

***Case Study 1: An Evidence-Based Practice Review Report***

***Theme: School (setting) based interventions for children with special educational needs (SEN)***

***How effective is the Secret Agent Society Programme for improving social skills of children with Autism Spectrum Disorder?***

## 1.1 Summary

This review aimed to explore the effectiveness of the Secret Agent Society programme for targeting the social and emotional skills of children aged 8 to 12 years with autism spectrum disorder (ASD). The programme follows a structured social-emotional curriculum with ongoing skill development through the use of the programme computer game and in-person practice at home and school.

A systematic literature search yielded six studies which met the inclusion criteria. These studies are evaluated using a Weight of Evidence framework (Gough, 2007). All studies reported positive gains in social skills post-intervention, however limitations in methodological study design and statistical analysis made it difficult to make conclusions regarding the generalisability of these findings. Future recommendations for research are outlined.

## 2. 1 Introduction

According to the DSM-5 (American Psychiatric Association, 2013), autism is typically defined by persistent difficulties with social communication and social interaction as well as restricted and repetitive patterns of behaviour, activities or interests including sensory difficulties.

The idea of an 'autism spectrum' accounts for the fact that every child with autism is different. For example, a child described to have social interaction difficulties could be unfazed by social interactions or known to inappropriately interact with others (Frederickson & Cline, 2015). Despite this, it is widely accepted that children with social difficulties often present with difficulty interpreting verbal and non-verbal language, feelings and intentions of others (American Psychiatric Association, 2013). This can make it particularly hard for these individuals to initiate and maintain social friendships.

Due to the difficulties children with autism have initiating and maintaining relationships, this places them at risk of victimisation and bullying from their typically developing peers (Rowley et al., 2012; Schroeder et al., 2014; Symes & Humphrey, 2010). Moreover, due to the lack of relationships these children often miss out on the supportive network of friends and peers known to be protective factors against the negative effects of bullying and victimisation (Chamberlain et al., 2007). This exclusion from peer and friendship groups results in fewer opportunities for these children to engage with others, preventing them from developing their social skills (Schroeder et al., 2014). Therefore, it would feel beneficial for these children to receive

support to develop their social skills, helping them to access and navigate the social world around them.

## 2.2 Secret Agent Society (SAS) Programme

The Secret Agent Society (SAS) Programme was developed by Dr Renae Beaumont, a Clinical Psychologist in Australia to support the social and emotional skills of children aged 8 to 12 years with autism spectrum disorder. The programme typically consists of group meetings for the child, delivered between 11 and 22 sessions for a total of 45-90 minutes. There are an additional two to four sessions post intervention for skill maintenance. These sessions are designed to cover the main components of the Secret Agent Society social and emotional curriculum (The Social Skills Training Institute, 2019), summarised in table 1 below.

**Table 1**

*Outline of key components of the SAS programme social and emotional curriculum*

Components
1. Recognising emotions in themselves and others
2. Expressing their feelings appropriately
3. Coping with feelings of anger and anxiety, including the use of specific relaxation techniques
4. Understanding friendship
5. Communicating and playing with others
6. Coping with teamwork, games and competitiveness
7. Building and maintaining friendships

8. Coping with mistakes, transitions and other difficult situations
  9. Solving social problems
  10. Recognising and managing bullying
  11. Recognising personal strengths and goals
- 

To support the children receiving the intervention, group sessions for parents are incorporated within the programme. These sessions consist of an initial two-hour information session so parents are able to fully understand the requirements of the programme, followed by a further 11 to 22 group meetings of 30-45 minutes or four two-hour sessions over the course of the programme. These group meetings provide opportunities for networking and discussion between parents and facilitators for sharing success and problem-solving difficulties.

Alongside these child sessions and parent support, classroom teachers are regularly updated on the skills and content the children are learning and provided with tips and strategies to support the application and generalisation of their social skills within school. Between sessions children are required to participate in weekly tasks, including the Secret Agent Society computer game and completion of a reflective journal. Children also have access to the 'skill tracker system', a tool that aims to track progress and reward the child's skill development at home and at school.

### 2.3 Theoretical Basis of the Secret Agent Society Programme

The SAS Programme follows an eclectic approach, with underpinnings in various psychological theories, such as theory of mind, executive dysfunction

and weak central coherence. The programme also adopts principles, such as applied behavioural analysis and cognitive behavioural therapy.

Theory of Mind is described as the ability to impute mental states to oneself and others (Premack et al., 1978). It is an important skill that we use in everyday life, particularly when communicating and interacting with others. Individuals with autism tend to experience difficulty with theory of mind tasks (Baron-Cohen, 1985) where the adoption of a view other to their own is required. The SAS programme aims to address these needs by teaching children how to detect their own thoughts and feelings as well as those of others.

Executive functioning skills are commonly agreed as mental processes that encompass working memory, flexible thinking and self-control. Many individuals with autism may face challenges with their executive functioning, particularly with planning, organising and showing flexibility in their thought and actions (Pennington & Ozonoff, 1996). This affects their ability to plan, predict and initiate and respond to social interactions with others. The digital missions and step-by-step problem-solving within the SAS programme are used to specifically target executive functioning by helping them to develop their flexible thinking and planning.

Central coherence refers to an innate ability to make sense of information by putting together individual components to form a 'bigger picture' with the added ability to then generalise this across contexts (Pellicano et al., 2006). It is suggested that individuals with autism have weak central coherence and therefore struggle to perceive connections or make sense of information,

such as subtle verbal or non-verbal clues in everyday interactions. The SAS programme aims to help children develop these skills, such as tone of voice and body language to accurately interpret social situations.

The SAS programme uses positive reinforcement principles of applied behaviour analysis (Matson et al., 2012) to increase desired behaviours. Within the programme, this is achieved through the use of the home-school diary. The diary awards points for use of focus skills at home and school which are exchanged for rewards. These rewards help to maintain motivation to continue engaging with the programme.

Cognitive behavioural therapy is highly structured therapy that aims to challenge unhelpful and irrational thoughts an individual may have whilst also providing strategies and techniques to help cope in challenging situations (Anderson & Morris, 2006). The SAS programme aims to provide children with useful strategies to cope with unpleasant feelings they may feel in certain situations, such as mindfulness or other relaxation techniques.

## 2.4 Rationale for Review

The percentage of pupils with an Education, Health and Care (EHC) plan within the UK has increased once again in the year 2020/2021, with the most common type of need for an EHC plan being Autistic Spectrum Disorders (GovUK, 2021). Due to these ongoing increases in the commissioning of EHC plans there is emphasis placed upon the education and inclusion of children with autism in mainstream provision or mainstream settings with an ASD resource base attached. The role of the Educational Psychologist (EP) in the UK is becoming an increasingly supportive one whereby advice is

sought by schools on the best guidance and evidence-based interventions to support these children effectively within their settings. This literature review will support EPs to make an informed judgement about the effectiveness of the Secret Agent Society programme for supporting the social skills of children with autism. Therefore, the question of this literature review is 'How effective is the Secret Agent Society Programme for improving social skills of children with autism spectrum disorder?'

## 2.5 Literature Search

It is important to note that within the DSM-5 (American Psychiatric Association, 2013), the terms 'Autistic Disorder', 'Asperger syndrome', 'Rett's Disorder', 'Childhood Disintegrative Disorder' and 'Pervasive Developmental Disorder – not otherwise specified (PDD-NOS)' were replaced by the term 'Autism Spectrum Disorder'. Therefore, studies which were conducted before the change in DSM criteria for autism will be included as long as they would now be classified within 'Autism Spectrum Disorders'. In addition, the term 'ASD' or 'autism' will be used throughout this literature review to refer to all Autism Spectrum Disorders.

A search of the literature was conducted on 11<sup>th</sup> January, 2022 using six databases, namely Educational Resources Information Center (ERIC), Google Scholar, Medline, PsycINFO, SCOPUS and Web of Science. Search terms used in this literature search are outlined in Table 2.

