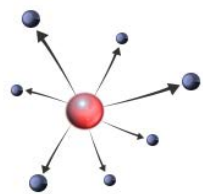

Guideline Supplementary Paper

New Zealand Autism Spectrum Disorder Guideline supplementary paper on
social skills groups for children and young people with ASD



INSiGHT
Research Ltd

Independent Network of Specialists in Guidelines & Health Technology

With the support of the New Zealand Autism Spectrum Disorder
Living Guideline Group

1 March 2016

© Ministry of Health 2016

Published by: INSIGHT Research Ltd
181 Blighs Rd, Strowan 8052, New Zealand

ISBN (Electronic): 978-0-947491-72-7

Copyright



This work is licensed under the Creative Commons Attribution 4.0 International licence. In essence, you are free to: share ie, copy and redistribute the material in any medium or format; adapt ie, remix, transform and build upon the material. You must give appropriate credit, provide a link to the licence and indicate if changes were made.

Funding and independence

This work was funded by the New Zealand Ministry of Health and sponsored by the New Zealand Ministry of Education.

The work was researched and written by INSIGHT Research Ltd employees or contractors. Appraisal of the evidence, formulation of recommendations and reporting are independent of the Ministries of Health and Education.

Statement of intent

INSIGHT Research produces evidence-based best practice guidelines, health technology assessments and literature reviews to help health care practitioners, policy-makers and consumers make decisions about health care in specific clinical circumstances. The evidence is developed from systematic reviews of international literature and placed within the New Zealand context.

Guidelines, including supplementary papers, are not intended to replace the health practitioner's judgment in each individual case.

Suggested citation

Marita Broadstock. New Zealand Autism Spectrum Disorder Guideline supplementary paper on social skills groups for children and young people with ASD. Christchurch: INSIGHT Research; 2016.

Currency review date: 2020

HP6364

Contents

List of Tables	vi
About the evidence review	vii
Purpose.....	vii
Scope of the evidence review.....	vii
Definitions	viii
Target audience	ix
Treaty of Waitangi	ix
Recommendation development process.....	ix
Summary	x
Summary of new recommendations	x
1 Introduction	1
1.1 Social skills in ASD.....	1
Social challenges in ASD	1
Social skills training.....	2
Social skills groups.....	3
1.2 Recommendations relating to social skills groups in the NZ ASD Guideline	3
1.3 Objectives of the current review update.....	4
2 Social skills groups in children and young people with ASD	5
2.1 Scope and methods	5
Research question	5
Sample.....	5
Study designs	5
Intervention/exposure.....	5
Comparator	6
Outcomes	6
Identification and selection of studies for inclusion	6
Critical appraisal of included studies	8
2.2 Body of evidence.....	8
Included studies	8
Systematic reviews	8
Primary studies	9
Study location and setting.....	9
Study design.....	9
Participants.....	9
Interventions	10
Outcomes assessed	11
Social competence.....	11
Quality of life	11
Specific skills or sub-domains of social competence	13
Blinding.....	14
2.3 Quality of included studies.....	14
2.4 Narrative appraisal of studies	19
Systematic reviews	19
Reichow et al (2013).....	19

Kaat & Lecavalier (2014)	19
Primary studies	20
PEERS and CFT studies	20
Laugeson et al (2009)	20
Frankel et al (2010)	21
Gantman et al (2012)	23
Schohl et al (2014)	24
Yoo et al (2014).....	25
Intensive Skillstreaming	26
Lopata et al (2010)	26
Thomeer et al (2012).....	27
Multimodal social skills group	27
Beaumont & Sofronoff (2008).....	27
Peer-mediated social skills group	29
Koenig et al (2010).....	29
Affection-focussed social skills group.....	30
Andrews et al (2013)	30
2.5 Synthesis of results	31
Systematic reviews	31
Social competence	33
Quality of life.....	33
Social communication.....	35
Problem behaviours.....	35
Emotion recognition	35
Emotion management.....	37
Empathy	37
Affection	37
Quality of play/interaction.....	37
Programme specific skills/knowledge.....	37
Maintenance	39
2.6 Limitations and future research directions	39
Introduction	39
Sample size	40
Sample characteristics and recruitment.....	40
Diagnosis of ASD.....	42
Control groups	42
Outcome measures.....	42
Assessment and maintenance	43
Study design	44
Moderators and mediators of treatment response	44
Programme standardisation and fidelity	44
Programme intensity.....	45
Programme components.....	45
Subject characterisation.....	45
Additional strategies	46
Variability in results.....	46
2.7 Summary and conclusions	47
Overview	47

Summary of main findings.....	48
Social competence	48
Quality of life.....	49
Social communication.....	49
Problem behaviours.....	49
Emotion recognition, management and empathy	49
Affection	49
Quality of play.....	50
Programme specific skills/knowledge.....	50
Maintenance.....	50
Systematic reviews	50
Conclusions	51
3 Recommendation development	53
Preamble.....	53
Revision of existing recommendations	53
New recommendation	53
Appendix 1: Methods	52
A1.1 Contributors	52
Living Guideline Group members	52
Ex-officio LGG members	52
INSIGHT Research	53
Declarations of competing interest	53
Acknowledgements	53
A1.2 Review scope.....	53
A1.3 Research question	54
A1.4 Search strategy.....	54
Search databases	54
A1.5 Appraisal of studies.....	55
Assigning a level of evidence	55
Appraising the quality of included studies.....	56
Completing evidence tables.....	56
Appraisal of primary and secondary studies.....	56
A1.6 Preparing recommendations	57
Developing recommendations	57
A1.7 Consultation	58
Appendix 2: Abbreviations and glossary	60
A2.1 Abbreviations and acronyms	60
Miscellaneous Terms	60
Tests, scales and measures.....	61
Databases.....	62
A2.2 Glossary.....	62
Epidemiological and statistical terms.....	62
Topic Specific Terms.....	66
Appendix 3: Evidence Tables of included studies.....	67
References.....	86

List of Tables

New recommendation relevant to social skills groups in children and young people with ASD	x
New good practice points relevant to social skills groups in children and young people with ASD	x
Table 1.1: Recommendations relevant to social skills groups in children and young people in the NZ ASD Guideline.....	4
Table 2.1: Inclusion and exclusion criteria for selection of studies	7
Table 2.2: Characteristics and results of included primary studies organised by type of social skills group	16
Table 2.3: Characteristics and results of included primary studies for social competence outcomes.	32
Table 2.4: Characteristics and results of included primary studies for quality of life outcomes.....	34
Table 2.5: Characteristics and results of included primary studies for social communication outcomes.....	34
Table 2.6: Characteristics and results of included primary studies for problem behaviour outcomes.	36
Table 2.7: Characteristics and results of included primary studies for emotional recognition, emotion management, empathy, and affection.....	36
Table 2.8: Characteristics and results of included primary studies for quality of play during get-togethers.	38
Table 2.9: Characteristics and results of included primary studies for knowledge of programme-specific skills.	38
Table 3.1: New recommendations relevant to social skills groups in children and young people with ASD.....	55
Table 3.2: New good practice points relevant to social skills groups in children and young people with ASD.	55
Table A1.2: Guide to grading recommendations	58

About the evidence review

Purpose

The New Zealand Autism Spectrum Disorder Guideline (the ASD Guideline) [1] was published in April 2008. As part of their commitment to the implementation of the guideline, New Zealand's Ministry of Health and Ministry of Education agreed to establish a Living Guideline process in 2009. This process is where a guideline is regularly updated and refined to reflect new evidence and changing user needs.

Updates within the living guideline process are required when the recommendations in the guideline are no longer considered valid in view of research evidence that has emerged since the guideline's literature searches were conducted. A multidisciplinary team from the Living Guideline Group (LGG), an advisory group responsible for identification of areas for update, consideration of new evidence and reporting on any implications for guideline recommendations.

This supplementary report describes a systematic review which aims to provide an evidence-based synthesis of research published in or beyond 2004 relating to group-based social skills training for children and young people with Autism Spectrum Disorder (ASD). This review updates the evidence considered in the ASD Guideline [1]. Also reported are revised and new recommendations pertinent to the topic developed by the Living Guideline Group following their consideration of the systematic review.

The systematic review was undertaken by INSIGHT Research to support the work of the New Zealand Autism Spectrum Disorder Guideline's Living Guideline Group. The methodology followed is consistent with that undertaken for previous supplementary reports of the LGG which were developed and conducted by the New Zealand Guideline Group [2-6].

The systematic review and the entire living guideline process was funded by the New Zealand Ministry of Health, and sponsored by the New Zealand Ministry of Education.

Scope of the evidence review

The current review aims to systematically update the evidence relating to the effectiveness of social skills groups for children and young people (aged 6-21 years) with Autism Spectrum Disorder (ASD). Eligible studies were randomised controlled trials. The evaluated intervention was group-based social skills training; that is, a series of group-based training sessions which focus on developing social skills for young people with ASD. Comparison groups were wait list controls, or those receiving usual care and/or no treatment. Eligible outcomes included social competence, social communication, emotion recognition, quality of social interaction/play, problem

behaviour, and observed specific behaviours; quality of life (including anxiety and depression); and programme knowledge. The Living Guideline Group identified this area as worthy of updating and one which could lead to revised or additional recommendations in the ASD Guideline [1].

This document needs to be read in the context of the original New Zealand ASD Guideline [1].

Definitions

Autism Spectrum Disorder (ASD) is a condition that affects communication, social interaction and adaptive behaviour functioning. In the DSM-5 [7], four pervasive developmental disorder subcategories specified in the diagnostic manual's predecessor, the DSM-IV [8], are now subsumed into one broad category of autism spectrum disorder. These subtypes are autistic disorder, Asperger's disorder (Asperger syndrome), childhood disintegrative disorder (CDD), and pervasive developmental disorder not otherwise specified (PDD-NOS). The name pervasive developmental disorder (PDD) has now been changed to Autism Spectrum Disorder (ASD), the term used throughout the original New Zealand ASD Guideline [1].

The diverse range of disability and intellectual function expressed by people across the autism spectrum requires that a wide range of services and approaches be employed to reflect the heterogeneity of the condition.

In this review unless otherwise stated, *social skills groups* refer to interventions which provide structured sessions in social skills training in small groups of people of a similar age group and with similar social problems. A session typically includes teaching a specific skill, demonstration of the skill through role playing, practice of the skill, and individualised feedback. Groups meet on a regular basis, typically for 1-2 hours, for several weeks, facilitated by at least one trained instructor/therapist. Parents are typically provided training in concurrent sessions to encourage their children to practice newly learned skills at home.

Whilst it is acknowledged that the term "high functioning" is not universally favoured in the context of autism, in this report, the term "high functioning" is used to refer to people with higher cognitive functioning either as established by intelligence tests (generally indicated by full IQ scores of 70 or above), or through the diagnosis of "high-functioning autism" or Asperger syndrome (under DSM-IV criteria). In light of the removal of Asperger syndrome as a separate diagnostic classification in DSM-5, these distinctions may no longer be used clinically.

Target audience

This evidence review and guidance update is intended primarily for the providers of professional health and education services for New Zealanders with ASD. It is also expected that the recommendations will be accessed by people with ASD and their families.

