for intervention duration observed in the present review may be attributed to the heterogeneous nature of the CBI evaluated across studies. Potential confounding factors such as the presence of a facilitator (Golan & Baron-Cohen, 2006; Jouen et al., 2017) or integrating CBI with social skills groups (Beaumont & Sofronoff, 2008; Lopata et al., 2016; Rice et al., 2015) may have lengthen the duration of the intervention beyond that actually reported. Additionally, underrepresentation of older autistic individuals (21%) in this review could in part have contributed to the observed absence of significant moderating effect of age. Previous meta-analysis on technological based interventions in ASD reported comparable non-significant moderating effects for intervention duration, age and IQ on intervention outcomes (Grynszpan et al., 2014). The authors concluded that this was likely due to heterogeneous methodologies, small sample sizes, and the lack of representation from older and lower functioning participants.

4.2. Future implications

The findings of this review provide insights into the potential for Serious Game design principles to enhance social emotional outcomes of autistic individuals. Given the present finding that Serious Games may specifically enhance the distant generalisation and transferability outcomes of social emotional CBI, future research may seek to identify the underlying mechanisms of this effect. We speculate that the opportunity for individuals to experiment with skills in a multi-modal environment may plausibly underpin the increased efficacy of Serious Game designs in improving distant generalisation skills and transfer to other skills. For example, CBI incorporating narrative elements and increasing levels of difficulty such as the Secret Agent Society (Beaumont & Sofronoff, 2008) and JeStimule (Serret et al., 2014) offer the unique opportunity for individuals to contextualise their learning across multiple contexts and graded levels of difficulty, rather than being limited by a singular static environment. This approach likely results in more robust skill development and, a greater generalisation of learnt skills.

Future research could identify the most appropriate types of game designs for autistic individuals. While the present findings take an important step forward in demonstrating the relevance of Serious Game design in CBI, such interventions will only be as effective as they are accepted by their intended populations. Approaches to developing CBI should include co-production techniques involving autistic individuals in determining the appropriateness and motivational appeal of the CBI features (Frauenberger, Good, & Keay-Bright, 2011). Further research engaging autistic individuals is required to identify specific narrative styles promoting their learning and motivation (Tang, Falkmer, Bölte, & Girdler, 2018).

5. Conclusion

The results from this review suggest that currently available social emotional CBI present some limitations in the application of Serious Game design principles. The meta-analyses revealed that implementing social emotional CBI resulted in a small to large overall improvements in social emotional skills in autistic individuals, with the largest effect for close generalisation skills. Social emotional improvements were found to be moderated by the increased application of Serious Game principles. This result highlights the relevance of Serious Game principles in enhancing the outcomes of social emotional CBI, with specific reference to improving distant generalisation and transferability skills. The Serious Game Assessment Tool developed as part of this review enables future research to systematically evaluate the implementation of Serious Game principles in the design of game-based interventions. Future research may consider using the five Serious Game design principles as a potential avenue for guiding the development of prospective interventions.

Acknowledgements

The authors would like to acknowledge Matthew Albrecht for providing software support, Richard Parsons and Torbjorn Falkmer for their statistical advice. Funding for this study was provided by the Cooperative Research Centre for Living with Autism (Autism CRC) [Project Number 3.032RS]. JT was supported by funding from an Australian Postgraduate Award scholarship from the Australian Federal Government. The funding bodies had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript. Approval was obtained from Autism CRC for submission of manuscript for publication.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at https://doi.org/10.1016/j.rasd.2019.101412.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)* (5th ed.). Washington, DC: American Psychiatric Publication.

Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.). *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 3–20). New York: Oxford University Press.
Baron-Cohen, S. (2006). The hyper-systemizing, assortative mating theory of autism. *Progress in Neuro-psychopharmacology & Biological Psychiatry, 30*(5), 865–872. https://doi.org/10.1016/j.pnpbp.2006.01.010.
Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17.