**Table 2**

*Terms used for Literature Database Search*

Database	Search Term
ERIC	
	“Secret Agent Society”
Google Scholar	
	OR
Medline	
	“Secret Agent Society: Operation Regulation”
PsycINFO	
	OR
SCOPUS	
	“Junio?r Detective Training Program**”
Web of Science	

*Note. Search terms separated by ‘OR’ ensure that the database considers alternative terms for the same concept. Quotation marks are used for exact phrases of concepts (e.g. “Secret Agent Society”), asterisks for truncations and question marks indicate wildcards*

Search results were filtered to include peer reviewed studies and those written in English. The search identified a total of 61 studies. Once studies were filtered by title and duplicates were removed, 16 studies remained. These studies were then screened by abstract using the inclusion and exclusion criteria (Table 3) leaving a total of 13 studies. Full text screening excluded a further seven studies, leaving six studies to be reviewed in relation to this systematic review’s research question (Table 4). Figure 1 illustrates the literature search and selection process in a flowchart.

Appendix A lists full references for studies excluded at abstract and full-text screening.



**Table 3***Inclusion and exclusion criteria used for study screening*

<b>Study feature</b>	<b>Inclusion criteria</b>	<b>Exclusion criteria</b>	<b>Rationale</b>
1. Intervention	The study must follow the main theoretical components of the social skills 'Secret Agent Society' intervention or 'Junior Development Training Programme'. Adaptions to interventions are acceptable as long as they do not change the theoretical or curriculum components of the programme	The study does not use the 'Secret Agent Society' as the social skills intervention  The 'Secret Agent Society: Operation Regulation' adaptation as the programme does not follow a social skills curriculum	To ensure that the reviewer is examining the effectiveness of the 'Secret Agent Society' intervention in improving social skills of children with ASD

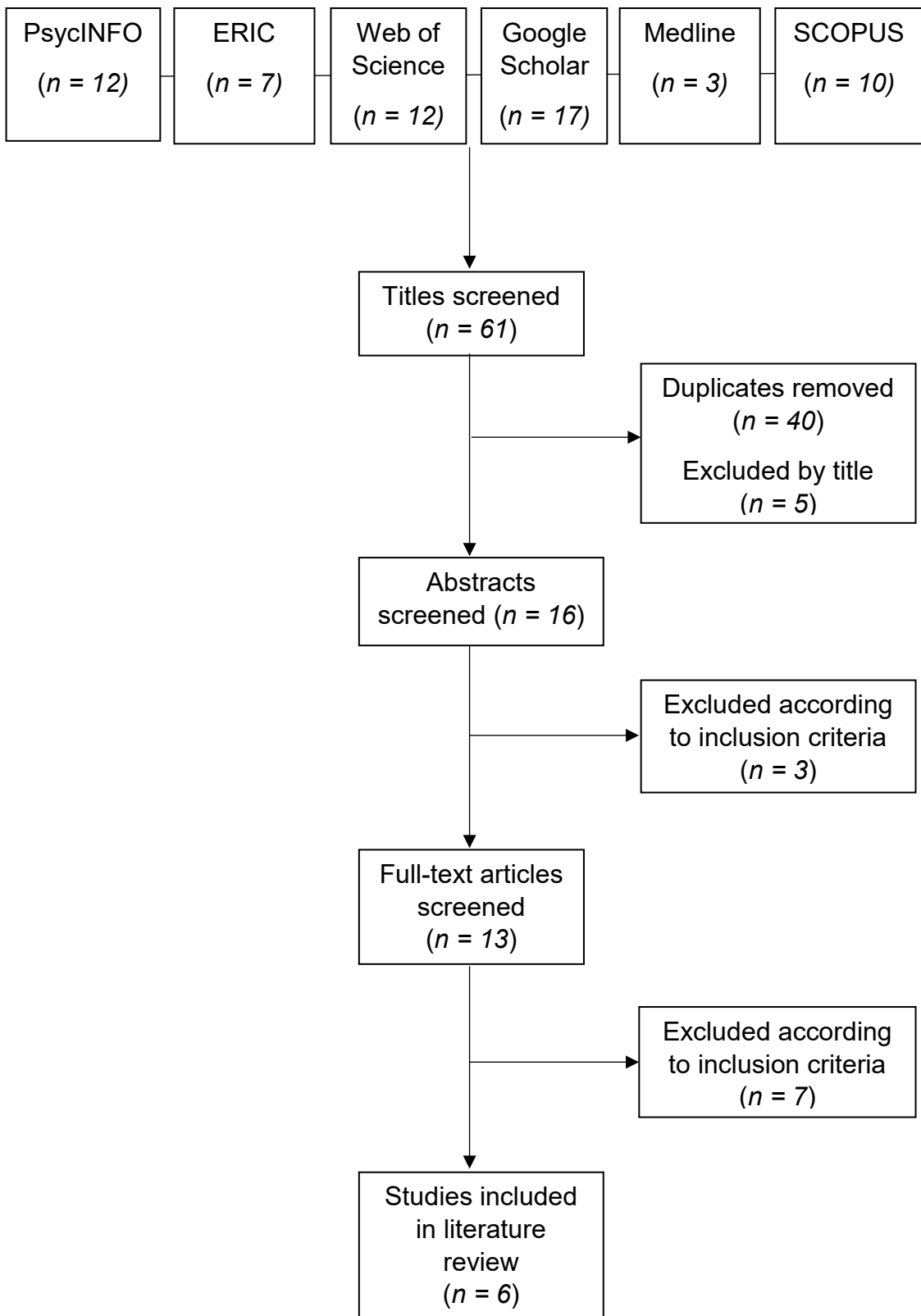
<b>Study feature</b>	<b>Inclusion criteria</b>	<b>Exclusion criteria</b>	<b>Rationale</b>
2. Participants	The study must include children or young people who have a diagnosis of Autism Spectrum Disorder (ASD). This includes previous terms that were classified under the 'Autism Spectrum Disorder' classification in the DSM-5	The study includes typically-developing children or young people or children and young people with special educational needs other than Autism Spectrum Disorder (ASD) or Asperger Syndrome  Participants have a co-morbid disorder that is the child's primary need	To ensure the reviewer can critically evaluate the effectiveness of the intervention for targeting social skills of children and young people with Autism Spectrum Disorder
3. Measures	The study must focus on the social skill outcomes of the children receiving the intervention or other control group	The study does not focus on the social skill outcomes of the children receiving the intervention or other control group	To ensure that the studies relate to the research question

<b>Study feature</b>	<b>Inclusion criteria</b>	<b>Exclusion criteria</b>	<b>Rationale</b>
4. Study design and methodology	The study must use a group-based or single case experimental study design	The study does not use a group-based or single case experimental design study design	To ensure the reviewer can critically evaluate the effectiveness of the intervention by comparing outcomes between or within participants
5. Type of article	The articles must be published in peer-reviewed journals	The articles were not published in a peer reviewed journal	To ensure the studies have been reviewed by independent researchers in the field and are of high-quality
6. Setting and delivery	The study must be delivered within a school, research or home setting by school staff, therapist(s) or parent(s) and	The study is not delivered within a school, research or home setting by school staff, therapist(s) or parent(s) and carer(s)	To allow generalisation of results to school settings and home environments

<b>Study feature</b>	<b>Inclusion criteria</b>	<b>Exclusion criteria</b>	<b>Rationale</b>
	carer(s)		
7. Language	The study must be published in English	The study is not published in English	To ensure research can be understood
8. Country	The study must have been conducted in an OECD country	The study has not been conducted in an OECD country	To ensure participants and systems in which these studies have been conducted have similar demographics and values. This will allow generalisation of results to the UK