Treaty of Waitangi

INSIGHT Research acknowledges the importance of the Treaty of Waitangi to New Zealand, and considers the Treaty principles of partnership, participation and protection as central to improving Māori health.

INSIGHT Research's commitment to improving Māori health outcomes means we attempt to identify points in the guideline or evidence review process where Māori health must be considered and addressed. In addition, it is expected that Māori health is considered at all points in the guideline or evidence review in a less explicit manner.

Recommendation development process

The research questions were identified and prioritised by the Living Guideline Group and were used to inform the search of the published evidence. A one day, face-to-face meeting of the Living Guideline Group was held on 11 November 2014 where evidence was reviewed and existing recommendations revised and/or new recommendations developed.

INSIGHT Research follows specific structured processes for evidence synthesis. Full methodological details are provided in **Appendix 1**. This appendix also includes details of the Living Guideline Group membership and lists the organisations that provided feedback during the consultation period. **Appendix 2** presents a glossary of key epidemiological and topic-specific terms, abbreviations and acronyms. **Appendix 3** presents full evidence tables of included studies.

Summary

Summary of new recommendations

New recommendation relevant to social skills groups in children and young people with ASD

Reference	New recommendation	Grade
4.2.1a	Facilitated and structured social skills groups should be considered for high functioning children and young people with ASD.	B

New good practice points relevant to social skills groups in children and young people with ASD

New Reference	New Good Practice Points	Grade
4.2.1b	Social skills groups approaches in New Zealand need to be responsive to the cultural and linguistic diversity of the group participants.	✓
4.2.1c	Decisions about participating in social skills groups should be guided by whether a person with ASD values it, and whether they are expected to benefit from it.	✓

1 Introduction

1.1 Social skills in ASD

Social challenges in ASD

Difficulties in social interaction are a defining feature of autism spectrum disorder (ASD) and an area of significant vulnerability, regardless of language, cognitive ability or age. Impairments relevant to social functioning are diverse and wideranging, affecting speech, linguistic conventions and interpersonal interactions [9].

Whilst there is variability in presentation of social deficits in individuals on the spectrum, features of social function that can be affected in ASD include [9-14]:

- poor initiation of social contact
- difficulties interpreting verbal and nonverbal social cues, including gaze, gesture and posture
- impairment in the ability to identify faces, and read facial expressions and emotion
- difficulty understanding and expressing emotions and affection appropriately
- a tendency to focus on particular topics of interest in speech and attention
- poor interpretation of contextual elements of the social environment
- poor insight into the emotional components of relationships
- lack of ability in taking the perspective of the other (theory of mind)
- challenges with speech prosody (variation of pitch and inflection)
- challenges in understanding the social pragmatics of communication (following social rules of conversations such as turn taking, taking the listener's perspective, and nuances of language)
- difficulty interpreting nonliteral language such as sarcasm, metaphor, and irony
- impaired ability to regulate affect and self monitor behaviour.

Heightened sensory sensitivity, and comorbid conditions such as obsessive-compulsive disorder and anxiety, may also contribute to avoidance of social contact [10].

Young people on the spectrum can be at greater risk of social isolation, teasing, bullying, peer victimisation, lower self-esteem, and dropping out of school [15, 16]. Social skills deficits and peer rejection can negatively impact quality of life, and can contribute to problems in peer acceptance, romantic relationships, daily living, and

academic and vocational outcomes [17, 18]. Mental health problems including anxiety and depression, which are commonly comorbid in ASD, may also be exacerbated [16].

For many people with ASD, social deficits are not explained by a lack of social interest [9] and they can be painfully aware of their own social difficulties. Social problems can be heightened in adolescence and early adulthood when friendship networks with peers and “fitting in” are particularly important. Analysis of a large US longitudinal cohort study of adolescents enrolled in special education found that adolescents with ASD were significantly more likely to never see friends out of school (43%), never to get called by friends (54%), and never to be invited to social activities (50%) when compared with adolescents with other disabilities [19].

Being at the higher functioning end of the autism spectrum can introduce additional challenges for people with ASD as their greater independence and mainstream schooling can open them up to more social contact with neurotypical peers [20].

Social skills training

Social skills training uses behavioural and social learning techniques to teach social skills necessary to build and foster relationships with others [17]. The primary outcome of social skills training is social competence. Social competence is about being able to manage and contribute to the social interactions we have. It has been defined as the ability to achieve personal goals in social interaction while simultaneously maintaining positive relationships with others over time and across situations [21]. Being socially competent involves many elements, including: the ability to regulate emotions, knowledge and experience of social interactions, and understanding social situations and customs [22].

Teaching techniques used in social skills training typically include direct instruction, modeling, role-playing, shaping, feedback, and reinforcement interactions [17]. There are a range of modalities of social skills training, including applied behaviour analysis-based approaches, cognitive behavioural components, naturalistic techniques (particularly for young children), parent training, peer training, visual supports (e.g., social stories, scripts and visual activity training), video modeling, individualised clinician training, and the subject of the current review, *social skills groups* [23].

Social skills training can target specific behaviours that can be taught, modelled and practiced in training (eg., introducing oneself), as well as generalised improvements in social behaviour that may not be specifically targeted (eg., staying calm when teased) [24]. Social skills training programmes can also focus on social areas such as the development of friendships [25] and showing and responding appropriately to signs of affection [10].

Social skills groups

Development and evaluation of social skills groups is a burgeoning area of intervention and research in autism, particularly where offered to children and adolescents with average or above average cognitive and verbal skills.

Social skills groups for people on the autism spectrum allow participants to learn, practice and enhance social skills among other young people in a supportive environment. Groups include a small number of people with ASD of a similar age group but may be supplemented by non-autistic peers who model appropriate social skills and provide opportunities for practising these skills. Groups meet on a regular basis, typically for 60-90 minutes, for several weeks, facilitated by at least one trained instructor/therapist [26].

Social skills groups tend to be more structured than informal social groups. A session typically includes teaching a specific skill, demonstration of the skill through role playing, practice of the skill, and individualised feedback. Based on behavioural principles, this process permits immediate reinforcement of targeted skills which is theorised to increase the likelihood of the skill being used again [26]. Some group approaches also provide training for parents to encourage their children to practice newly learned skills at home as directed through set homework tasks.

1.2 Recommendations relating to social skills groups in the NZ ASD Guideline

Social skills groups are referred to explicitly twice in the New Zealand ASD Guideline [1], though no specific recommendation is given on their use.

Part 3 (Education for Learners) of the ASD Guideline [1] includes a brief paragraph about social skills groups (pg. 106) which is included as an instructor-led teaching strategy under Section 3.2b on social development. Described as a common technique, the intervention is described as relying on teaching and practising topics such as body language, emotion recognition and understanding, and conversational skills in a group of target children. Referring to one broad narrative review [27], the ASD Guideline notes that “studies have shown good participant and parent satisfaction, but only modest improvement in target social skills,” adding that social skills groups may improve a “participant’s mood and self-image”.

It should be noted that social skills groups are distinguished from peer mediated strategies in the ASD Guideline (endorsed in Recommendation 3.2.2.4) with respect to the source of instruction. Peer mediated approaches are based around interactions with trained, typically developing peers whereas social skills groups involve instruction from adult instructors/therapists.

Table 1.1: Recommendations relevant to social skills groups in children and young people in the NZ ASD Guideline

Original Reference	Original Recommendation	Grade
4.2.1	The development of social skills and community support groups for young people and adults should be undertaken to minimise and avoid problems.	C

In Part 4 (“Treatment and management of ASD”) of the ASD Guideline [1], social skills groups are referred to in the body text within section 4.2 on “Problem minimisation and avoidance” (pg. 139).

“Promising outcomes have been described for adults with ASD following participation in relatively inexpensive and time-efficient social skills groups and community support schemes.”

This statement references Recommendation 4.2.1 (pg. 138), which may be impacted by the current review on the effectiveness of social skills groups (see **Table 1.1**).

“The development of social skills and community support groups for young people and adults should be undertaken to minimise and avoid problems.”
(Grade C)

Part 5 (Living in the community) of the ASD Guideline [1] also mentions social skills groups explicitly in Section 5.2.d where the approach is listed among a range of strategies that can support young people with ASD.

As described in **Section 2.8** of the current supplementary report, the Living Guideline Group were tasked with reconsidering the wording and evidence grade of recommendation in relation to social skills groups specifically, and the possibility of developing additional recommendations, based on the updated evidence regarding social skills groups presented below.

1.3 Objectives of the current review update

The objectives of this review update were to:

- systematically identify, select, appraise and synthesise research evidence published since January 2004 relating to the effectiveness of social skills groups in children and young people with autism spectrum disorder; and to
- consider this evidence as it supplements that of the original ASD Guideline [1] to revise existing recommendations and/or develop new ones.

2 Social skills groups in children and young people with ASD

This chapter describes the findings of a systematic review update relating to social skills groups in children and young people with ASD. It also reports the development of new and revised recommendations by the Living Guideline Group to supplement the ASD Guideline [1] on this topic.

2.1 Scope and methods

Research question

The review update's research question was:

- How effective are social skills groups for assisting children and young people with ASD?

Sample

Included were children and young people aged 6-21 years (or for samples with a mean age within this range) diagnosed with Autism Spectrum Disorder (ASD) as classified by or consistent with DSM-IV [8] or DSM-5 [7].

A minimum number of 10 people with ASD was required for the treatment group.

Study designs

Eligible designs were randomised controlled trials which compare children and young people with ASD randomised to receive a social skills group intervention with those randomised to wait list control, usual care or no intervention. Changes in outcomes measured at baseline and immediately post-intervention (or an equivalent time for the no-treatment comparator group) were compared for the treatment and control groups. Longer-term followup of outcomes and measurement of maintenance of treatment effects were also reported.

Intervention/exposure

Included studies evaluated social skills groups. These provide structured training in the practical development of social skills in small group format, tailored specifically for people with ASD and include at least two people with ASD. A social skills group session typically includes a structured lesson on a specific skill including brief didactic instruction, demonstration of the skill (e.g., through role play), behavioural rehearsal of the skill, discussion, and individualised performance feedback. Homework in the way of

socialisation assignments may be set. Social skills groups meet on a regular basis, typically for 60-90 minutes for several weeks, and are facilitated by at least one trained instructor/therapist [13, 26].

Optional features include the involvement of peers and parents in the programme. Some groups include peer models who can demonstrate appropriate social skills and facilitate their practise. However these groups remain led by a instructor/therapist, are group-based, and are distinct from purely peer-mediated interventions. Some approaches also encourage parents to participate in the social skills groups as trainers or to attend separate, parallel sessions where they are taught how to encourage their child's social engagement. Multimodal interventions incorporating additional learning tools such as computers were eligible for inclusion in the current review where group-based social skills training from a instructor/therapist remained the main component of every session.

Comparator

Comparison groups could include people receiving no intervention, being in a wait list control, or receiving usual care. Comparisons between social skills group intervention and an alternative intervention were also included.

Outcomes

Measures of social competence were the primary outcome. To provide a comprehensive indication of outcome assessment, the current review included all outcomes relevant to sub-domains of social competence including social communication, emotion recognition, quality of social interaction/play, problem behaviour, and observed specific behaviours (e.g., counts of social initiations, length of conversations). Outcomes relevant to quality of life (including anxiety and depression, friendship quality) for the study participant (i.e., child/young person) as well as their parent/caregiver, and programme knowledge, were also eligible for inclusion.