Bauminger-Zviely, N., Eden, S., Zancanaro, M., Weiss, P. L., & Gal, E. (2013). Increasing social engagement in children with high-functioning autism spectrum disorder using collaborative technologies in the school environment. *Autism, 17*(3), 317–339. https://doi.org/10.1177/1362361312472989.
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.org/10.1111/j.1469-7610.2008.01920.x.
Berggren, S., Fletcher-Watson, S., Milenkovic, N., Marschik, P. B., Bölte, S., & Jonsson, U. (2017). Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. *Developmental Neurorehabilitation,* 1–14. https://doi.org/10.1080/17518423.2017.1305004.
Bernard-Opitz, V., Sriram, N., & Nakhoda-Sapuan, S. (2001). Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. *Journal of Autism and Developmental Disorders, 31*(4), 377–384.
Bernardini, S., Porayska-Pomsta, K., & Smith, T. J. (2014). ECHOES: An intelligent serious game for fostering social communication in children with autism. *Information Sciences, 264*, 41–60. https://doi.org/10.1016/j.ins.2013.10.027.
Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hodges, & J. C. Valentine (Eds.). *Handbook of research synthesis and meta-analysis* (pp. 221–236). New York, USA: Russell Sage Foundation.
Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97–111. https://doi.org/10.1002/jrsm.12.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Multiple outcomes or time-points within a study. In M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein (Eds.). *Introduction to meta-analysis* (pp. 225–238). Chichester, UK: John Wiley and Sons, Ltd.
Bölte, S., Ciaramidaro, A., Schlitt, S., Hainz, D., Kliemann, D., Beyer, A., et al. (2015). Training-induced plasticity of the social brain in autism spectrum disorder. *The British Journal of Psychiatry, 207*(2), 149–157. https://doi.org/10.1192/bjp.bp.113.143784.
Bölte, S., Feineis-Matthews, S., Leber, S., Dierks, T., Hubl, D., & Poustka, F. (2002). The development and evaluation of a computer-based program to test and to teach the recognition of facial affect. *International Journal of Circumpolar Health, 61*(Suppl 2), 61–68.
Bölte, S., Golan, O., Goodwin, M. S., & Zwaigenbaum, L. (2010). What can innovative technologies do for Autism Spectrum disorders? *Autism, 14*(3), 155–159. https://doi.org/10.1177/1362361310365028.
Bölte, S., Hubl, D., Feineis-Matthews, S., Prvulovic, D., Dierks, T., & Poustka, F. (2006). Facial affect recognition training in autism: Can we animate the fusiform gyrus? *Behavioral Neuroscience, 120*(1), 211–216. https://doi.org/10.1037/0735-7044.120.1.211.
Bölte, S., Mahdi, S., de Vries, P. J., Granlund, M., Robison, J. E., Shulman, C., et al. (2017). *The Gestalt of Functioning in Autism spectrum Disorder: Results of the International Conference to Develop Final Consensus ICF core Sets.* Manuscript submitted for publication.
Catalano, C. E., Luccini, A. M., & Mortara, M. (2014). Guidelines for an effective design of serious games. *International Journal of Serious Games, 1*(1), 1–13. https://doi.org/10.17083/ijsg.v1i1.8.
Cheng, Y., Luo, S. Y., Lin, H. C., & Yang, C. S. (2018). Investigating mobile emotional learning for children with autistic spectrum disorders. *International Journal of Developmental Disabilities, 64*(1), 25–34. https://doi.org/10.1080/20473869.2016.1206729.
Choque Olsson, N., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., et al. (2017). Social skills training for children and adolescents with autism spectrum disorder: A randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 56*(7), 585–592. https://doi.org/10.1016/j.jaac.2017.05.001.
Clark, B. G., Magill-Evans, J. E., & Koning, C. J. (2015). Youth with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 30*(1), 57–64. https://doi.org/10.1177/1088357614522289.
Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004.
Constantino, J. N., & Gruber, C. P. (2007). *Social responsiveness scale (SRS).* Los Angeles, CA: Western Psychological Services.