**Figure 1**

*Flow diagram of study selection process*



**Table 4**

*Full references of studies included in this systematic literature review*

---

<b>Study Reference</b>
1. Beaumont, R., Rotolone, C., & Sofronoff, K. (2015). The secret agent society social skills program for children with high-functioning autism spectrum disorders: A comparison of two school variants. <i>Psychology in the Schools, 52</i> (4), 390–402. <a href="https://doi.org/10.1002/PITS.21831">https://doi.org/10.1002/PITS.21831</a>
2. Einfeld, S. L., Beaumont, R., Clark, T., Clarke, K. S., Costley, D., Gray, K. M., Horstead, S. K., Redoblado Hodge, M. A., Roberts, J., Sofronoff, K., Taffe, J. R., & Howlin, P. (2018). School-based social skills training for young people with autism spectrum disorders. <i>Journal of Intellectual and Developmental Disability, 43</i> (1), 29–39. <a href="https://doi.org/10.3109/13668250.2017.1326587">https://doi.org/10.3109/13668250.2017.1326587</a>
3. Tan, Y. L., Mazzucchelli, T. G., & Beaumont, R. (2015). An evaluation of individually delivered secret agent society social skills program for children with high-functioning autism spectrum disorders: A pilot study. <i>Behaviour Change, 32</i> (3), 159–174. <a href="https://doi.org/10.1017/bec.2015.7">https://doi.org/10.1017/bec.2015.7</a>
4. Beaumont, R., Walker, H., Weiss, J., & Sofronoff, K. (2021). Randomized Controlled Trial of a Video Gaming-Based Social Skills Program for Children on the Autism Spectrum. <i>Journal of Autism and Developmental Disorders, 51</i> (10), 3637–3650. <a href="https://doi.org/10.1007/S10803-020-04801-Z">https://doi.org/10.1007/S10803-020-04801-Z</a>

---

---

### Study Reference

---

5. Sofronoff, K., Silva, J., & Beaumont, R. (2017). The Secret Agent Society Social-Emotional Skills Program for Children With a High-Functioning Autism Spectrum Disorder: A Parent-Directed Trial. *Focus on Autism and Other Developmental Disabilities*, 32(1), 55–70. <https://doi.org/10.1177/1088357615583467>
  6. 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, and Allied Disciplines*, 49(7), 743–753. <https://doi.org/10.1111/J.1469-7610.2008.01920.X>
- 

### 2.6 Mapping the Field

The six studies identified through the systematic literature search focused upon the Secret Agent Society intervention that had been implemented with pupils with ASD. Table 5 shows the key features of each study.

**Table 5***Mapping the Field using key features of each study*

Author	Study Design	Location	Participants	Intervention	Outcome Measures	Key Findings
Beaumont et al. (2015)	Quasi-experimental design	Mainstream school context  Australia	69 children (64 males; 5 females)  Aged 7-12  Autism Spectrum Disorder diagnosis with a full-scale IQ score of 79 or higher on Wechsler Abbreviated Scale of Intelligence  Assigned into two conditions: Intervention 35 Control 34	Secret Agent Society 10-week programme with adjustments made for school setting  Delivered by school staff  10-week existing social skills programme in school with flexible use of Secret Agent Society Computer Game Pack materials	Social Skills Questionnaire parent version (SSQ-P)  Social Skills Questionnaire teacher version (SSQ-T)  Measures pre and post intervention and at 6-week follow-up	Parent reported social skill gains post-intervention. Gains maintained at follow-up  Teacher reported social skill gains post-intervention. Slight social skill gains at follow-up



Author	Study Design	Location	Participants	Intervention	Outcome Measures	Key Findings
Einfeld et al. (2018)	Quasi-experimental design	ASD primary and secondary 'satellite' classes run by specialist Autism educators Australia	84 children (75 male; 9 female) Aged 8-14 Autism Spectrum Disorder diagnosis Assigned into two conditions: Intervention 26 Waitlist 58	Secret Agent Society programme delivered over 10-13 weeks Delivered by school staff 9 x 90-minute sessions Treatment as usual – Aspect Comprehensive Approach for Education	Social Skills Questionnaire parent version (SSQ-P) Social Skills Questionnaire teacher version (SSQ-T) Measures taken pre and post intervention and at 12-month follow-up	Parent reported social skill gains post-intervention

Author	Study Design	Location	Participants	Intervention	Outcome Measures	Key Findings
Tan et al. (2015)	Single case experimental design	Therapist delivery to family (parent, child and sibling)  Australia	3 children (all male)  Aged 8-11  Autism Spectrum Disorder diagnosis with a full-scale IQ score of 85 or above on Wechsler Abbreviated Scale of Intelligence	Secret Agent Society programme over 9 weeks  Delivered by programme creator and parent  75-minute sessions (first 60 minutes delivered by therapist)	Social Skills Questionnaire parent version (SSQ-P)  Social Skills Questionnaire teacher version (SSQ-T)  Staggered start (baseline), then measures post-intervention and 6-week follow-up	Parent reported social skill gains for all children. Gains were maintained at follow-up

Author	Study Design	Location	Participants	Intervention	Outcome Measures	Key Findings
Beaumont et al. (2021)	Randomised controlled trial	Parent delivery with therapist support  Australia	70 children (60 male; 10 female)  Aged 7 to 12  Autism Spectrum Disorder diagnosis with at least low average cognitive functioning, defined by their full-scale IQ or verbal and perceptual indices on standardised intelligence tests  Randomly allocated into two conditions: Intervention 35 Control 35	Secret Agent Society 10-week programme with adjustments made for parent accessibility  Delivered by parents with support from trained facilitators  10 x 30-minute sessions  Similarly structured 10-week programme without the social or emotional skills component	Social Skills Questionnaire parent version (SSQ-P)  Social Skills Questionnaire teacher version (SSQ-T)  Measures pre and post intervention and at 6-week follow-up	Parent reported social skill gains post-intervention  Parent reported social skill gains greater for SAS programme post-intervention and at follow-up

Author	Study Design	Location	Participants	Intervention	Outcome Measures	Key Findings
Sofronoff et al. (2017)	Quasi-experimental within subject design	Parent delivery with therapist support  Australia	41 children (36 male; 5 female)  Aged 7 to 12  Autism Spectrum Disorder diagnosis with a full-scale IQ of 85 and above on a cognitive assessment	Secret Agent Society 10-week programme with adjustments to self-directed format for parent delivery  Delivered by parents  10 x 90-minute sessions	Social Skills Questionnaire parent version (SSQ-P)  Measures pre and post intervention and at 6-week follow-up	Gains in social skills post-intervention

Author	Study Design	Location	Participants	Intervention	Outcome Measures	Key Findings
Beaumont & Sofronoff (2008)	Randomised controlled trial	Therapist led Australia	49 children (44 male; 5 female) Aged 7 to 11 Autism Spectrum Disorder diagnosis with a full-score IQ score of 85 and above on WISC-III Randomly allocated into two conditions: Intervention 26 Wait-list 23	7-week Junior Detective Training Programme now known as Secret Agent Society Delivered by trained facilitators and parents Week 1 & 2 60-minutes playing computer game and then 60-minutes in small group therapy Week 3 & 4 45-minutes playing computer game and then 75-minutes in small group therapy	Social Skills Questionnaire parent version (SSQ-P) Social Skills Questionnaire teacher version (SSQ-T) Measures taken pre and post intervention and at 5-month follow-up	Parent reported social skill gains post-intervention

*Note.* SSQ-P and SSQ-T refer to the parent and teacher versions of the Social Skills Questionnaire

## 3.1 Critical Evaluation

### 3.2 Weight of Evidence

To critically assess each study within this review a weight of evidence (WoE) framework (Gough, 2007) was used. This framework consists of three components to weight studies and produce an overall WoE score (WoE D), namely methodological quality (WoE A), appropriateness of design (WoE B) and topic relevance (WoE C). WoE scores for each study are summarised in Table 6. Further information on how each WoE score was calculated can be found in Appendix B.

WoE A assessed the methodological quality of the studies by using the Gersten et al. (2005) and Horner et al. (2005) coding protocols to assess studies with a quasi-experimental study and single case experimental study design respectively.

WoE B explored the appropriateness of the evidence for the literature review question. As the question for this literature review focused on the effectiveness of the Secret Agent Society programme, the Petticrew and Roberts (2003) criteria was used to judge the appropriateness of the study design against the question type. Research author and impact on bias was also considered and formed another criterion within WoE B.

WoE C aims to assess the relevance of the evidence to the review question. To explore whether the studies were able to answer the review question, consideration was given towards intervention fidelity, participants, outcome measures and applicability to school settings within the UK.