Measures of programme acceptability such as satisfaction were excluded as they are considered to be poorly predictive measures of intervention effectiveness.

Identification and selection of studies for inclusion

Selection criteria were applied to abstracts to identify articles for retrieval, and then to retrieved full text articles, to identify included studies. Selection criteria for included and excluded studies are described in **Table 2.1**.

Search strategies were limited to publications from January 1 2004 onwards to ensure capture of articles published since the search was conducted for the original ASD Guideline [1]. Studies already appraised in the ASD Guideline were excluded from the current review regardless of date of publication.

Table 2.1: Inclusion and exclusion criteria for selection of studies

Characteristic	Inclusion criteria
Publication type	Studies published January 1 2004 or later.
Participant characteristics	Children and young people aged 6-21 years (or a sample with mean age within this range) diagnosed with Autism Spectrum Disorder (ASD) as classified by or consistent with DSM-IV-TR or DSM-5.
Sample size	Sample with ≥ 10 people receiving treatment.
Study Design	Randomised controlled trials Recently published secondary studies (systematic reviews and/or meta-analyses) that had a clear and relevant review question, used at least one electronic bibliographic database, included at least one study eligible for the current review, and published in or since 2012.
Exposure	Participation in a social skills training group including at least 2 people with ASD and at least one trained therapist/instructor.
Comparator	No intervention, wait list or usual treatment.
Outcome	Validated outcomes relevant to the study participant (i.e., child/young person) as well as their parent/caregiver including any of the following: <ul style="list-style-type: none"> • social competence • sub-domains of social competence, including social communication, emotion recognition, quality of social interaction/play, problem behaviour, observed specific behaviours • quality of life (including anxiety, depression of child/adolescent and parent/care-giver), and • programme knowledge.
Characteristic	Exclusion criteria
Publication type	Non-systematic reviews, correspondence, editorials, expert opinion articles, comments, articles published in abstract form, conference proceedings, dissertations, book chapters, or news items. Unpublished data
Attrition	Studies with $>50\%$ attrition from either arm of trial (unless adequate statistical methodology has been applied to account for missing data).
Language	Non-English language articles
Scope	Studies that did not provide separate analyses/syntheses of results relevant to the scope of the review (e.g., with respect to age group and diagnosis). Studies cited in the original ASD Guideline [1].
Study Design	Quasi-randomised studies, uncontrolled studies, observational studies, case series, n of one/single case experimental studies.
Outcome	Measures of programme acceptability such as satisfaction.

Nine bibliographic, health technology assessment, and guideline databases were included in the systematic search. The search was conducted in May-June 2014 and identified 1793 articles. References of retrieved articles were also cross-checked to identify additional articles.

Critical appraisal of included studies

Included studies were assigned “levels of evidence” which correspond to an evidence hierarchy [28]. This hierarchy ranks the quality of research designs which are broadly associated with particular methodological strengths and limitations. For the current review, the evidence hierarchy placed systematic reviews of randomised controlled trials as representing the most robust evidence (level 1 evidence), following by randomised controlled trials (level II).

Within each study design, studies can be conducted with varying degrees of rigour. This was reflected in assessment of methodological quality (including study validity, effect size, precision of results, applicability and generalisability) using design-specific validated instruments. Quality was coded as either good (+), uncertain (?), or poor (-).

Full details of review methods including search strategies, appraisal of study quality and data extraction are presented in **Appendix 1**.

2.2 Body of evidence

Included studies

Twelve studies met selection criteria and were eligible for appraisal and inclusion, two of which were recently published systematic reviews/meta analyses. The Evidence Tables for included studies are presented in **Appendix 3**. Throughout tables and text, studies are ordered according to the following hierarchy: study type (systematic reviews then primary studies), level of evidence (highest first), chronology of year of publication (oldest first), first author’s surname (alphabetical order).

Systematic reviews

Two recently published systematic reviews on the review topic were identified [24, 26].

A systematic review and meta analysis published by the Cochrane Collaboration [26] considered the effectiveness of social skills groups for children and adolescents aged 6 to 21 years with ASD. Randomised controlled trials published to 2011 were included comparing participants of social skills groups with no intervention/wait list control groups with social competence as the primary outcome.

A more recent review [24] was a methodology (rather than an efficacy) review of group-based social skills studies for people aged 18 years or under with ASD. A range of

study designs were considered including single case experimental studies, and uncontrolled and controlled group studies published between 2000 and 2012.

Primary studies

Ten primary studies were appraised [10, 13, 25, 29-35].

Study characteristics and outcomes for the included RCTs are summarised in **Table 2.2** organised broadly by intervention type.

Study location and setting

Seven primary studies were conducted in the United States, two studies were undertaken in Australia [10, 29], and one in South Korea [35]. Whilst the review was open to research in any context, and it was often not stated precisely where the social skills groups were held, studies tended to be conducted in clinical settings delivered by staff within specialised research centres or university-based treatment clinics.

Study design

All 10 randomised controlled trials compared a social skills training group intervention with a wait list control condition. All studies reported assessments of outcomes at baseline, and immediately post intervention for the treatment group or after an equivalent time for the control group. And so an intervention of 12 weeks duration would report assessments at baseline and after 12 weeks, although this post-test assessment does not represent 12 weeks follow-up post intervention. For 4 studies [10, 29, 34, 35], further follow-up was reported for the intervention group although this was not able to be controlled as the wait list control group had commenced their intervention by that time. Instead such follow-up assessment was used to measure within-group maintenance of any observed treatment effects.

Some studies reported on outcomes post intervention for the wait list control groups once they undertook their delayed intervention. However as such outcomes do not have a concurrent control group this data is not reported as evidence of treatment effect.

Participants

Sample sizes ranged from 17 to 68 (10 to 35 receiving the intervention), with 443 participants represented across the 10 studies. Nine of the 10 included studies included participants described as having High Functioning Autism (HFA), Asperger syndrome (AS), or PDD-NOS, with one study including children with a range of PDDs [31]. All but two studies included only participants who could be described as being cognitively high functioning in terms of having a minimum full IQ of at least 70, or being diagnosed with HFA or AS. The two studies which did not meet these criteria only included participants with a minimum verbal IQ of 65 [35] or 70 [32].

Six of the 10 studies [13, 25, 32, 33, 34, 35] also specified minimum verbal IQ levels in their selection criteria, some also describing participants as “verbally fluent”.

Six of the trials involved children with a mean age ranging from 8.4 - 9.7 years [10, 25, 29, 31, 33, 34]. Three studies concerned adolescents, mean ages between 13.6 – 14.6 years [13, 32, 35]. One study included young adults aged between 18 and 23 years with an average age of 20.4 years [30].

Samples were predominantly male (range: 70-96%).

In four of the trials, the majority of the sample (85% or more) were Caucasian [13, 31, 33, 34], with more mixed ethnicities represented in 3 studies including Caucasian (42-66%), Asian, Hispanic, and others [25, 30, 32]. Sample ethnicity was not reported for 3 studies: two undertaken in Queensland, Australia, [10, 29], and one in South Korea [35].

Interventions

Multiple approaches to social skills training were employed teaching a broad range of social skills. For convenience of reporting in **Table 2.2**, the social skills training (SST) groups assessed have been grouped into five categories:

- Program for the Education and Enrichment of Relational Skills (PEERS) [13, 30, 32, 35] or Children’s Friendship Training (which PEERS was adapted from for use with teens) [25] interventions;
- intensive “Skillstreaming” intervention employing multiple group sessions per day [33, 34];
- multi-modal SST (Junior Detective Training Programme) including a computer-based component [29];
- peer-mediated SST where each group included neurotypical peer tutors [31]; and
- SST focussed on improving the giving of, and response to, appropriate displays of affection [10].

These are fairly arbitrary groupings to reflect distinguishing features.

All 10 studies appear to have been manualised, and monitored programme fidelity to the curriculum in at least a sample of social skills groups using checklists. However the degree to which studies maintained fidelity was only reported in 3 studies, ranging from 75% for the evaluation of the peer mediated intervention [31] to 94-95% in the two Skillstreaming studies [33, 34]. The reliability of completion of fidelity checks was assessed in some studies; for example, by rating sub-samples of video-taped sessions [29].

Studies varied in intensity ranging from 10 hours (weekly 2-hour sessions over 5 weeks) for the intervention targeting affection [10] to the very intensive 46 hours (five 70-minute sessions per day, 5 days per week, over 5 weeks) for Skillstreaming [33, 34]. In between in programme intensity were PEERS and CFT interventions ranging

from 14 to 21 hours (60-90 minute weekly sessions held over 12-14 weeks) [13, 25, 30, 32, 35], 16 hours (2 hourly weekly sessions run over 8 weeks) for the multi-modal intervention [29], and 20 hours (75-minute weekly sessions run over 16 weeks) for the skills training intervention which included trained peer tutors [31].

All but one study [31] included weekly sessions for parents to attend, usually held concurrently with social skills group sessions. Homework tasks were usually set to be completed between sessions.

Outcomes assessed

A large and diverse range of outcomes were measured in the included studies, varying in type (rating scales, observations, and performance measures), informant (affected child/young person, parent/caregiver, teacher, clinician/assessor) and domain/construct measured ranging from specific behaviours to global improvement in social functioning.

Outcome measures used for each area of interest and in which included studies are described below.

Social competence

These rating scales assess a broad range of social skills and deficits that can characterise autism spectrum disorder. Most commonly used assessments of social competence included the Social Skills Rating Scale (SSRS) (all 5 PEERS/CFT studies) [13, 25, 30, 32, 35] and the Social Responsiveness Scale (SRS) (4 PEERS studies) [13, 30, 33, 34].

The Behavior Assessment System for Children (BASC-2) was used in the 2 Skillstreaming studies [33, 34].

Used in single studies were the following: Emotion Regulation and Social Skills Questionnaire (ERSSQ) [29]; Social Skills Questionnaire (SSQ) [29]; Social Communication Questionnaire (SCQ) [35]; Social Competence with Peers Questionnaire (SCPQ) [10]; and the Social Competence Inventory (SCI) [31].

In addition, several diagnostic scales were used as measures of ASD symptomatology with higher symptoms indicative of lower social competence. These included the Clinical Global Impressions-Improvement (CGI-I) [31]; the Autism Diagnostic Observation Schedule (ADOS) [35]; the Asperger Syndrome Diagnostic Scale (ASDS) [35]; and the Socialisation domain of the Vineland Adaptive Behavior Scale (VABS) [35].

Quality of life

Outcomes assessed that impact on quality of life included measures of mood, including social anxiety using the Social Interaction Anxiety Scale (SIAS) [13]; state and trait anxiety for parent and adolescent using the State and Trait Anxiety Inventory (STAI-S, STAI-T) [35]; parent ratings of child's anxiety using the Spence Child Anxiety Scale (SCAS) [10]; and depression of the child (Child Depression Inventory, CDI) or parent (Beck Depression Inventory, BDI) [35].

Also assessed were measures of the child/young person's perceived loneliness through the Loneliness Scale (LS), and Social and Emotional Loneliness Scale for Adults (SELSA) [25, 30].