Del Re, A. C. (2013). *compute.es: Compute effect sizes version 0.2-2.* R package.
Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology, 11*(1), 37–50.
Del Re, A. C., & Hoyt, W. T. (2014). *MAd: Meta-analysis with mean differences version 0.8-2.* R package.
Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–Based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634.
Faja, S., Aylward, E., Bernier, R., & Dawson, G. (2008). Becoming a face expert: A computerized face-training program for high-functioning individuals with autism spectrum disorders. *Developmental Neuropsychology, 33*(1), 1–24. https://doi.org/10.1080/87565640701729573.
Faja, S., Webb, S. J., Jones, E., Merkle, K., Kamara, D., Bavaro, J., et al. (2012). The effects of face expertise training on the behavioral performance and brain activity of adults with high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(2), 278–293. https://doi.org/10.1007/s10803-011-1243-8.
Fletcher-Watson, S. (2014). A targeted review of computer-assisted learning for people with autism spectrum disorder: Towards a consistent methodology. *Review Journal of Autism and Developmental Disorders, 1*(2), 87–100. https://doi.org/10.1007/s40489-013-0003-4.
Fletcher-Watson, S., Petrou, A., Scott-Barrett, J., Dicks, P., Graham, C., O'Hare, A., et al. (2015). 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.
Frauenberger, C., Good, J., & Keay-Bright, W. (2011). Designing technology for children with special needs: Bridging perspectives through participatory design. *CoDesign, 7*(1), 1–28. https://doi.org/10.1080/15710882.2011.587013.
Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Meir-Goren, N., et al. (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. https://doi.org/10.1007/s00787-017-0968-0.
Friedrich, E., Sivanathan, A., Lim, T., Suttie, N., Louchart, S., Pillen, S., et al. (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. https://doi.org/10.1007/s10803-015-2523-5.
Gates, J. A., Kang, E., & Lerner, M. D. (2017). Efficacy of group social skills interventions for youth with autism spectrum disorder: A systematic review and meta-analysis. *Clinical Psychology Review, 52*, 164–181. https://doi.org/10.1016/j.cpr.2017.01.006.
Goh, D. H., Ang, R. P., & Tan, H. C. (2008). Strategies for designing effective psychotherapeutic gaming interventions for children and adolescents. *Computers in Human Behavior, 24*(5), 2217–2235. https://doi.org/10.1016/j.chb.2007.10.007.
Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., et al. (2010). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders, 40*(3), 269–279. https://doi.org/10.1007/s10803-009-0862-9.
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 Psychopatholgy, 18*(2), 591–617. https://doi.org/10.1017/S0954579406060305.
Goodwin, M. S. (2008). Enhancing and accelerating the pace of autism research and treatment: The promise of developing innovative technology. *Focus on Autism and Other Developmental Disabilities, 23*(2), 125–128. https://doi.org/10.1177/1088357608316678.
Gordon, I., Pierce, M. D., Bartlett, M. S., & Tanaka, J. W. (2014). Training facial expression production in children on the autism spectrum. *Journal of Autism and Developmental Disorders, 44*(10), 2486–2498. https://doi.org/10.1007/s10803-014-2118-6.
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*(Supplement C), 195–211. https://doi.org/10.1016/j.compedu.2017.05.002.
Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. *Autism, 18*(4), 346–361. https://doi.org/10.1177/1362361313476767.
Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews, 42*, 279–297. https://doi.org/10.1016/j.neubiorev.2014.03.013.

Habgood, M. P. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences, 20*(2), 169–206. https://doi.org/10.1080/10508406.2010.508029.
Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review, 20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6.
Heimann, M., Nelson, K. E., Tjus, T., & Gillberg, C. (1995). Increasing reading and communication skills in children with autism through an interactive multimedia computer program. *Journal of Autism and Developmental Disorders, 25*(5), 459–480.
Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions, Vol. 5.1.0*. Chichester, England; Hoboken, NJ: The Cochrane Collaboration.
Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.
Hirvikoski, T., Mittendorfer-Rutz, E., Boman, M., Larsson, H., Lichtenstein, P., & Bölte, S. (2016). Premature mortality in autism spectrum disorder. *The British Journal of Psychiatry, 208*(3), 232–238. https://doi.org/10.1192/bjp.bp.114.160192.
Hopkins, I., Gower, M., Perez, T., Smith, D., Amthor, F., Casey Wimsatt, F., et al. (2011). Avatar assistant: Improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders, 41*(11), 1543–1555. https://doi.org/10.1007/s10803-011-1179-z.
Howlin, P., Moss, P., Savage, S., & Rutter, M. (2013). Social outcomes in mid-to later adulthood among individuals diagnosed with autism and average nonverbal IQ as children. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(6), 572–581. https://doi.org/10.1016/j.jaac.2013.02.017.
Jeffries, T., Crosland, K., & Miltenberger, R. (2016). Evaluating a tablet application and differential reinforcement to increase eye contact in children with autism. *Journal of Applied Behavior Analysis, 49*(1), 182–187. https://doi.org/10.1002/jaba.262.
Jonsson, U., Alaie, I., Lofgren Wilteus, A., Zander, E., Marschik, P. B., Coghill, D., et al. (2017). Annual research review: Quality of life and childhood mental and behavioural disorders- A critical review of the research. *Journal of Child Psychology and Psychiatry, 58*(4), 439–469. https://doi.org/10.1111/jcpp.12645.
Jouen, A.-L., Narzisi, A., Xavier, J., Tilmont, E., Bodeau, N., Bono, V., et al. (2017). GOLIAH (Gaming open Library for Intervention in Autism at Home): A 6-month single blind matched controlled exploratory study. *Child and Adolescent Psychiatry and Mental Health, 11*(1), 17. https://doi.org/10.1186/s13034-017-0154-7.
Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education.* San Francisco: Pfeiffer.
Kmet, L. M., Lee, R. C., & Cook, L. S. (2011). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields.* Alberta, Canada: Alberta Heritage Foundation for Medical Research.
Krug, D. A., Arick, J. R., & Almond, P. (2008). *Autism screening instrument for educational planning: Background and development.* Austin, TX: Western Psychological Services.
LaCava, P. G., Golan, O., Baron-Cohen, S., & Myles, B. S. (2007). Using assistive technology to teach emotion recognition to students with Asperger syndrome: A pilot study. *Remedial and Special Education, 28*(3), 174–181. https://doi.org/10.1177/07419325070280030601.
Lacava, P. G., Rankin, A., Mahlios, E., Cook, K., & Simpson, R. L. (2010). A single case design evaluation of a software and tutor intervention addressing emotion recognition and social interaction in four boys with ASD. *Autism, 14*(3), 161–178. https://doi.org/10.1177/1362361310362085.
Lee, L., Packer, T. L., Tang, S. H., & Girdler, S. (2008). Self-management education programs for age-related macular degeneration: A systematic review. *Australasian Journal on Ageing, 27*(4), 170–176. https://doi.org/10.1111/j.1741-6612.2008.00298.x.
Levy, A., & Perry, A. (2011). Outcomes in adolescents and adults with autism: A review of the literature. *Research in Autism Spectrum Disorders, 5*(4), 1271–1282. https://doi.org/10.1016/j.rasd.2011.01.023.
Lopata, C., Thomeer, M. L., Rodgers, J. D., Donnelly, J. P., & McDonald, C. A. (2016). RCT of mind reading as a component of a psychosocial treatment for high-functioning children with ASD. *Research in Autism Spectrum Disorders, 21*, 25–36. https://doi.org/10.1016/j.rasd.2015.09.003.
Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr, Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223.
Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685.
Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review, 34*(1), 73–86. https://doi.org/10.1016/j.cpr.2013.11.002.