**Table 6***Overall Weight of Evidence Ratings*

Study	WoE A	WoE B	WoE C	WoE D
Beaumont et al. (2015)	1	1.5	2.4	1.63 (low)
Einfeld et al. (2018)	1	2	2.6	1.87 (medium)
Tan et al. (2015)	2	1.5	2	1.83 (medium)
Beaumont et al. (2021)	2	2.5	1.4	1.97 (medium)
Sofronoff et al. (2017)	1	1.5	1.6	1.37 (low)
Beaumont & Sofronoff (2008)	1	2.5	2.4	1.97 (medium)

*Note.* 1 - 1.6 (low), 1.7 – 2.3 (medium), 2.4 – 3 (high)

### 3.3 Participants

All six studies included participants with a confirmed ASD diagnosis from a paediatrician or other professional. Although four of these studies didn't use the DSM V criteria (Beaumont et al., 2015; Beaumont & Sofronoff, 2008; Einfeld et al., 2018; Tan et al., 2015), participant diagnoses across all studies fulfilled the DSM V criteria (American Psychiatric Association, 2013) for Autism Spectrum Disorder. This resulted in high WoE C scores for all studies.

To ensure that children and young people can access the content of the programme, studies which detailed a confirmed diagnosis of ASD and full-scale IQ score criteria on a specific cognitive assessment received higher weighting on WoE C (Beaumont et al., 2015; Beaumont & Sofronoff, 2008;

Einfeld et al., 2018; Tan et al., 2015). Only one study (Einfeld et al., 2018) made considerations for participants who may have had a cognitive assessment within the past two years. Beaumont et al. (2021) used teacher ratings of academic performance to gauge cognitive ability when cognitive assessment scores were not available. This resulted in a lower WoE C score for this study.

Two studies (Einfeld et al., 2018; Sofronoff et al., 2017) experienced significant levels of participant attrition with many children changing schools or choosing to withdraw from the process. Another study (Beaumont & Sofronoff, 2008) did not report any information on participant attrition. High levels of attrition and omitted information were reflected in low WoE A scores.

### 3.4 Study design

Two of the included studies (Beaumont et al., 2021; Beaumont & Sofronoff, 2008) were randomised controlled trials (RCTs). RCTs are considered to be the most appropriate study design for answering effectiveness questions due to increased experimental control over the independent variable. Therefore, these two studies were awarded a high WoE B rating for study design.

According to Ellopoulos et al. (2005), quasi-experimental designs with a control group are the most robust type of quasi-experimental study designs. Therefore, studies without a control group (Sofronoff et al., 2017) received lower WoE A and WoE B ratings. However, two studies (Beaumont et al., 2015; Einfeld et al., 2018) utilised wait-list control groups as a way to increase their methodological quality, this resulted in higher WoE B scores. One study (Tan et al., 2015) utilised a single case experimental study design



(SCED), however didn't demonstrate experimental effect at three time points, a key component of well-designed SCEDs. Therefore, this study scored low on WoE A and WoE B.

### 3.5 Intervention

Only one of the studies (Einfeld et al., 2018) met all intervention fidelity criteria and received a high WoE C score. A further three studies (Beaumont et al., 2015; Beaumont & Sofronoff, 2008; Sofronoff et al., 2017) adjusted the programme to make delivery viable and as such met three of the WoE C criteria and scored medium on the intervention fidelity WoE C score. One study (Beaumont et al., 2021) met only two criteria as the programme was adapted from the original structure and was delivered in fewer sessions than specified within the manual. This resulted in a low WoE C score.

Fidelity of programme implementation was mentioned in all but one of the included studies (Sofronoff et al., 2017). Two of the studies (Beaumont et al., 2015; Beaumont et al., 2021) used basic information to assess intervention fidelity, such as checklists completed by the facilitator. Two studies (Beaumont & Sofronoff, 2008; Einfeld et al., 2018) used a checklist and videotaped a proportion of sessions which were then independently rated. Differences in assessing fidelity of intervention were reflected within WoE A scores.

All six studies were conducted in Australia. This is an OECD country and therefore has a similar economic and social policy to the UK. Despite this, there will still be differences between education systems and as such

generalisability to the UK is low. Therefore, all studies received a lower WoE C rating for country.

To ensure study findings can be generalised to school settings, studies delivered in a school setting by teaching staff were awarded a high WoE C application to school setting score. This was the case for two studies (Beaumont et al., 2015; Einfeld et al., 2018). The four other studies included in this review (Beaumont et al., 2021; Beaumont & Sofronoff, 2008; Sofronoff et al., 2017; Tan et al., 2015) were delivered outside of the school setting by trained facilitators, resulting in low WoE C application to school setting scores.

### 3.6 Measures

The Emotion Regulation and Social Skills Questionnaire (ERSSQ) was used in all studies. Within an evaluative study of the measure (Butterworth et al., 2014), they found that both the parent and teacher versions of the ERSSQ have been found to have high internal consistency (Cronbach's alpha >.90) alongside good concurrent validity with the Social Skills Questionnaire (SSQ; Spence, 1995). However, they failed to establish the criterion validity of the ERSSQ teacher version. At the time of Beaumont et al.'s (2021) publication, a normative sample had yet to be determined meaning there is no information on the suitability of its use for specific populations. Furthermore, there are no cut-off scores meaning that clinically significant changes pre and post intervention and at follow-up cannot be identified. Due to these issues, it was felt that the Social Skills Questionnaire provided a better measure of a child's 'true' social skills competency. Therefore, this was the outcome measure used to answer the review question.

All six studies included in this review used the Social Skills Questionnaire to measure participants' social skills. The questionnaire consists of 30 statements that describe different social behaviours. The teacher and parent are asked to rate each statement on a scale of zero to two (zero = not true, two = mostly true) for how true it is for their child. A total score is then calculated by adding the ratings of each item. A higher score indicated greater social functioning. All but one study (Sofronoff et al., 2017) used both the teacher (SSQ-T) and parent (SSQ-P) versions of the questionnaire. It should be noted that the Social Skills Questionnaire hasn't been re-standardised since normative data for the measure was gathered in 1995 (Spence, 1995). Therefore, studies which included a measure of generalised social skills would have received the highest WoE C rating. None of the studies used a measure of generalised social skills, such as child observation and as such received a lower WoE C score.

Three studies (Beaumont & Sofronoff, 2008; Einfeld et al., 2018; Sofronoff et al., 2017) included appropriate use of pre, post and follow-up measures to observe an effect of the intervention on participant social skills. This was reflected in WoE A scores. Two studies (Beaumont et al., 2015; Beaumont et al., 2021) aimed to collect follow-up data but experienced difficulties collecting data due to the time-point falling in school holidays or at the end of the school academic year. One study (Beaumont et al., 2021) gathered post-intervention measures 10-weeks after the intervention had finished rather than at the cessation of the programme. This would have had an impact upon the data collected and as such this was acknowledged in the WoE A rating.

### 3.7 Conflict of interest

All studies included within this review had Dr Renae Beaumont as one of the study authors. This highlights a natural conflict of interest as they receive a financial contribution when the SAS programme is sold. Therefore, studies conducted by independent researchers would have received the highest WoE B rating. Five studies (Beaumont et al., 2021; Beaumont & Sofronoff, 2008; Einfeld et al., 2018; Sofronoff et al., 2017; Tan et al., 2015) disclosed a conflict of interest and therefore received a medium WoE B score. The remaining study (Beaumont et al., 2015) provided enough information to suggest there is a potential conflict of interest, however this isn't explicitly stated. This resulted in a low WoE B score.

### 3.8 Study findings and effect sizes

All included studies except one (Tan et al., 2015) reported effect sizes which is reflected in the WoE A score. Two studies (Beaumont et al., 2015; Beaumont & Sofronoff, 2008) reported an eta-squared effect size, two studies used a partial eta-squared effect size and Einfeld et al. (2018) reported significant coefficients. Although these coefficients were unable to be converted to Cohen's *d*, reported means and standard deviations for the intervention and wait-list groups post-intervention meant an effect size could be calculated using the online Campbell Collaboration Calculator. Outcomes in studies reporting an eta-squared or partial eta-squared effect size were converted to the common effect size of Cohen's *d* using the Psychometrica website (Lenhard & Lenhard, 2016).