Friendship quality was assessed through the Friendship Quality Scale (FQS) [13, 32], or in terms of perceived popularity using the Piers-Harris Self-Concept Scale (PHS) [25].

Specific skills or sub-domains of social competence

- **Emotion recognition** was assessed in three studies using different performance measures. These included one administered via computer using the Diagnostic Analysis of Nonverbal Accuracy (DANVA2) [33, 34], and two involving presentation of photographs requiring ratings/recognition of facial expression, and posture [29].
- **Emotion management** knowledge was assessed in a study where the child was read scenarios and asked to offer suggestions about how to cope with anxiety (“Dylan is being teased”) and with bullying (“James and the maths test”) [29].
- **Empathy** of the child/adolescent was rated by the parent using the Empathy Quotient (EQ) [30]
- **Affection** was targeted in one intervention and was the focus of three outcomes for children [10]. Three parent-administered measures were used to assess the following: the capacity to engage in affectionate behaviour outside immediate family (using the Affection For Others Questionnaire, AOQ); the ability to determine the purpose of affection (Walk in the Forest Test, WFT); and difficulty in general affection (excessive and inadequate affection) (General Affection Questionnaire, GAQ).
- **Problem behaviours** were measured in three studies, all relating to PEERS/CFT, in ratings by teachers, parents, and children/young people. Measures included a subscale of the SSRS scale [13, 32], and for teacher ratings of withdrawal and aggression using the Pupil Evaluation Inventory (PEI) [25].
- **Idiomatic language** was the only specific measure related to social communication. It was assessed by the Comprehensive Assessment of Spoken Language (CASL) [33].
- **Quality of play** was measured using two similar scales, the QPQ (Quality of Play Questionnaire) and QSQ (Quality of Socialization Questionnaire). These were completed by children/adolescents and/or parents for the five PEERS/CFT interventions [13, 25, 30, 32, 35]. Forms of the measures sometimes used included quantification of the number of get-togethers the child/adolescent hosted, and was invited to, in the last month; where these occurred, whether any conflict occurred, and whether engagement activities (socially interactive) or disengagement activities (minimally interactive activities such as computers and television) were undertaken.
- **Programme-specific skills/knowledge** was assessed in 6 studies, using researcher-developed tools. The four PEERS evaluations employed either the TASSK (Test of Adolescent Social Skills Knowledge) or a version adapted for young adults (TYASK) [13, 30, 32, 35]. The two Skillstreaming evaluations [33, 34] included assessor rated measures of child’s programme knowledge

(Skillstreaming Knowledge Assessment, SKA), and ratings of observed skills taught in the programme using the Adaptive Skillstreaming Checklist (ASC).

- Satisfaction outcomes were measured in some studies as a measure of programme acceptability (but not effectiveness). Whilst these are not reported in the body of this report they are provided in the Evidence Tables reporting detailed information for each included study (see **Appendix 3**).

Most outcomes were based on validated rating scales, however some tools were developed by researchers for their study and reliability and validity were not comprehensively determined. Performance measures were assessed for emotional recognition tasks in 3 studies, programme knowledge (SKA) in 2 studies, and emotion management tasks in one study. Direct behavioural assessments of global autistic symptoms were made in only one study using the ADOS diagnostic tool [35]. However, no direct observational counts of targeted social skills were made in naturalistic environments (e.g., counts of social initiation observed in the playground).

Blinding

Due to the nature of the study designs, the child/young person receiving the intervention, their caregiver/parent, and the instructors/therapists providing the intervention were unable to be blinded to the group allocation and therefore outcomes were open to detection and performance biases. That is, the way outcomes were reported, and/or the way the individual was treated, may have varied in ways not strictly related to the intervention. Some studies included teachers who were able to be blinded to group assignment, although response rates were often poor for these informants.

2.3 Quality of included studies

Studies were assigned levels of evidence and quality codes according to methods previously described (see **Section 2.1**, and **Appendix 1**, Section A1.5).

Two systematic reviews [24, 26] included RCTs and were ranked at level II in the NHMRC hierarchy of evidence [28]. The methodological review of group-based social skills studies [24] was coded as being of uncertain quality (?). It employed systematic though limited searching, coding of predetermined quality criteria, and presented detailed tables of included studies. A Cochrane Collaboration systematic review [26] was coded as being of good quality (+). The review was based on an extensive search involving multiple databases and explicit selection criteria, and employed independent selection and appraisal using validated quality coding by two reviewers. Effect sizes were calculated and random effects meta analyses performed.

Of the 10 primary studies appraised [10, 13, 25, 29-35] all were randomised controlled trials and ranked at level II of the NHMRC hierarchy of evidence [28]. Of these, 5 were

coded as being of good quality (+) [13, 25, 29, 34, 35] and 5 were rated as being of uncertain quality (?) [10, 30-33]. No studies were assessed as being of poor (-) quality.

Table 2.2: Characteristics and results of included primary studies organised by type of social skills group

Reference	Quality, country	Intensity	Sample	Outcomes (assessment tool) (informant) with significant treatment improvements	Outcomes (assessment tool) (informant) with no significant treatment improvements
PEERS/Children's Friendship Training					
Laugeson et al (2009) [32]	Quality: ? US	12 90-minute weekly sessions =18 hours	n=33 13-17 years (M=14.6) 85% male TG: n=17, CG: n=16	- social competence (SSRS) (PARENT) - friendship quality (FQS) (ADOLESCENT) - PEERS social skills knowledge (TASSK) (ADOLESCENT) - hosted get-togethers (QPQ) (ADOLESCENT)	- social competence (SSRS) (TEACHER) - problem behaviour (SSRS) (ADOLESCENT) - problem behaviour (SSRS) (PARENT) - problem behaviour (SSRS) (TEACHER) - invited, conflict, get-togethers (QPQ) (ADOLESCENT) - hosted, invited, conflict, get-togethers (QPQ) (PARENT)
Frankel et al (2010) [25]	Quality: + US	12 90-minute weekly sessions =18 hours	n=68 2 nd - 5 th grade (M age=8.4 years) 85% male TG: n=35, CG: n=33	- loneliness (LS) (CHILD) - friendship quality/popularity (PHS) (CHILD) - disengaged activity (QPQ) (PARENT) - self-control (SSRS) (PARENT) - hosted get-togethers (QPQ) (PARENT)	- assertion, externalising, internalising (SSRS) (PARENT) - withdrawal, aggression (PEI) (TEACHER) - invited, conflict, engagement in, get-togethers (QPQ) (PARENT)
Gantmen et al (2012) [30]	Quality: ? US	14 60-minute weekly sessions =14 hours	n=17 18-23 years (M=20.4) 70% male TG: n=10, CG: n=7	- social competence (SRS) (PARENT) - social competence (SSRS) (PARENT) - social loneliness (SELSA) (YOUNG ADULT) - empathy (EQ) (PARENT) - PEERS social skills knowledge (TYASK) (ADULT) - hosted, invited, get-togethers (QSQ) (PARENT)	- hosted, invited to get-togethers (QSQ) (ADULT)
Schohl et al (2014) [13]	Quality: + US	14 90-minute weekly sessions =21 hours	n=58 11-16 years (M=13.6) 52% male (90%) TG: n=29, CG: n=29	- social competence (SRS) (PARENT) - problem behaviours (SSRS) (PARENT) - problem behaviours (SSRS) (TEACHER) - social anxiety (SIAS) (ADOLESCENT) - PEERS social skills knowledge (TASSK) (ADOLESCENT) - hosted, or invited, get-togethers (QSQ) (ADOLESCENT)	- social competence (SSRS) (PARENT) - social competence (SSRS) (ADOLESCENT) - social competence (SRS) (TEACHER) - friendship quality (FQS) (ADOLESCENT) - hosted, invited, conflict at get-togethers (QSQ) (PARENT) - conflict at get-togethers (QSQ) (ADOLESCENT)
Yoo et al (2014) [35]	Quality: + South Korea	14 90-minute weekly sessions =21 hours	n=47 12-18 years (M=14.0) 96% male TG: n=23, CG: n=24	- social competence (VABS) (PARENT) - social competence/ASD symptoms (ADOS) (ASSESSOR) - depression (CDI) (ADOLESCENT) - parent's state anxiety (STAI-S) (PARENT) - PEERS social skills knowledge (TASSK) (ADOLESCENT)	- social competence (SSRS) (ADOLESCENT) - social competence/ASD symptoms (SCQ) (PARENT) - social competence/Asperger behaviour (ASDS) (PARENT) - state or trait anxiety (STAI-S/T) (ADOLESCENT) - parent's trait anxiety (STAI-T) (PARENT) - parent's depression (BDI) (PARENT) - hosted, invited, conflict at get-togethers (QPQ) (PARENT) - hosted, invited, conflict at get-togethers (QPQ) (ADOLESCENT)

Table 2.2: Characteristics and results of included primary studies grouped by type of social skills group (*continued*)

Reference	Quality, country	Intensity	Sample	Outcomes (reported by) (assessment tool) with significant treatment improvements	Outcomes (reported by) (assessment tool) with no significant treatment improvements
Intensive Skillstreaming					
Lopata et al (2010) [33]	Quality: ? US	5 70-minute sessions per day, 5 per week, over 5 weeks =46 hours	n=36 7-12 years (M=9.5) 94% male TG: n=18, CG: n=18	<ul style="list-style-type: none"> - social competence (SRS) (PARENT) - withdrawal (BASC-2) (PARENT) - idiomatic language (CASL) (CHILD) - skills taught in programme (ASC) (PARENT) - Skillstreaming programme knowledge (SKA) (ASSESSOR) 	<ul style="list-style-type: none"> - social skills behaviour (BASC-2) (PARENT) - emotion recognition (DANVA-2) (CHILD)
Thomeer et al (2012) [34]	Quality: + US	5 70-minute sessions per day, 5 per week, over 5 weeks =46 hours	n=35 7-12 years (M=9.3) 86% male TG: n=17, CG: n=18	<ul style="list-style-type: none"> - social competence (SRS) (PARENT) - social skills behaviour (BASC-2) (PARENT) - skills taught in programme (ASC) (PARENT) - idiomatic language (CASL) (CHILD) - Skillstreaming programme knowledge (SKA) (ASSESSOR) 	<ul style="list-style-type: none"> - withdrawal (BASC-2) (PARENT) - emotion recognition (DANVA-2) (CHILD)

Table 2.2: Characteristics and results of included primary studies grouped by type of social skills group (continued)