Malinverni, L., Mora-Guiard, J., Padillo, V., Valero, L., Hervas, 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. https://doi.org/10.1016/j.chb.2016.01.018.
Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 35*(6), 655–668. https://doi.org/10.1080/13803395.2013.809700.
Mazurek, M. O., Shattuck, P. T., Wagner, M., & Cooper, B. P. (2012). Prevalence and correlates of screen-based media use among youths with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(8), 1757–1767. https://doi.org/10.1007/s10803-011-1413-8.
McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282. https://doi.org/10.11613/BM.2012.031.
Miller, N., Wyatt, J., Casey, L. B., & Smith, J. B. (2018). Using computer-assisted instruction to increase the eye gaze of children with autism. *Behavioral Interventions, 33*(1), 3–12. https://doi.org/10.1002/bin.1507.
Mitgutsch, K., & Alvarado, N. (2012). Purposeful by design?: A serious game design assessment framework. *Paper Presented at the Proceedings of the International Conference on the Foundations of Digital Games*.
Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059.
Olsson, N. C., Flygare, O., Coco, C., Görling, A., Råde, A., Chen, Q., et al. (2017). Social skills training for children and adolescents with autism spectrum disorder: A randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 56*(7), 585–592. https://doi.org/10.1016/j.jaac.2017.05.001.
Ramdoss, S., Machalicek, W., Rispoli, M., Mulloy, A., Lang, R., & O'Reilly, M. (2012). Computer-based interventions to improve social and emotional skills in individuals with autism spectrum disorders: A systematic review. *Developmental Neurorehabilitation, 15*(2), 119–135. https://doi.org/10.3109/17518423.2011.651655.
Rao, P. A., Beidel, D. C., & Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: A review and recommendations. *Journal of Autism and Developmental Disorders, 38*(2), 353–361. https://doi.org/10.1007/s10803-007-0402-4.
Reichow, B., Steiner, A. M., & Volkmar, F. (2012). Social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *The Cochrane Database of Systematic Reviews, 2012*(7), 1–50. https://doi.org/10.1002/14651858.CD008511.pub2.
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. *Journal of Autism and Developmental Disorders, 45*(7), 2176–2186. https://doi.org/10.1007/s10803-015-2380-2.
Rooney, P. (2012). A theoretical framework for serious game design: Exploring pedagogy, play and fidelity and their implications for design process. *International Journal of Game-Based Learning, 2*(4), 41–60. https://doi.org/10.4018/ijgbl.2012100103.
RStudio Team (2016). *RStudio: Integrated development for r.* Retrieved fromBoston, MA: RStudio, Inc. http://www.rstudio.com/.
Russo-Ponsaran, N. M., Evans-Smith, B., Johnson, J., Russo, J., & McKown, C. (2016). Efficacy of a facial emotion training program for children and adolescents with autism spectrum disorders. *Journal of Nonverbal Behavior, 40*(1), 13–38. https://doi.org/10.1007/s10919-015-0217-5.
Russo-Ponsaran, N. M., Evans-Smith, B., Johnson, J. K., & McKown, C. (2014). A pilot study assessing the feasibility of a facial emotion training paradigm for school-age children with autism spectrum disorders. *Journal of Mental Health Research in Intellectual Disabilities, 7*(2), 169–190. https://doi.org/10.1080/19315864.2013.793440.
Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire.* Los Angeles, CA: Western Psychological Services.
Schopler, E., Reichier, R. J., & Renner, B. R. (1986). *The childhood autism rating scale (CARS), Vol. 2nd*. Torrance, CA: Western Psychological Services.
Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): Preliminary development of a UK screen for mainstream primary-school-age children. *Autism, 6*(1), 9–31. https://doi.org/10.1177/1362361302006001003.
Seida, J. K., Ospina, M. B., Karkhaneh, M., Hartling, L., Smith, V., & Clark, B. (2009). Systematic reviews of psychosocial interventions for autism: An umbrella review. *Developmental Medicine and Child Neurology, 51*(2), 95–104. https://doi.org/10.1111/j.1469-8749.2008.03211.x.