All studies reported positive effects from participants participating in the SAS programme as outlined in Table 7. Three studies (Beaumont et al., 2021; Beaumont & Sofronoff, 2008; Einfeld et al., 2018) reported small, large and very large gains in social skills when compared to a control group. However, one study (Beaumont et al., 2015) found no difference in social skill gains between groups. It should be noted however, that the control group in this study received an existing social skills programme and utilised game pack materials from the SAS programme.

There is a large discrepancy between the approach and quality of statistical analyses used. One study (Einfeld et al., 2018) conducted a regression analysis to determine the relationship between the SAS programme and social skills. Despite this analysis showing a positive relationship between the SAS programme and social skills over time, this type of analysis does not allow causation to be implied. Therefore, we cannot explicitly state that the SAS programme resulted in social skill gains in participants. Another study (Tan et al., 2015) used a single case experimental study design but did not demonstrate experimental effect at three different time points. Therefore, despite the study findings reporting a positive effect on participant social skills after engaging in the SAS programme, it cannot be concluded that this effect was due to the SAS programme.

All studies experienced issues with participant attrition and missing data post-intervention and at follow-up. Due to these issues, Cohen's *d* effect sizes were unable to be calculated. Only two studies (Beaumont et al., 2015; Beaumont et al., 2021) aimed to address this issue by conducting analyses

using an intention to treat principle. Therefore, pre-intervention scores were assigned to post-intervention and follow-up scores or assigned the same score as the previous time point. These methods assume that the child made no improvement over time. Despite these attempts to mitigate the impact of missing data upon statistical analyses, findings from all studies within this review should be interpreted with caution. It should be noted that one study (Sofronoff et al., 2017) analysed attrition data and found an interaction with level of education, parental traits of autism and age.

Due to the issues with data collection, only one study (Beaumont et al., 2015) was able to analyse teacher reported data on social skills. The other five studies (Beaumont et al., 2021; Beaumont & Sofronoff, 2008; Einfeld et al., 2018; Sofronoff et al., 2017; Tan et al., 2015) relied heavily upon parent reported data on social skills. Positive gains in participant social skills post-intervention were on the parent version of the Social Skills Questionnaire, which highlights the potential issue of social desirability. Parents may answer the questionnaire in a more favourable light to report more desirable attributes, such as their children having better social skills. Therefore, the findings of these studies should be interpreted cautiously.

G\* Power calculation to work out the required sample size to determine a statistically significant effect was conducted in only two studies (Beaumont et al., 2021; Sofronoff et al., 2017). Due to participant attrition and incompleteness of outcome measures post-intervention and at follow-up, the study was under-powered according to their a-priori G\* Power analysis. Therefore, although these studies found a large and very large effect size we should be

cautious interpreting the effect sizes as they may not be truly representative of the social skill gains made.

**Table 7***Effect sizes for study outcomes relevant to the review question*

Study	Outcome Measure	Outcomes relevant to the review question	Effect Size	Descriptor	WoE D
Beaumont et al. (2015)	SSQ-P ( $n = 32$ )	Social skill gains in both conditions over time	$d = 1.00$	Large	1.97 (medium)
	SSQ-T ( $n = 34$ )	Social skill gains in both conditions over time	$d = 1.09$	Large	
			No difference in social skill gains between groups	N/A	
	SSQ-P ( $n = 33$ )	SAS pre-post intervention social skill gains	$d = 0.39$	Medium	
	SSQ-T	SAS pre-post intervention social skill gains	$d = 0.57$	Medium	
	SSQ-T	SAS follow-up social skill gains	$d = 0.03$	Very small	
Einfeld et al. (2018)	SSQ-P	Social skill gains in intervention ( $n = 21$ ) vs. wait-list ( $n = 36$ )	$d = 0.24$	Small	2.13 (medium)
Tan et al. (2015)	SSQ-P	Gains in social skills post-intervention	N/A		



Study	Outcome Measure	Outcomes relevant to the review question	Effect Size	Descriptor	WoE D
Beaumont et al. (2021)	SSQ-P	Post-intervention social skill gains for treatment group ( $n = 31$ ) vs. control ( $n = 33$ )	$d = 1.06$	Large	2.1 (medium)
	SSQ-P	Follow-up social skill gains for treatment group ( $n = 25$ ) vs. control ( $n = 32$ )	$d = 0.84$	Large	
	SSQ-P	Pre ( $n = 35$ ) to post ( $n = 31$ ) intervention social skill gains	$d = 2.73$	Very large	
Sofronoff et al. (2017)	SSQ-P ( $n = 54$ )	Social skill gains over time	$d = 1.34$	Very large	1.1 (low)
Beaumont & Sofronoff (2008)	SSQ-P ( $n = 26$ )	SAS pre-post intervention social skill gains	$d = 2.17$	Very large	1.83 (medium)

*Note.* SSQ-P and SSQ-T measures refer to the Social Skills Questionnaire Parent and Teacher versions respectively. Effect sizes are categorised according to the following categories (Cohen, 1988) – 0.2 (small), 0.5 (medium), 0.8 (large)

## 4.1 Conclusions

The aim of this review was to evaluate the effectiveness of the Secret Agent Society programme for improving the social skills of children with ASD. A systematic literature search yielded six studies that met the outlined inclusion criteria. The weight of evidence framework (Gough, 2007), was used to critique findings and provide an overall weight of evidence score. Four studies (Beaumont et al., 2021; Beaumont & Sofronoff, 2008; Einfeld et al., 2018; Tan et al., 2015) received a medium overall weight of evidence score. The remaining two studies (Beaumont et al., 2015; Sofronoff et al., 2017) received a low overall weight of evidence score. All studies reported gains in social skills after participating in the SAS programme, however conclusions were mainly drawn from parent reported data and should be interpreted cautiously. There was a large discrepancy between study approaches and quality of statistical analyses used to examine the results, making it difficult to make conclusions about the impact of the programme upon social skills in children with ASD. Moreover, due to effect sizes being incalculable for two studies, it wasn't possible to compare magnitude of social skill gains between studies.

Future research would benefit from independent researchers to eliminate the bias presented from the programme creator being involved in the research. Moreover, time of data collection should be considered to increase questionnaire completion. If age, education level and parental traits of autism impact being able to complete questionnaires then support for these individuals to complete questionnaires should be put in place. In addition, all but one study adapted the programme in some way to make delivery viable

which suggests that the programme may be intensive to implement within a school setting. It would be beneficial for a piece of qualitative work to be carried out to hear the experiences of facilitators in delivering the programme and to highlight any issues with programme implementation.

In summary, the current evidence base is insufficient to conclude that the SAS programme is an effective intervention for improving social skills in children with ASD.

## 5.1 References

- Anderson, S., & Morris, J. (2006). Cognitive Behaviour Therapy for People with Asperger Syndrome. *Behavioural and Cognitive Psychotherapy*, 34(3), 293-303. doi:10.1017/S1352465805002651
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*. 5th ed. Arlington, VA: American Psychiatric Association.
- Baron-Cohen, S., Leslie, A.M., and Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, 4, 37-46. doi:10.1016/0010-0277(85)90022-8
- Butterworth, T. W., Hodge, M., Sofronoff, K., Beaumont, R., Gray, K. M., Roberts, J., ... & Einfeld, S. L. (2014). Validation of the emotion regulation and social skills questionnaire for young people with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 44(7), 1535-1545. doi:10.1007/s10803-013-2014-5
- Chamberlain, B., Kasari, C., Rotheram-Fuller, E., (2007). Involvement or isolation? The social networks of children with autism in regular classrooms. *Journal of Autism and Developmental Disorders*, 37(2), 230-242. doi:10.1007/s10803-006-0164-4
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.).
- Ellopoulos, G. M., Harris, A. D., Lautenbach, E., & Perencevich, E. (2005). A Systematic Review of Quasi-Experimental Study Designs in the Fields

of Infection Control and Antibiotic Resistance. *Clinical Infectious Diseases*, 41(1), 77-82. doi:10.1086/430713

Frederickson, N., & Cline, T. (2015). *Special Educational Needs, Inclusion and Diversity: A Textbook* (3rd edition). Buckingham: OU Press.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional children*, 71(2), 149-164. doi:10.1177/001440290507100202