Reference	Quality, country	Intensity	Sample	Outcomes (reported by) (assessment tool) with significant treatment improvements	Outcomes (reported by) (assessment tool) with no significant treatment improvements
Multimodal skills training with computer game component (Junior Detective Training Program)					
Beaumont and Sofronoff (2008) [29]	Quality: + Australia	8 2-hourly weekly sessions =16 hours	n=49 7-11 years (M=9.7) 90% male TG: n=26, CG: n=23	- social competence (SSQ) (PARENT) - social competence (ERSSQ) (PARENT) - emotion management (anxiety) (ASSESSOR) - emotion management (bullying) (ASSESSOR)	- social competence (SSQ) (TEACHER) - social competence (ERSSQ) (TEACHER) - emotion recognition: facial expression (ASSESSOR) - emotion recognition: posture cues expression (ASSESSOR)
Peer-mediated skills training					
Koenig et al (2010) [31]	Quality: ? US	16 75-minute weekly sessions =20 hours	N=42 8-11 years (M=9.2) 77% male TG: n=24, CG: n=18	- proportion of responders indicating change in social functioning (CGI-I) (PARENT)	- social competence (SCI) (PARENT)
Affection-focussed skills training					
Andrews et al (2013) [10]	Quality: ? Australia	5 2-hour weekly sessions =10 hours	n=58 7-12 years (M=9.0) 81% male (81%) TG: n=29, CG: n=29	- affection for others (AOQ) (PARENT)	- assess purpose of affection (WFT) (PARENT) - excessive or inadequate affection (GAQ) (PARENT) - social competence (SCPQ) (PARENT) - symptoms of child anxiety (SCAC) (PARENT)

Key: ? indicating rating of uncertain quality; + indicating rating of good quality; ADI-R=Autism Diagnostic Interview-Revised; ADOS=Autism Diagnostic Observation Schedule-Generic; AOQ=Affection for Others Questionnaire; AS=Asperger syndrome; ASC=Adaptive Skillstreaming Checklist; ASDS=Asperger Syndrome Diagnostic Scale; ASD=autism spectrum disorder; BASC-2=Behavior Assessment System for Children – 2nd edition; BDI=Beck Depression Inventory; CASL=Comprehensive Assessment of Spoken Language; CBCL=Child Behavior Checklist; CDI=Child Depression Inventory; CG=control group; CGI-I=Clinical Global Impressions-Improvement; ERSSQ=Emotion Regulation and Social Skills Questionnaire; EQ=Empathy Quotient; FQS=Friendship Qualities Scale; GAQ=General Affection Questionnaire; IQ=intelligence quotient; LS=Loneliness Scale; M=mean; PEERS=Program for the Education and Enrichment of Relational Skills; PEI=Pupil Evaluation Inventory; PHS=Piers-Harris Self-Concept Scale; QPQ=Quality of Play Questionnaire; QSQ=Quality of Socialization Questionnaire; RCT=randomised controlled trial; SCAS=Spence Child Anxiety Scale; SCI=Social Competence Inventory; SCPQ-P=Social Competence with Peers Questionnaire; SCQ=Social Communication Questionnaire; SELSA=Social and Emotional Loneliness Scale for Adults; SKA=Skillstreaming Knowledge Assessment; SIAS=Social Interaction Anxiety Scale; SSQ=Social Skills Questionnaire; SRS=Social Responsiveness Scale; trait versions; TASSK-R=Test of Adolescent Social Skills Knowledge-Revised; TG=treatment group; US=United States of America; TYASK=Test of Young Adult Social Skills Knowledge; VABS=Vineland Adaptive Behavior Scale; WFT=Walk in the Forest Test.

2.4 Narrative appraisal of studies

A narrative critique of included studies' individual strengths and limitations is provided in this section. Full details are presented in the appendicised Evidence Tables (**Appendix 3**). Results are summarised across all studies in **Table 2.2**. For ease of comparison, results specific to each outcome are also presented in separate tables (**Tables 2.3 – 2.9**) referred to in the synthesis of main findings (**Section 2.5**).

Systematic reviews

Two systematic reviews were included [24, 26].

Reichow et al (2013)

A Cochrane collaboration systematic review [26] examined the impact on social competence of participating in social skills groups compared to no intervention controls in young people aged 6-21 years with ASD. Five RCT's (with 196 participants) published to 2011 were identified as eligible and were included in a meta analysis. Modest gains over controls were found for the intervention group for social competence (ES=0.47, 95% CI=0.16 to 0.78, p=0.003) based on 4 studies, and friendship quality (ES=0.41, 95% CI=0.02 to 0.81, p=0.04) from 2 studies, and in a single study, for decreased loneliness (ES=-0.66, 95% CI=-1.15 to -0.17, p=0.008). No treatment effects were found for emotional recognition (2 studies), social communication as related to understanding of idioms (one study), or child or parental depression (one study). The authors concluded that there was emerging evidence for effectiveness of social skills group interventions for some young people with ASD and called for more research, especially with respect to improvements in quality of life.

Kaat & Lecavalier (2014)

This recent systematic review [24] was a methodological critique of group-based social skills training (SST) for children and young people with ASD. There were 48 studies appraised applying a range of study designs, including 13 studies described as randomised controlled trials (though 2 of these could not be verified as employing fully randomised group allocation). The authors suggested that it would be

“presumptive to conduct an efficacy review across all SST interventions given the differences in program design, targeted skills, instructional methods, and the various outcome reported.”

The review described common limitations of the evidence base, including lack of independent confirmation of diagnoses, and limited sample scope and size. However the reviewers also observed that study quality had increased in recent years, with most using or adapting manualised treatments and using validated outcome measures. Recommendations were provided for the design of future trials.

Primary studies

Table 2.2 summarises key findings of the 10 primary studies appraised relevant to the effectiveness of social skills groups in children and adolescents with ASD [10, 13, 25, 29-35].

Narrative summaries are provided below for each primary study organised broadly by intervention approach in the following order: Program for the Education and Enrichment of Relational Skills (PEERS)/Childhood Friendship Training (CFT) (5 studies), intensive intervention (Skillstreaming) (2 studies); multimodal Junior Detective intervention (1 study), peer mediated intervention (1 study), and affection focused intervention (1 study).

PEERS and CFT studies

Laugeson et al (2009)

A social skills programme called PEERS (Program for the Education and Enrichment of Relational Skills), developed from the Children's Friendship Training programme [25], was evaluated in an RCT conducted in Southern California [32]. The manualised PEERS programme provides concrete rules and steps for common social situations including conversational skills; peer entry and exiting skills; how to develop friendship networks; practising good sportsmanship; how to handle teasing, bullying and arguments with peers; changing bad reputations; and how to be a good host during get-togethers with friends. PEERS provides didactic instruction, role playing behaviour rehearsal, coaching with performance feedback, weekly socialisation assignments and homework review. Parents attend separate concurrent sessions to learn how to act as social coaches and encourage social engagement.

In the trial, 33 verbally fluent adolescents aged 13-17 years (mean age of 14.6 years) were randomised to either the social skills group (PEERS) (n=17) or to a wait list control group (n=16), with repeated measures at baseline and 12 weeks (immediately post PEERS completion or equivalent time for controls). PEERS was provided to small groups of approximately 7 people over 12, weekly, 90-minute sessions. Students were drawn from a range of educational settings and ethnic backgrounds and there was a high rate of trial attendance and completion (92%) with regular fidelity checks to ensure the curriculum was followed.

There were no differences at baseline between the treatment and control groups in socio-demographic, interpersonal relationship skills (VABS), and outcome variables. Repeated measures Group (Treatment vs Control) X Time (pre vs post) MANOVAs were conducted for adolescent and parent outcomes, and teacher outcomes, separately. Of 12 outcomes assessed, treatment effects were evident for 4. Adolescents receiving PEERS improved compared with baseline in their self-reported friendship quality, knowledge of PEERS skills (measured using the TASSK instrument), and their perceived number of hosted get-togethers. PEERS participants' parents also

reported improvements in their teens' social skills (SSRS). There were no improvements on these variables for the control group over time.

There were no significant treatment effects found for the remaining outcomes including adolescent, parent or teacher reported problem behaviour (SSRS), adolescent reported ratings of get-togethers (invited or conflict at) or parent ratings of get-togethers (hosted, invited or conflict at) (QPQ); or teacher reported social skills (SSRS).

Some limitations acknowledged by the authors of this trial include that the ASD diagnoses of the sample were not independently verified. With respect to outcomes, no direct observational data was collected, parents were not blind to allocation and whilst independent blinded assessments were requested from teachers, these analyses were limited by a poor return rate of questionnaires (39%).

Frankel et al (2010)

Children's Friendship Training (CFT) is a manualised, parent-assisted social skills programme for children that specifically targets ecologically valid friendship skills. It was developed at UCLA outpatient clinics in Southern California and was adapted for teens to form PEERS by overlapping research teams at the same institution. CFT was not explicitly developed for children with ASD and has been used with children with a range of social needs, including ADHD, Fetal Alcohol Spectrum Disorder, and Adjustment Disorder. This RCT was conducted to investigate the feasibility and efficacy of CFT for children with ASD [25].

The trial included 68 children with verified "high functioning autism" (HFA) who attended (at least most of the school day) 2nd to 5th grade classes in mainstream schools. The children were aged 8.4 years on average. To be eligible, students needed to demonstrate some basic social knowledge and ability, such as being able to switch topics in a conversation, having the capacity for joint attention and basic social reciprocity, and knowing the rules of at least two common board games (e.g., chess) and common school yard games (e.g., handball).

A unique aspect of CFT compared to other social skills training group interventions in this review is that the groups included typically developing people without a diagnosis of ASD who had significant social skills problems. The mixing of ASD and non-ASD children was said to serve the programme aim of integrating children with ASDs into groups of typically developing children [36]. The non-ASD children participating in the CFT groups but not included in analyses had ADHD (46%), Adjustment Disorder and/or Oppositional Defiant Disorder (22%), anxiety disorder (5%), Fetal Alcohol Syndrome Disorder (1%), mood disorder (1%), learning disability (1%), or no diagnosis (25%).

Participants with ASD were randomised to receiving CFT treatment (N=35) or assigned to a wait list control group (N=33). The Children's Friendship Training groups each included 10 children across 1-2 school grade levels representing a range of diagnoses, with no more than 4 children with ASD in each treatment group. The programme aims to teach social etiquette, friendship skills and specific rules of behaviour which are used by peers over 14, weekly, 60-minute modules following similar topics as for PEERS.

Parents attend concurrent sessions, and are encouraged to supervise play dates between their child and other children who are not part of the social skills group. Outcomes were measured at baseline and immediately post-intervention (or 14 weeks post baseline for controls).

There were no significant differences between groups at baseline in socio-demographic, IQ, VABS, and outcome variables. ANCOVAs found statistically significant improvement for the treatment group compared with controls for 5 of 13 outcomes. Specifically, following Children's Friendship Training, participants reported lower perceived loneliness (LS) and higher friendship quality/perceived popularity (PHS). Mothers of CFT participants reported that their children hosted more get-togethers (QPQ), had decreased use of disengaged activities during get-togethers (minimally interactive activities such as computers and television) (QPQ), and increased in their self-control during provocation (SSRS). No significant treatment effects were found for parent-reported ratings of get-togethers (invited to, conflict at, or engagement in) (QPQ); in social skills relating to assertion (making friends, playing well), externalising (intrusive and aggressive behaviour), and internalising (social withdrawal) (SSRS); or in teacher-reported outcomes relating to pupil withdrawal, or aggression (PEI).

The authors concluded that the Children's Friendship Training programme was feasible and cost effective with modest improvements evident for the treatment group on child measures of popularity and loneliness, and parent measures of social skill and play date behaviour. It should be noted that the trial eligibility was limited to enrolling children who were somewhat functionally, verbally and socially able as previous use of the CFT had demonstrated that some basic social ability and knowledge of common games was necessary for full participation in the social skills group training. Similar to many studies in this area, no direct observational data were collected and parents were not blind to allocation and as active implementers of treatment were therefore open to detection and performance bias in their assessments. Whilst teachers were blinded to allocation, analyses revealed that they were less likely to report on children rated as having more externalising (intrusive, aggressive) behaviours at baseline. This suggests a response bias toward children who were less aggressive initially.