Serret, S., Hun, S., Iakimova, G., Lozada, J., Anastassova, M., Santos, A., et al. (2014). Facing the challenge of teaching emotions to individuals with low-and high-functioning autism using a new serious game: A pilot study. *Molecular Autism, 5*(1), 1. https://doi.org/10.1186/2040-2392-5-37.

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.
Silver, M., & Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or Asperger syndrome to recognize and predict emotions in others. *Autism, 5*(3), 299–316. https://doi.org/10.1177/1362361301005003007.
Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., et al. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal, 343*(1), 1–8. https://doi.org/10.1136/bmj.d4002.
Swettenham, J. (1996). Can children with autism be taught to understand false belief using computers? *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 37*(2), 157–165. https://doi.org/10.1111/j.1469-7610.1996.tb01387.x.
Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., et al. (2010). Using computerized games to teach face recognition skills to children with autism spectrum disorder: The Let's Face it! Program. *Journal of Child Psychology and Psychiatry, 51*(8), 944–952. https://doi.org/10.1111/j.1469-7610.2010.02258.x.
Tang, J., Falkmer, M., Bölte, S., & Girdler, S. (2018). Designing a serious game for youth on the autism spectrum: Perspectives from end-users and professionals. *Journal of Autism and Developmental Disorders, 49*(3), 978–995. https://doi.org/10.1007/s10803-018-3801-9.
Thomeer, M. L., Rodgers, J. D., Lopata, C., McDonald, C. A., Volker, M. A., Toomey, J. A., et al. (2011). Open-trial pilot of Mind Reading and in vivo rehearsal for children with HFASD. *Focus on Autism and Other Developmental Disabilities, 26*(3), 153–161. https://doi.org/10.1177/1088357611414876.
Thomeer, M. L., Smith, R., Lopata, C., Volker, M., Lipinski, A., Rodgers, J., et al. (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. https://doi.org/10.1007/s10803-015-2374-0.
Tobin, M. C., Drager, K. D. R., & Richardson, L. F. (2014). A systematic review of social participation for adults with autism spectrum disorders: Support, social functioning, and quality of life. *Research in Autism Spectrum Disorders, 8*(3), 214–229. https://doi.org/10.1016/j.rasd.2013.12.002.
Turner-Brown, L. M., Perry, T. D., Dichter, G. S., Bodfish, J. W., & Penn, D. L. (2008). Brief report: Feasibility of social cognition and interaction training for adults with high functioning autism. *Journal of Autism and Developmental Disorders, 38*(9), 1777–1784. https://doi.org/10.1007/s10803-008-0545-y.
Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders, 43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5.
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03.
White, S. W., Abbott, L., Wieckowski, A. T., Capriola-Hall, N. N., Aly, S., & Youssef, A. (2018). Feasibility of automated training for facial emotion expression and recognition in autism. *Behavior Therapy, 49*(6), 881–888. https://doi.org/10.1016/j.beth.2017.12.010.
White, S. W., Keonig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: A review of the intervention research. *Journal of Autism and Developmental Disorders, 37*(10), 1858–1868. https://doi.org/10.1007/s10803-006-0320-x.
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.
Wolstencroft, J., Robinson, L., Srinivasan, R., Kerry, E., Mandy, W., & Skuse, D. (2018). A systematic review of group social skills interventions, and meta-analysis of outcomes, for children with high functioning ASD. *Journal of Autism and Developmental Disorders, 48*(7), 2293–2307. https://doi.org/10.1007/s10803-018-3485-1.
Wong, C., & Kasari, C. (2012). Play and joint attention of children with autism in the preschool special education classroom. *Journal of Autism and Developmental Disorders, 42*(10), 2152–2161. https://doi.org/10.1007/s10803-012-1467-2.
Yusoff, A. (2010). *A conceptual framework for serious games and its validation.* Southampton, UK: University of Southhampton.