Gough, D. (2007). Weight of Evidence: A framework for the appraisal of the quality and relevance of evidence. In J.Furlong, A. Oancea (Eds.) *Applied and Practice-based Research. Special Edition of Research Papers in Education*, 22(2), 213-228. doi:10.1080/02671520701296189

GovUK – Explore education statistics (2021). Special educational needs in England. <https://explore-education-statistics.service.gov.uk/find-statistics/special-educational-needs-in-england>

Horner, R. H., Carr, E. G., Halle, J., Mcgee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165–179. doi:10.1177/001440290507100203

Lenhard, W., & Lenhard, A. (2016). *Calculation of Effect Sizes*. doi:10.13140/RG.2.2.17823.92329

- Matson, J.L., Turygin, N.C., Beighley, J., Rieske, R., Tureck, K., & Matson, M.L. (2012). Applied behavior analysis in autism spectrum disorders: Recent developments, strengths, and pitfalls. *Research in Autism Spectrum Disorders*, 6(1), 144-150. doi:10.1016/j.rasd.2011.03.014
- Mills, E. (2019). *How effective are technology-assisted, parent Naturalistic Developmental Behavioural Interventions for supporting the development of social-communication skills in young children with Autism Spectrum Disorder?* [UCL]. <https://www.ucl.ac.uk/educational-psychology/resources/CS1Mills18-21.docx.pdf>
- Pellicano, E., Maybery, M., Durkin, K., & Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: “Weak” central coherence and its relationship to theory of mind and executive control. *Development and Psychopathology*, 18(1), 77-98. doi:10.1017/S0954579406060056
- Pennington, B.F. and Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 37, 51-87. doi:10.1111/j.1469-7610.1996.tb01380.x
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health*, 57, 527–529. doi:10.1136/jech.57.7.527
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(4), 515-526. doi:10.1017/S0140525X00076512

- Rowley, E., Chandler, S., Baird, G., *et al.* (2012). The experience of friendship, victimization and bullying in children with an autism spectrum disorder: Associations with child characteristics and school placement. *Research in Autism Spectrum Disorders*. 6(3) 1126-1134. doi:10.1016/j.rasd.2012.03.004
- Schroeder, J.H., Cappadocia, M.C., Bebko, J.M. *et al.* (2014). Shedding Light on a Pervasive Problem: A Review of Research on Bullying Experiences Among Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-013-2011-8.
- The Social Skills Training Institute (2019). *Secret Agent Society Guide for Professionals*. <https://www.sst-institute.net/sites/sst-commercekickstart.dd/files/SAS%20Guide%20for%20Professionals%202019.pdf>
- Šimundić, A.-M. (2013). Bias in research Lessons in biostatistics. *Biochemia Medica*, 23(1), 12–17. doi:10.11613/BM.2013.003
- Spence, S. H. (1995). Social skills questionnaire. Social skills training: Enhancing social competence with children and adolescents: *Research and Technical Supplement*. Windsor: NFER-Nelson
- Symes, W. and Humphrey, N. (2010). Peer-group indicators of social inclusion among pupils with autism spectrum disorders (ASD) in mainstream secondary schools: A comparative study. *School Psychology International*. 31(5), 478-494. doi:10.1177/0143034310382496

Wood, H. (2021). *How effective is the Fun FRIENDS intervention at reducing anxiety outcomes in young children?* [UCL].

[https://www.ucl.ac.uk/educational-psychology/resources/CS1Wood\\_20-23.pdf](https://www.ucl.ac.uk/educational-psychology/resources/CS1Wood_20-23.pdf)



## 6.1 Appendices

### 6.2 Appendix A – Excluded Studies

*Studies excluded at abstract screening with exclusion criteria number(s)*

Full study reference	Exclusion criteria number(s)
Lee, V., Roudbarani, F., Tablon Modica, P., Pouyandeh, A., & Weiss, J. A. (2022). Adaptation of cognitive behavior therapy for autistic children during the pandemic: A mixed-methods program evaluation. <i>Evidence-Based Practice in Child and Adolescent Mental Health</i> , 7(1), 76-93. doi:10.1080/23794925.2021.1941432	1 (Intervention) 3 (Measures)
MacEville, D., & Brosnan, G. (2020). Adapting an emotional regulation and social communication skills group programme to teletherapy, in response to the COVID-19 pandemic. <i>Irish Journal of Psychological Medicine</i> , 1-6. doi:10.1017/ipm.2020.109	3 (Measures) 4 (Study design)
Maughan, A. L., & Weiss, J. A. (2017). Parental outcomes following participation in cognitive behavior therapy for children with autism spectrum disorder. <i>Journal of Autism and Developmental Disorders</i> , 47(10), 3166-3179. doi:10.1007/s10803-017-3224-z	3 (Measures)

*Studies excluded at full-text screening with exclusion criteria number(s)*

Full study reference	Exclusion reason
<p>Beaumont, R. B., Pearson, R., &amp; Sofronoff, K. (2019). A novel intervention for child peer relationship difficulties: The Secret Agent Society. <i>Journal of Child and Family Studies</i>, 28(11), 3075-3090. doi:10.1007/s10826-019-01485-7</p>	2 (Participants)
<p>Beaumont, R. B., Smith-Merry, J., Costley, D., Howlin, P., Sofronoff, K., Roberts, J., ... &amp; Einfeld, S. L. (2019). Implementation, Evaluation and Maintenance of a Social-Emotional Skills Training Program for Children with an Autism Spectrum Disorder in a Specialist School Setting. <i>International journal of special education</i>, 34(1), 95-108.</p>	3 (Measures)
<p>Costley, D., Baldwin, S., Clark, T., Howlin, P., Taffe, J., Beaumont, R., . . . Sofronoff, K. (2020). The Association Between Parent Engagement and Child Outcomes in Social Skills Training Programs: Discovering the Secret Agent Society in Partnership. <i>Australasian Journal of Special and Inclusive Education</i>, 44(1), 46-59. doi:10.1017/jsi.2020.2</p>	3 (Measures)
<p>Cullen, A. M. Emotions and Social Problem Solving: Using the Secret Agent Society Computer Game with a Pupil with an Autism Spectrum Disorder in an Irish Mainstream School Setting. In <i>LEARN</i> (p. 86).</p>	3 (Measures) 4 (Study design)
<p>Sauvé, J. S., O’Haire, C., Hall, H., Lane, C., &amp; Hudson, B. O. (2018). Adapting a social skills intervention for children with Autism within an Urban specialty</p>	2 (Participants)

Full study reference	Exclusion reason
<p>community clinic. <i>Evidence-Based Practice in Child and Adolescent Mental Health</i>, 3(4), 219-235. doi:10.1080/23794925.2018.1483214</p>	
<p>Tan, Y. L., Mazzucchelli, T. G., &amp; Beaumont, R. (2015). An evaluation of individually delivered secret agent society social skills program for children with high-functioning autism spectrum disorders: A pilot study. <i>Behaviour Change</i>, 32(3), 159–174. doi:10.1017/bec.2015.7</p>	4 (Study design)
<p>Thomson, K., Burnham Riosa, P., &amp; Weiss, J. A. (2015). Brief report of preliminary outcomes of an emotion regulation intervention for children with autism spectrum disorder. <i>Journal of Autism and Developmental Disorders</i>, 45(11), 3487-3495. doi:10.1007/s10803-015-2446-1</p>	1 (Intervention) 3 (Measures)
<p>Weiss, J. A., Thomson, K., Burnham Riosa, P., Albaum, C., Chan, V., Maughan, A., ... &amp; Black, K. (2018). A randomized waitlist-controlled trial of cognitive behavior therapy to improve emotion regulation in children with autism. <i>Journal of Child Psychology and Psychiatry</i>, 59(11), 1180-1191. doi:10.1111/jcpp.12915</p>	1 (Intervention)

## 6.3 Appendix B – Weight of Evidence (WoE)

### 6.3.1 Weight of Evidence (WoE) A – Methodological Quality

Two coding protocols, namely Gersten et al. (2005) and Horner et al. (2005) coding protocols were used to weigh the evidence of the studies included in this systematic literature review. The coding protocols were used to assess the methodological quality of two randomised controlled trials, two quasi-experimental between subject studies, a quasi-experimental within subject design and a single case experimental study.