A follow-up study [36] of 24 of the 66 study participants investigated maintenance of treatment effects 1-5 years post intervention (mean age of 12.6 years). As this data was uncontrolled, it was excluded from the current review and was not formally appraised, however the findings are relevant to longer term maintenance. Compared to baseline, at follow-up the participants were reported by their parents to have been invited on significantly more play dates and showed less conflict during them, and the children reported having increased social skills with fewer problem behaviours. However the study was marred by considerable attrition, and the fact that these outcomes are uncontrolled and may improve over time with maturation.

Gantman et al (2012)

A group of UCLA researchers with overlapping authorship to another included PEERS trial [32] conducted a small RCT to evaluate a PEERS training intervention adapted for young adults [30]. Modifications included additional modules on dating etiquette, handling peer pressure and avoiding exploitation, and to the level of caregiver involvement required in the programme with an emphasis on the young adult's need for social independence. There were 14, weekly 90-minute sessions held for the young adults, with parents attending separate concurrent sessions.

The small study included 17 "high functioning" (IQ above 70) young adults aged 18-23 years (mean age of 20.4 years) with "English fluency" who were attending college at least part-time, and living or in daily interaction with their caregivers. Participants were randomised to the PEERS social skills group (n=10) or to a wait list control group (n=7), with repeated measures at baseline and at 14 weeks (immediately post intervention for the PEERS group) with no further follow-up.

There were no significant group differences at baseline in age, IQ, or social or ASD symptom outcomes. Outcome scores were converted to difference scores (post-test minus baseline) with a significant group effect found by MANOVA in multivariate analyses. Univariate tests suggested significant improvements in outcomes over time for the PEERS participants compared with those in the wait list control group. Specifically, in outcomes reported by the young adults themselves, the treatment group had increased knowledge of PEERS (TYASK), and reduced social loneliness (SELSA). In caregiver reported outcomes, the PEERS participants were observed to have reduced social competence/ASD symptoms (SRS total, and some subscales), improved social competence (SSRS total, and some subscales), and improved empathy (EQ). Non-parametric tests indicated that the PEERS group had increased hosted, and invited, get-togethers as reported by care-givers (QSQ).

The results for non-significant tests were not clearly stated but appear to include young adults' reported frequency of get-togethers, and some subscales of the SRS and SSRS. Whilst attempts were made to gather independent reports on some outcomes from participants such as College teachers, data returned was limited, and not reported.

This small trial had several limitations in method and reporting. The participants' ASD diagnoses were not independently verified. There was no direct observational data collected, and no independent unblinded assessments were collected. The authors also noted that there is a lack of appropriate ASD outcome measures validated for adults.

With respect to the significant findings, there was no adjustment to alpha for multiple tests and as many were only significant at the 0.05 level, they may have been chance effects. The sample size was also very small, reducing statistical power. It was not clear precisely what sub-scale tests were conducted as only significant univariate results were reported. No follow-up data was collected beyond immediate post-test and so maintenance of effects were not monitored.

Schohl et al (2014)

A recent RCT [13] broadly replicated the PEERS evaluation by Laugesan et al, 2009 [32] included in the current review. Set in a medium-sized US city, the trial included 58 adolescents with higher functioning ASD aged 11-16 years (mean age=13.6 years, SD=1.5); most of whom were male (81%), and caucasian (90%). Included were adolescents with an ADOS-verified ASD diagnosis, parent-reported social problems, verbal IQ of 70 or above, English fluency, and an interest in joining the social skills group.

Participants were randomised to either the intervention (PEERS) (n=29) or to a wait list control group (n=29), with repeated measures at baseline and at 14 weeks, immediately following completion of the PEERS programme for the treatment group. PEERS was provided to small groups of up to 10 people over 14, weekly, 90-minute sessions. Homework compliance and attendance was enforced and participation rate post group allocation was high (92%).

Fourteen outcome measures were assessed at baseline and follow-up, with rating scales completed by the adolescents, parents and/or teachers. There were no significant differences between groups at baseline in demographic or outcome measures. Analyses included repeated measures MANOVAs, ANOVAs and confirmatory post hoc paired t-tests (with Bonferroni correction applied).

Significant intervention effects were found for 7 of the 14 outcome variables assessed. Compared to wait list controls, participation in a PEERS social skills group led to significant improvements between baseline and follow-up. Specifically, adolescents receiving PEERS reported greater knowledge of PEERS concepts and skills (measured using the TASSK instrument), increased number of both hosted and invited get-togethers (QSQ), and decreased social anxiety (SIAS). Parents reported lower scores for social competence/ASD symptoms (SRS-P), and problem behaviours (SSRS-P) for children receiving PEERS. Teachers also reported lower problem behaviour scores for students in the intervention social skills group.

By contrast, no significant treatment effects were found for 7 other variables: adolescent-reported conflict at get-togethers (QSQ-A), and friendship quality (FQS); parent-reported ratings of get-togethers (hosted, invited, conflict) (QSQ-P), and social skills (SSRS-P); and teacher-rated social skills (SSRS-T) or social competence/ASD symptoms (SRS-T).

This well conducted study replicates many of the findings of the earlier evaluation of PEERS [32] with the addition of reduced social anxiety. Intervention effects were evident for half of the outcomes, including most of the adolescent-reported ones. However, there were a number of study limitations. Whilst parents were not able to be blinded to PEERS allocation, the teacher assessments were blinded, however there was a large amount of missing data for the teachers' assessments which may have decreased power in analysis. No behavioural observations were measured.

Yoo et al (2014)

This RCT attempted to replicate and validate the cross-cultural utility of PEERS in South Korea [35]. There were 47, verbally fluent Korean adolescents aged 12-18 years in the trial; 23 receiving PEERS and 24 in the wait list control group. All but 2 were male. Inclusion and exclusion criteria, treatment and procedures were similar to Laugesan et al (2009), however great care was taken to make the translated PEERS training curriculum culturally appropriate. Korean versions of the outcome rating scales were also used. A large number of outcomes were assessed at baseline and 14 weeks post-test. Maintenance at 3 months follow-up was also assessed for a subset of measures, though it is unclear whether these were reported for just the treatment group, or also included the control group's data after receiving delayed treatment.

There were no significant differences at baseline for key socio-demographic measures and some outcomes measures (analyses were not reported for all outcomes at baseline), although a difference was noted on one sub-scale (imagination) of one instrument (ADOS).

Repeated measures Group (treatment and control) by Time (baseline and 14 weeks) ANOVAs found that the PEERS group, compared with control group, were significantly improved at follow-up on several outcomes. These included PEERS social skills knowledge (TASSK-R); interpersonal relationship skills (socialisation domain of VABS); ASD symptoms (measured by assessor observation using the ADOS); and adolescent depression (CDI). In addition, mother's state anxiety (STAI-S) was decreased at follow-up. Models adjusting for likely covariates led to similar results although a treatment effect for parent-reported adolescent emotional problems (CBCL) appears to be confounded by baseline differences in adjusted means for this measure.

By contrast, there were no significant treatment effects for get-togethers (as measured by the QPQ); social behaviour (SSRS); ASD symptoms (SCQ, SRS); Asperger syndrome behavioural characteristics (ASDS); state or trait anxiety for the adolescent (STAI-S/T); or trait anxiety (STAI-T) or depression (BDI) for the parent. The lack of treatment effects for the SRS and SSRS is in contrast to the effects found in previous PEERS evaluations [13, 32], however the Korean study used adolescent ratings of social behaviour in the SSRS in contrast to the parent-reported SSRS in the other PEERS trials. The authors suggest that some adolescents may have limited insight into their difficulties, and the observations of parents, teachers, and in this study, directly observed improvements in social skills, may be more reliable [35].

Investigating maintenance effects at 3 months, scores did not change significantly from assessments made at the completion of PEERS. An exception was PEERS knowledge (TASSK) which was decreased, but was still higher than at baseline. Unfortunately, direct behavioural ratings/global behavioural assessment using the ADOS and the VABS were not assessed at 3 months.

The authors concluded that the PEERS social skills intervention with modest cultural adjustment appears to be efficacious for teens with ASD in South Korea. The study was reasonably well conducted and has advantages over the previous PEERS studies

[13, 32]. Global behavioural assessments (using the ADOS and VABS) were made and revealed improvements on these ratings post PEERS. Unfortunately the ratings were not blinded to allocated condition and therefore were open to the potential for bias in assessment. Further, there were no third party (e.g., teacher) assessments. Another drawback of the study was that many assessment tools were applied and many tests run for subdomain scores without any correction to alpha. This may lead to a bias toward finding significant treatment effects that are actually chance effects.

Intensive Skillstreaming

Lopata et al (2010)

A US-based RCT [33] investigated the feasibility and efficacy of a manualised “Skillstreaming” social skills approach for children with ASD. The intensive programme (46 hours) was run over 5 weeks, with five 70-minute treatment cycles per day. Each cycle included 20 minutes of instruction using a multi-step sequence including direct instruction, modeling, role-play, performance feedback, and setting of homework. Intensive instruction was followed by 50 minutes of therapeutic activity to practice and reinforce taught skills. Activities targeted social skills, face-emotion recognition, interest expansion, and interpretation of non-literal language. Parents attended weekly 90-minute sessions covering the programme content and received training to encourage their child’s generalisation of skills at home.

Participants were 36 children with “high functioning” (IQ above 70) ASD aged 7-12 years (mean age of 9.5 years), the high majority of whom were Caucasian (89%), and male (94%). ASD diagnoses were determined by written report and not independently verified by the researchers, although historical ADI-R results were made available to support ASD diagnoses for 80% of participants. Participants were stratified into three age groups (7-8, 9-8, 11-12 years) prior to randomising into one of three treatment groups of the same age range (n=18, 6 children in each), or the wait list control group (n=18).

Between group (treatment versus control) ANCOVA’s were conducted to predict the 5 week post-test scores, with baseline scores as covariates. Applying a Bonferroni corrected alpha, significant improvement in the treatment group compared with controls was observed for 5 of 7 outcome measures including: parent-reported decreased social competence/ASD symptoms (SRS), reduced withdrawal behaviours (BASC-2), and increased ratings of social skills taught in the Skillstreaming programme (ASC); and improvements in child-reported idiomatic language (CASL), and programme knowledge (SKA). Standardised effect size estimates were generally in the medium and large range. There were no significant treatment effects for parent-reported social skills behaviour (BASC-2) or in the emotion recognition tasks completed by the child via computer (DANVA-2).

High levels of treatment fidelity were reported. For the treatment group, 50% had post-test ASD social impairments (on the SRS) that decreased from the severe to mild-to-

moderate range, or from mild-to-moderate to normal range, compared with no children in the wait list control group changing to a less severe score.

In common with most intervention studies, no direct observational data was collected and parents were unable to be blinded to allocation. Staff completed some outcome measures but only for participants in the treatment group. Maintenance of treatment effects was not investigated.

Thomeer et al (2012)

This RCT [34] is a replication and extension of Lopata et al's trial [33] investigating Skillstreaming using almost identical methods and in a similarly sized and characterised sample (see **Table 2.2**). The two main differences were that the Thomeer et al study addressed limitations of the earlier trial with respect to independently verifying ASD diagnoses using a gold standard screening tool (ADI-R), and including a follow-up assessment at 2-3 months to assess maintenance of treatment effects.

The study results were very similar to the earlier RCT. Again there were improvements in 5 of 7 outcomes in the treatment group compared with the wait list control group. These included parent-reported decreased social competence/ASD symptoms (SRS), increased ratings of social skills taught in the Skillstreaming programme (ASC), and reduced social skills behaviours (BASC-2); and improvements in child-reported programme knowledge (SKA), and idiomatic language (CASL), again with standardised effect size estimates generally in the medium and large range. There were no significant treatment effects for parent-reported withdrawal behaviour (BASC-2) or in the emotion recognition tasks (DANVA-2).

Essentially the results are the same for the two RCT's evaluating Skillstreaming except that the BASC-2 subscale relating to social skills behaviour was improved in this trial where as the subscale relating to withdrawal behaviour was not, a reverse of the findings of the earlier Lopata et al trial [33].

The study [34] also considered maintenance of effects by conducting follow-up assessment 2-3 months after post-test for the parent-report measures in the treatment group only. Results found that outcomes remained significantly improved at follow-up compared with baseline for knowledge of Skillstreaming programme skills (ASC) and social skill behaviour subscale (of the BASC-2), but were not maintained for social competence (SRS) and the withdrawal behaviour subscale (of the BASC-2).

Multimodal social skills group

Beaumont & Sofronoff (2008)

An Australian RCT evaluated a manualised, multimodal social skills group programme called the Junior Detective Training Program (JDTP) [29]. This programme is aimed at enhancing the emotional understanding and social skills of children with Asperger syndrome (AS). Over 8 weeks, 2-hourly sessions are run weekly in small groups of three children and two therapists. Parents attend separate concurrent sessions and

hand-outs are provided for teachers to support the skills taught. The first four sessions include a computer game component where the central character is a “junior detective” and needs to decode suspects’ thoughts and feelings. Computer-animated cartoon and human characters are used to teach children how to recognise complex emotions (e.g., guilt, embarrassment, suspicion, teasing) from non-verbal and environmental characteristics, and prosody of speech. Skills are applied in “virtual reality” missions. The small group therapy component was aimed at generalising computer game content and teaching additional social skills including generic strategies to solving social problems. Features of the group sessions included modeling new skills, role-plays, practise with peers, group discussions, and setting of “home missions”¹.

The trial included 49 children with reported previous diagnoses of AS aged 7-11 years (mean age of 10 years), 26 in the treatment group and 23 in the wait list control group. Outcomes were measured at pre-test to establish baseline scores, and at 7 weeks post-test (immediately post intervention for the treatment group). Maintenance of any change was assessed for the treatment group at 6 weeks, and 5 months, post programme completion. Participants missing a session were required to attend a make-up session leading to high completion rates, with regular, reliable, fidelity checks conducted.

Analyses confirmed that there were no significant differences between groups at baseline in age, IQ, ASD symptoms, and outcome variables. Repeated measures mixed-model Group (Treatment vs Control) X Time (pre- vs post-test) MANOVAs were conducted for parent-reported social skills outcomes. A significant Group by Time interaction and follow-up analyses confirmed that the social skills group improved in social competence (SSQ), and social skills (ERSSQ), whereas the control group did not. The treatment effect brought post-test results into the typical (“normal”) range suggesting their clinical significance. The same outcomes were assessed by teachers but a poor response rate (only 7 of the treatment group) precluded reliable investigation of treatment effects.

Two scenario-based tasks assessed children’s (7-11 years) knowledge of emotion management techniques relating to coping with anxiety, and with bullying. A significant Group by Time interaction and follow-up analyses confirmed an improvement in social knowledge in the social skills group participants for both measures, with no significant improvement evident in the control group over time. Finally, two simple emotion recognition tasks relating to facial expression, and posture cues, were also investigated. Both treatment and wait list control groups improved over time with no significant group effect, or Group by Time interaction, found.

Improvements for the social skills group condition for social competence and social skills (SSQ and ERSSQ) were maintained at 6-week and 5-month follow-up in the treatment group.

¹ Note that this study was excluded from the Cochrane systematic review [26] as the intervention included an individual computerised component. However, it was included in the methodological review [24] considered in the current review. As aspects relating to social skills group training were present in every session of the multimodal Junior Detective intervention, the study was also considered eligible for the current review.

This study suggests that the multimodal Junior Detective social skills group intervention can lead to clinically significant improvements in social skills, competence and emotional understanding in children with Asperger syndrome. There were some limitations to the research. The diagnoses were based on reported diagnosis by a paediatrician with some confirmation from an ASD symptomatology tool but these were not independently verified using appropriate diagnostic screening tools. Parent-report measures were unable to be blinded and whilst teacher-reports were, there were too few responses to permit analyses. There was no direct observational data in actual peer interactions to support whether the increases in knowledge around managing emotions was transferred to real life. The improvements evident for both treatment and control groups for the emotion recognition tasks may indicate practice effects, and the authors suggest that the tasks may have been too simplistic and open to ceiling effects.

Finally, as a multimodal intervention, it is not clear whether some components of the Junior Detective Training Program intervention (such as the computer game feature) were key to reported improvements in some social outcomes, or could have been omitted or replaced by other training mediums (such as picture books or filmed scenes).

Peer-mediated social skills group

Koenig et al (2010)

In Connecticut, US, an RCT evaluated a manualised social skills group programme involving trained peer tutors (without ASD) recruited from a local school [31]. Over 16 weeks, small groups of 4-5 children with ASD and two same-aged neurotypical peer tutors were led by two therapists through weekly, 75-minute sessions on social skills training. Training involved socialising activities including playing cooperatively, taking turns, listening to one another, solving a problem together, and tolerating frustration and change. After 3 sessions, individualised treatment plans were developed for each child with specific behaviors targeted for change for the remaining sessions.

The participants were 42 children aged 8-11 years old (mean age of 9 years) with study verified ASD and an IQ of at least 70, almost all of whom were Caucasian, randomised to either treatment group (n=24) or wait list control group (n=18). Treatment fidelity was moderately good and session attendance was high. There were no significant differences between groups at baseline in socio-demographic, diagnostic or outcome variables, although the treatment group did have marginally fewer children receiving medication than in the wait list control (6 vs 10).

Compared with the wait list control group, participants in the social skills group were more likely to be *treatment responders*; that is, being much or very much improved with respect to two target behaviours of social functioning (using the Clinical Global Impressions – Improvement or CGI-I scale). Specifically, at the 16-week post-test compared with baseline, about 70% of treatment group children showed improvement compared with none of the children in the wait list control group over the same period. Notably, whilst the independent raters of children's behaviour were blinded to group

allocation for this measure, their assessment was based on reports on children's behaviour from parents who were aware of whether their children had received treatment, which may have biased their reporting.

A secondary outcome measure was of social competence, with no significant differences found between the treatment of wait list control groups on changes over time in either the pro-social index, or social initiation index of the parent-reported scale. Data was missing for 5 participants however when analyses were re-run without these, the results were consistent with previous analyses.

A study limitation was that IQ scores of participants were drawn from unverified historical records from schools or clinics. There were no direct behavioural assessments of the children.

Affection-focussed social skills group

Andrews et al (2013)

An Australian trial [10] based in a University clinical psychology clinic evaluated a 5-week social skills group intervention which focused on improving giving and receiving signs of appropriate affection. The sample were 58 children aged 7-12 years with ASD (verified by paediatrician and validated assessment tool), all with IQ above 70, who were randomised (post stratification by gender and age) into treatment group (n=29) or wait list control (n=29). Participants were screened pre-selection for parent-reported "difficulties with affection".

Assessments were made at baseline, and 5 weeks post baseline (immediately following treatment for the intervention group) to determine treatment effects. In the treatment group only, an additional assessment assessed maintenance of effects 3 months later.

The social skills programme was of low intensity (10 hours) running 2-hour weekly small group sessions over five weeks. Groups include 3-4 children led by 2-3 trained therapists. Based on cognitive behavioural therapy (CBT) principles, the programme provided information, aided by visual tools, social stories, and role play exercises, to encourage children to develop and practise social strategies with an emphasis on affection. The five sessions included understanding and rating appropriate levels of affection, giving and receiving compliments, managing feelings, and identifying signs that a person needs affection (including facial expression, body language, verbal expressions, and tone of voice). Parents attended concurrent weekly sessions in large group format. Fidelity of the intervention was recorded but adherence was not reported.

No analyses were reported to compare groups at baseline across sociodemographic and diagnostic measures to determine whether randomisation had led to equivalent groups, and information on the ethnicity of the Australian sample was lacking. A repeated measures Group (Treatment vs Control) X Time (pre vs post) MANOVA found significant main effects and a significant Group X Time interaction. Subsequent univariate analyses for the five outcome variables revealed only one significant

interaction. This was for a scale assessing the capacity to engage in appropriate affectionate behaviour outside immediate family (AOQ), an index which did not differ between treatment or control groups at baseline. Analyses for 3 subscales of this measure revealed a significant difference for “giving affection” but not for receiving affection, or communicating empathy. Scores did not change significantly between post-test and 3 month follow-up in the treatment group suggesting results were maintained.

By contrast, no significant treatment effects were observed for the other four parent-rated outcomes including understanding the purpose of affection (WFT), difficulty in giving general affection (GAQ), social competence (SCPQ), or child anxiety (SCAS). Parents were not blind to allocation and no independent ratings or direct observational data was collected.

2.5 Synthesis of results

The current systematic review identified 10 RCTs evaluating social skills groups for children and young people aged 16-21 years with ASD that have been published since 2004, and 2 recently published systematic reviews with overlapping scope to the current review. As apparent from **Table 2.2**, the studies investigated a large and varied number of outcomes, measured by a range of assessment tools, using different types of informants. Given these complexities, to assist in drawing out patterns of results across studies separate tables have been created to present the results relevant to outcome domains separately (see **Tables 2.3 - 2.9**).

Systematic reviews

Two recently published systematic reviews on the topic were appraised as providing background to the current review. The Cochrane collaboration’s systematic review [26] included 5 RCT’s (with 196 participants) evaluating social skills groups in young people aged 6-21 years with ASD published to 2011. The authors concluded that there was emerging evidence for effectiveness of social skills group interventions for some young people with ASD with modest treatment gains found for social competence (4 studies), friendship quality (2 studies) and decreased loneliness (1 study). However there were no treatment effects found for emotional recognition (2 studies), understanding of idioms (one study), or child or parental depression (one study).

A more recent systematic review presented a methodological critique of group-based social skills training (SST) for children and young people with ASD [24]. In contrast to the more narrowly focused Cochrane review, this review included 48 studies across a range of study designs, 13 of which were described as randomised controlled trials (although 2 of these could not be verified as employing fully randomised group allocation). The authors suggested that it would be “presumptive to conduct an efficacy review across all SST interventions given the differences in program design, targeted skills, instructional methods, and the various outcome reported.”