Each randomised controlled study and quasi-experimental study was given a WoE A rating based upon the Gersten et al. (2005) coding protocol for group experimental and quasi-experimental research. Criteria to base WoE A rankings upon were adapted, as per Wood (2021), to appraise studies where they have met fewer essential criteria but more desirable criteria. Specified criteria can be found outlined in Table 10 below.

**Table 8**

*WoE A Criteria using the Gersten et al. (2005) coding protocol*

WoE A Rating	Criteria
3 (high)	Study meets at least 9 essential criteria and at least 4 desirable criteria
2 (medium)	Study meets at least 9 essential criteria and less than 4 desirable criteria OR Study meets 7 to 8 essential criteria and at least 4 desirable criteria

WoE A Rating	Criteria
1 (low)	Study meets 7 to 8 essential criteria and less than 4 desirable criteria OR Study meets less than 7 essential criteria and at least 4 desirable criteria
0 (very low)	Study meets less than 7 essential criteria and less than 4 desirable criteria

*Note.* If there is insufficient information to be able to code a specific criterion it will be coded as not meeting the criteria

Ratings assigned to studies based on these criteria are shown in Table 9.

**Table 9**

*WoE A Ratings for each study based on Gersten et al. (2005) coding protocol*

Study	Essential Criteria	Desirable Criteria	WoE A rating
Beaumont et al. (2015)	5	5	1 (low)
Einfeld et al. (2018)	5	5	1 (low)
Beaumont et al. (2021)	7	5	2 (medium)
Sofronoff et al. (2017)	6	4	1 (low)
Beaumont & Sofronoff (2008)	7	3	1 (low)

A coding matrix including the completed responses for the Gersten et al. (2005) coding protocol of each included study can be found in Appendix C.

In order to appraise the remaining study with a single case experimental design, the Horner et al. (2005) coding protocol was used. The WoE A

criteria for Horner et al. (2005) are outlined in Table 10 as defined in Mills (2019).

**Table 10***Criteria for WoE A using Horner et al. (2005) coding protocol*

Section	Score	Criteria
A	3	All criteria are satisfied
	2	Two criteria are satisfied
	1	One criterion is satisfied
	0	No criteria are satisfied
B	3	All criteria are satisfied
	2	Three or four criteria are satisfied
	1	One or two criteria are satisfied
	0	No criteria are satisfied
C	3	All criteria are satisfied
	2	Two criteria are satisfied
	1	One criterion is satisfied
	0	No criteria are satisfied
D	3	All criteria are satisfied
	2	Two criteria are satisfied
	1	One criterion is satisfied
	0	No criteria are satisfied
E	3	All criteria are satisfied
	2	Two criteria are satisfied
	1	One criterion is satisfied
	0	No criteria are satisfied
F	3	Experimental effects replicated across 3+ participants and in a unique setting
	2	Experimental effects replicated across 3+ participants
	1	Experimental effects are replicated across 2 participants
	0	Experimental effects are replicated with 1 or no participants
G	3	All criteria are satisfied
	2	Two or three criteria are satisfied
	1	One criterion is satisfied
	0	No criteria are satisfied
<b>Overall WoE A Rating</b>		Sum of A-G scores divided by the average (all 7 sections)

**Table 11***WoE A Criteria using the Horner et al. (2005) coding protocol*

WoE A Rating	Criteria
3 (high)	Average score of 2 to 3 across all 7 judgement areas
2 (medium)	Average score of 1 to 1.9 across all 7 judgement areas
1 (low)	Average score of 0 to 0.9 across all 7 judgement areas

Ratings assigned to this study are based on the criteria outlined in Table 12.

**Table 12***WoE A ratings based on Horner et al. (2005) coding protocol*

Study	Dimensions							Overall Score	WoE Ranking
	A	B	C	D	E	F	G		
Tan et al. (2015)	2	2	0	1	0	2	2	1.29	2 (Medium)

Examples of completed Gersten et al. (2005) and Horner et al. (2005) coding protocols can be found in Appendix D.

### 6.3.2 Weight of Evidence (WoE) B – Relevance of Methodology

The review question of this systematic literature review is regarding the effectiveness of the Secret Agent Society programme. According to Petticrew and Roberts (2003), randomised controlled trials are the best study design for answering this type of question, followed by quasi-experimental designs. The studies which used a quasi-experimental design, were weighted according to whether they used a control group, measures taken and



participant assignment to control groups. According to Ellopoulos et al. (2005), quasi-experimental designs with a control group are the most robust. In order to assess the effectiveness of the Secret Agent Society programme a third criterion was created to identify and assess the influence of bias upon the study. As the intervention creator was a researcher in all of these studies and were funded by different services, the studies were weighted differently according to whether a conflict of interest was disclosed. The criteria in which each study was weighted against are described in Table 13. Final ratings and an overall WoE B score are outlined in Table 14.

**Table 13**

*WoE B Criteria with rationale*

Criteria	Weighting	Rationale
A. Study Design	3 Randomised controlled trials	Petticrew and Roberts' (2003) typology of evidence adapted from Gray (1997), states that randomised controlled trials are the most appropriate type of study for an 'effectiveness' review question. This is followed by quasi-experimental and cohort studies. Quasi-
	2 Quasi-experimental study designs with a control group and/or non-random assignment to intervention or control	experimental study designs with a control group are the highest quality quasi-experimental designs and thus weighted higher
	1 Quasi-experimental designs with no control group and/or only pre and post measures, cohort studies and single case experimental designs	

Criteria	Weighting	Rationale
	0	Qualitative research, surveys, case studies or non-experimental evaluations
B. Research author(s)	3	Entirely Independent Researchers
	2	Researchers have disclosed or addressed a conflict of interest in the programme, e.g. author is creator and may receive financial contribution
	1	Researchers have provided information that suggests a possible conflict of interest but have not directly disclosed or addressed this
	0	Conflict of interest is not disclosed or addressed at all

**Table 14***WoE B Ratings*

Study	Criteria A	Criteria B	WoE B Rating
Beaumont et al. (2015)	2	1	1.5 (low)
Einfeld et al. (2018)	2	2	2 (medium)
Tan et al. (2015)	1	2	1.5 (low)
Beaumont et al. (2021)	3	2	2.5 (high)
Sofronoff et al. (2017)	1	2	1.5 (low)
Beaumont & Sofronoff (2008)	3	2	2.5 (high)

*Note.* 1 - 1.6 (low), 1.7 – 2.3 (medium), 2.4 – 3 (high)

### 6.3.3 Weight of Evidence (WoE) C – Topic Relevance

WoE C criteria were developed according to the topic relevance to the review question. Therefore, the criteria in Table 15 were created and each study received a 0-3 rating based on their average score across these 5 criteria. Due to the programme being examined as a school-based intervention for children with ASD, it was felt that intervention fidelity, participant diagnosis, outcome measures and applicability to school settings within the UK were particularly salient features to be explored.

**Table 15***WoE C Criteria with rationale*

Criteria	Scoring	Rationale	
A. Intervention fidelity:	3	Study meets all 4 criteria	To assess the effectiveness of the Secret Agent Society Programme
• Sessions follow the Secret Agent Society programme according to the original manual	2	Study meets 3 criteria	Studies are more relevant if they are implemented as per the original intervention manual
• Skill practise between sessions	1	Study meets 2 to 1 criterion	
• Delivered over 9 90-minute sessions or 18 45-minute sessions	0	Study meets no criteria	
• Facilitators trained to deliver the intervention			
B. Participants	3	Confirmed clinical diagnosis of ASD with a required cognitive ability score from a standardised assessment for recruitment. Standardised assessment used and score cut-offs should be stated	To ensure all participants were able to access the SAS programme content

Criteria	Scoring	Rationale
	2	Confirmed clinical diagnosis of ASD with a required cognitive ability score from a standardised assessment for recruitment. Standardised assessment and/or score cut-offs are not stated
	1	Confirmed clinical diagnosis of ASD without a required cognitive ability score for recruitment
	0	No information about ASD diagnosis or cognitive ability
C. Outcome measures	3	Outcome measures include a parent and teacher measure of their child's social skills with established reliability AND an ecological measure of social skill generalisation, e.g., playground observation of interaction with peers
	2	Outcome measures include a parent and teacher measure of their

Criteria	Scoring	Rationale
	child's social skills with established reliability	
	1 Outcome measures include either a parent OR teacher measure of their child's social skills with established reliability	
	0 Outcome measures include a measure of social skills without established reliability	
D. Application to school setting	3 Programme delivered in school setting by teaching staff	Studies where the Secret Agent Society Programme has been delivered in school settings are more relevant to this review
	2 Programme delivered in school setting by non-teaching trained facilitator	
	1 Programme delivered outside of school setting by trained facilitator	
	0 Programme delivered outside of school setting by untrained facilitator	

Criteria	Scoring	Rationale
E. Country	3 Study conducted in the UK	To explore the generalisability of study findings to the UK
	2 Study conducted in another OECD country	Studies conducted in an OECD country are more likely to have similarities to the UK education system
	1 Study not conducted in an OECD country	
	0 No information about study information	

**Table 16**

*WoE C Ratings*

Study	Criteria A	Criteria B	Criteria C	Criteria D	Criteria E	WoE C Rating
Beaumont et al. (2015)	2	3	2	3	2	2.4 (high)
Einfeld et al. (2018)	3	3	2	3	2	2.6 (high)
Tan et al. (2015)	2	3	2	1	2	2 (medium)
Beaumont et al. (2021)	1	1	2	1	2	1.4 (low)
Sofronoff et al. (2017)	2	2	1	1	2	1.6 (low)
Beaumont & Sofronoff (2008)	2	3	2	1	2	2.4 (high)

*Note.* 1 - 1.6 (low), 1.7 – 2.3 (medium), 2.4 – 3 (high)

## 6.4. Appendix C – Coding Matrix

**Table 17**

*Coding matrix to display methodological quality of included studies according to Gersten et al. (2005) coding protocol*

	Studies				
	Beaumont et al. (2015)	Einfeld et al. (2018)	Beaumont et al. (2021)	Sofronoff et al. (2017)	Beaumont & Sofronoff (2008)
<b><u>Essential Quality Indicators</u></b>					
<b>Describing participants</b>					
Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?	Yes	Yes	No	Yes	Yes
Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?	Yes	No	Yes	N/A	Yes
Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?	No	No	Yes	Yes	No



<b>Implementation of the intervention and description of comparison conditions</b>					
Was the intervention clearly described and specified?	No	No	Yes	Yes	Yes
Was the fidelity of implementation described and assessed?	No	Yes	Yes	No	Yes
Was the nature of services provided in comparison conditions described?	Yes	Yes	Yes	N/A	No
<b>Outcome Measures</b>					
Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?	No	No	No	No	No
Were outcomes for capturing the intervention's effect measured at the appropriate times?	Yes	Yes	No	Yes	Yes
<b>Data Analysis</b>					
Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study? Was a power analysis provided to describe the required sample size?	No	No	Yes	Yes	Yes

Did the research report include not only inferential statistics but also effect size calculations?	Yes	Yes	Yes	Yes	Yes
<b><i>Desirable Quality Indicators</i></b>					
Was data available on attrition rates among intervention samples? If so, was severe overall attrition documented?	Yes	Yes	No	No	Unknown
Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures?	Yes	No	No	Yes	No
Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?	Unknown	Yes	No	Unknown	Unknown
Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?	Yes	Yes	Yes	Yes	Yes
Was evidence of the criterion-related validity and construct validity of the measures provided?	Yes	No	Yes	No	No
Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist	No	Unknown	Unknown	No	No

following procedures specified), but also examine quality of implementation?

Was the nature of instruction or series documented in the comparison conditions?	No	No	Yes	N/A	No
Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?	No	Yes	Unknown	Unknown	Yes
Were results presented in a clear, coherent fashion?	Yes	Yes	Yes	Yes	Yes

---

## 6.5. Appendix D – Completed Coding Protocols

**Coding protocol:** Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C. & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional children*, 71(2), 149-164. doi:10.1177/001440290507100202

**Name of coder:**

**Date:** 21.01.2022

**Full study reference:** Beaumont, R., Rotolone, C., & Sofronoff, K. (2015). The secret agent society social skills program for children with high-functioning autism spectrum disorders: A comparison of two school variants. *Psychology in the Schools*, 52(4), 390–402. doi:10.1002/PITS.21831

**Research design:** Quasi-experimental design

**Type of publication:** Journal article

### Essential Quality Indicators

#### *Describing Participants*

Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?

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

Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?

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

Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?

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

### ***Implementation of the Intervention and Description of Comparison Conditions***

Was the intervention clearly described and specified?

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

Was the fidelity of implementation described and assessed?

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

Was the nature of services provided in comparison conditions described?

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

### ***Outcome Measures***

Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?

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

Were outcomes for capturing the intervention's effect measured at the appropriate times?

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

### ***Data Analysis***

Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?

- Yes
- No

- N/A
- Unknown/Unable to Code

Did the research report include not only inferential statistics but also effect size calculations?

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

***Desirable Quality Indicators***

Was data available on attrition rates among intervention samples? Was severe overall attrition documented?

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

Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures?

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

Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?

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

Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

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

Was evidence of the criterion-related validity and construct validity of the measures provided?

- Yes

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

Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

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

Was the nature of instruction or series documented in the comparison conditions?

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

Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

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

Were results presented in a clear, coherent fashion?

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

Quality Indicators	Total Indicators
Essential Quality Indicators	6
Desirable Quality Indicators	5

**Coding protocol:** Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S. & Wolery, M. (2005). The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education. *Exceptional Children*, 71(2), 165-179. doi:10.1177/001440290507100203

**Name of coder:**

**Date:** 21.01.22

**Full study reference:** Tan, Y. L., Mazzucchelli, T. G., & Beaumont, R. (2015). An evaluation of individually delivered secret agent society social skills program for children with high-functioning autism spectrum disorders: A pilot study. *Behaviour Change*, 32(3), 159–174. doi:10.1017/bec.2015.7

**Research design:** Single case study design

**Type of publication:** Journal article

<b>A. Description of participants</b>	<b>Yes/No</b>
Participants are described with sufficient detail to allow others to select individuals with similar characteristics (e.g. age, gender, disability, diagnosis)	Yes
The process for selecting participants is described with replicable precision	Yes
Critical features of the physical setting are described with sufficient precision to allow replication	No
<b>B. Dependent Variable</b>	
Dependent variables are described with operational precision	No
Each dependent variable is measured with a procedure that generates a quantifiable index	Yes
Measurement of the dependent variable is valid and described with replicable precision	Yes
Dependent variables are measured repeatedly over time	Yes
Data are collected on the reliability or interobserver agreement associated with each dependent variable, and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa = 60%)	No



<b>C. Independent Variable</b>	
Independent variable is described with replicable precision	No
Independent variable is systematically manipulated and under the control of the experimenter	No
Overt measurement of the fidelity of the implementation for the independent variable is highly desirable	No
<b>D. Baseline</b>	
The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable	Yes
Establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur	No
Baseline conditions are described with replicable precision	No
<b>E. Experimental Control / Internal Validity</b>	
The design provides at least three demonstrations of experimental effect at three different points in time	No
The design controls for common threats to internal validity (e.g. permits elimination of rival hypotheses)	No
The results document a pattern that demonstrates experimental control	No
<b>F. External Validity</b>	
Experimental effects are replicated across participants, settings, or materials to establish external validity	Yes
<b>G. Social Validity</b>	
The dependent variable is socially important	Yes
The magnitude of change in the dependent variable resulting from the intervention is socially important	Yes

Implementation of the independent variable is practical and cost effective No

Social validity is enhanced by the implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts No

---

<b>Total</b>	9
Average across 7 judgement areas	1.29
<b>Weight of Evidence (WoE) A Score</b>	<b>2</b>

---

WoE A Score 3: Average score of 2 – 3 across the seven judgement areas; WoE A Score 2: Average score of 1 – 1.9 across the seven judgement areas; WoE A Score 1: Average score of 0 – 0.9 across the seven judgement areas