Table 2.3: Characteristics and results of included primary studies for social competence outcomes.

Reference	Quality, country	Intensity	N, mean age (years)	Social competence (assessment tool) (informant)	
				Significant improvement	No significant improvement
PEERS/Children's Friendship Training					
Laugeson et al (2009) [32]	Quality ? US	18 hours	n=33 M age=14.6	Social competence (SSRS) (PARENT)	Social competence (SSRS) (TEACHER)
Frankel et al (2010) [25]	Quality + US	18 hours	n=68 M age=8.4	- <i>Subscales (self-control) (SSRS) (PARENT)</i>	- <i>Subscales (assertion, externalising, internalising) (PARENT) (SSRS)</i>
Gantmen et al (2012) [30]	Quality ? US	14 hours	n=17 M age=20.4	Social competence (SRS) (PARENT) Social competence (SSRS) (PARENT)	
Schohl et al (2014) [13]	Quality + US	21 hours	n=58 M age=13.6	Social competence (SRS) (PARENT)	Social competence (SSRS) (PARENT) Social competence (SSRS) (TEACHER) Social competence (SRS) (TEACHER)
Yoo et al (2014) [35]	Quality + South Korea	21 hours	n=47 M age=14.0	Social competence (VABS) (PARENT) Social competence/ASD symptoms (ADOS) (ASSESSOR)	Social competence (SSRS) (ADOLESCENT) Social competence/ASD symptoms (SCQ) (PARENT) Social competence/Asperger behaviour (ASDS) (PARENT)
Intensive Skillstreaming					
Lopata et al (2010) [33]	Quality ? US	46 hours	n=36 M age=9.5	Social competence (SRS) (PARENT) - <i>Subscales (Withdrawal) (BASC-2) (PARENT)</i>	- <i>Social skills behaviour (BASC-2) (PARENT)</i>
Thomeer et al (2012) [34]	Quality + US	46 hours	n=35 M age=9.3	Social competence (SRS) (PARENT) - <i>Social skills behaviour (BASC-2) (PARENT)</i>	- <i>Subscales (Withdrawal) (BASC-2) (PARENT)</i>
Multimodal skills training with computer game component (Junior Detective Training Program)					
Beaumont and Sofronoff (2008) [29]	Quality + Australia	16 hours	n=49 M age=9.7	Social competence (SSQ) (PARENT) Social competence (ERSSQ) (PARENT)	Social competence (SSQ) (TEACHER) Social competence (ERSSQ) (TEACHER)
Peer-mediated skills training					
Koenig et al (2010) [31]	Quality ? US	20 hours	N=42 M age=9.2	Proportion of responders indicating change in social functioning (CGI-I) (PARENT)	Social competence (SCI) (PARENT)
Affection-focussed skills training					
Andrews et al (2013) [10]	Quality ? Australia	10 hours	n=58 M age=9.0		Social competence (PARENT) (SCPQ)

Social competence

The principle target variable included in all studies was of social competence, using either a generalised index of social skills, or diagnostic assessment of characteristics of ASD and daily functioning as an associated indicator of social behaviour change. **Table 2.3** shows where social competence showed significant improvement following participation in social skills training over time (vs wait list controls), and where it didn't. Where reported, findings relating to subscales of social competence measures are also presented in **Table 2.3** and are italicised for ease of identification.

Excluding the subscale outcomes, there were significant treatment effects found for at least one parent-reported measure of social competence for 8 of the 10 social skills groups evaluations. These included 8 generalised measures of social competence, and 2 measures of adaptive functioning or ASD symptomatology (VAS, CGI-I). There was also a treatment effect for clinician assessed ASD-symptomatic behaviour measured (ADOS). By contrast, there were no parent-reported treatment effects for 3 generalised measures of social competence, or for 2 parent-reported measures of ASD or Asperger characteristic behaviour (SCQ, ASDS).

No improvement was found for those attending social skills groups compared with controls on 5 teacher-reported outcomes for social competence. Notably these outcomes were fraught by poor response rates and indication (in one study) of response bias in baseline characteristics of students for whom data was returned [25].

No significant treatment effects were found for social competence outcomes for the study targeting affection [10] which was a low intensity 10-hour intervention. One study reported only subscales of parent-reported SSRS and found mixed results. There was an improvement in self-control but no treatment effects for assertion, externalising or internalising social skills ratings.

Examination of the findings suggest that there was no clear association with respect to the relationship between treatment effects and study quality or sample characteristics. One study of PEERS [13] found treatment effects on one parent-reported social competence assessment tool (SRS) but not another (SSRS).

Quality of life

A number of outcomes were identified which are relevant to quality of life. However as they are quite distinct constructs they will be discussed separately. Results are presented in **Table 2.4**.

Mood (anxiety and depression) was examined in only 3 studies [10, 13, 35]. Two studies evaluating PEERS, both of good quality and including samples of adolescents of a similar age, found differing results. A US-based study found improvement in social anxiety [13] whilst a South Korean replication study [35] found no affect on child's self-reported state or trait anxiety, but did show improvement on self-reported depression. The Korean study also found improvement in parents' self reported state anxiety, but not for their self reported trait anxiety, or depression. A third study [10] which targeted

Table 2.4: Characteristics and results of included primary studies for quality of life outcomes.

Reference	Quality, country	Intensity	N, mean age (years)	Quality of life (assessment tool) (informant)	
				Significant improvement	No significant improvement
PEERS/Children's Friendship Training					
Laugeson et al (2009) [32]	Quality: ? US	18 hours	n=33 M age=14.6	Friendship quality (FQS) (ADOLESCENT)	
Frankel et al (2010) [25]	Quality: + US	18 hours	n=68 M age=8.4	Loneliness (LS) (CHILD) Friendship quality (perceived popularity) (PHS) (CHILD)	
Gantmen et al (2012) [30]	Quality: ? US	14 hours	n=17 M age=20.4	Social loneliness (SELSA) (YOUNG ADULT)	
Schohl et al (2014) [13]	Quality: + US	21 hours	n=58 M age=13.6	Social anxiety (adolescent) (SIAS) (ADOLESCENT)	Friendship quality (FQS) (ADOLESCENT)
Yoo et al (2014) [35]	Quality: + South Korea	21 hours	n=47 M age=14.0	Depression (CDI) (ADOLESCENT) Parent's state anxiety (STAI-S)* (PARENT)	State or trait anxiety (STAI-S/T) (ADOLESCENT) Parent's trait anxiety (STAI-T) (PARENT) Parent's depression (BDI) (PARENT)
Affection-focussed skills training					
Andrews et al (2013)	Quality: ? Australia	10 hours	n=58 M age=9.0		Symptoms of child anxiety (SCAC) (PARENT)

Table 2.5: Characteristics and results of included primary studies for social communication outcomes.

Reference	Quality, country	Intensity	N, mean age (years)	Social communication (assessment tool) (informant)	
				Significant improvement	No significant improvement
Intensive Skillstreaming					
Lopata et al (2010) [33]	Quality: ? US	46 hours	n=36 M age=9.5	Idiomatic language (CASL) (CHILD)	
Thomeer et al (2012) [34]	Quality: + US	46 hours	n=35 M age=9.3	Idiomatic language (CASL) (CHILD)	

the appropriate giving and receiving of affection in younger kids using a low intensity intervention (10 hours) found no impact on parent-reported child anxiety.

Four studies evaluating PEERS considered other quality of life outcomes relevant to the study participant's perceived loneliness, popularity, and friendship quality. With respect to loneliness, two studies found that those receiving PEERS/CFT reported reduced loneliness following the intervention, one study [25] included young children (mean age=8) and the other [30] included young adults (mean age=20).

Three PEERS/CFT studies investigated the effect of social skills groups on friendship quality. Two found positive treatment effects: a study of uncertain quality including adolescents [32] and a good quality study of children [25]; the other good quality of adolescents found no treatment effect [13].

Social communication

Only two studies, both evaluating the intensive Skillstreaming intervention in children, considered outcomes specifically related to social communication (see **Table 2.5**). Both found that measures of idiomatic language (the use of expressions that don't mean the same as their literal meaning) improved relative to controls following social skills group training.

Problem behaviours

Three studies, all evaluating PEERS/CFT interventions, included measures assessing problem behaviours (see **Table 2.6**). Of two considering general problem behaviour outcomes in adolescents following PEERS, one good quality study found reduced problem behaviours (according to parent and teacher report) [13] whereas the other study of uncertain quality [32] offering slightly less intense PEERS (18 hours of 21 hours) found no treatment effect for parent-, teacher-, and self-reported problem behaviour outcomes.

A third evaluation of children receiving Children's Friendship Training [25] found no treatment effect in outcomes for a teacher-reported measure of withdrawal and aggression.

Emotion recognition

Performance measures of emotion recognition were assessed in three studies of children with ASD. Two evaluations of intensive Skillstreaming interventions [33, 34] assessed children's reported recognition of emotions expressed by faces displayed via computer, and an evaluation of the Junior Detective multimodal intervention measured self-reported recognition of emotions shown in pictures of faces and posture. All three studies found no difference in emotion recognition in any measure (see **Table 2.7**).

Table 2.6: Characteristics and results of included primary studies for problem behaviour outcomes.

Reference	Quality, country	Intensity	N, mean age (years)	Problem behaviour (assessment tool) (informant)	
				Significant improvement	No significant improvement
PEERS/Children's Friendship Training					
Laugeson et al (2009) [32]	Quality: ? US	18 hours	n=33 M age=14.6		Problem behaviour (SSRS) (ADOLESCENT) Problem behaviour (SSRS) (PARENT) Problem behaviour (SSRS) (TEACHER)
Frankel et al (2010) [25]	Quality: + US	18 hours	n=68 M age=8.4		- withdrawal, aggression (PEI) (TEACHER)
Schohl et al (2014) [13]	Quality: + US	21 hours	n=58 M age=13.6	Problem behaviours (SSRS) (PARENT) Problem behaviours (SSRS) (TEACHER)	

Table 2.7: Characteristics and results of included primary studies for emotional recognition, emotion management, empathy, and affection

Reference	Quality, country	Intensity	N, mean age (years)	Emotional recognition, Emotion management, Empathy, Effecton (assessment tool) (informant)	
				Significant improvement	No significant improvement
PEERS/Children's Friendship Training					
Gantmen et al (2012) [30]	Quality: ? US	14 hours	n=17 M age=20.4	Empathy (EQ) (PARENT)	
Intensive Skillstreaming					
Lopata et al (2010) [33]	Quality: ? US	46 hours	n=36 M age=9.5		Emotion recognition (DANVA-2) (CHILD)
Thomeer et al (2012) [34]	Quality: + US	46 hours	n=35 M age=9.3		Emotion recognition (DANVA-2) (CHILD)
Multimodal skills training with computer game component (Junior Detective Training Program)					
Beaumont and Sofronoff (2008) [29]	Quality: + Australia	16 hours	n=49 M age=9.7	Emotion management (anxiety) (ASSESSOR) Emotion management (bullying) (ASSESSOR)	Emotion recognition: facial expression (ASSESSOR) Emotion recognition: posture cues expression (ASSESSOR)
Affection-focussed skills training					
Andrews et al (2013) [10]	Quality: ? Australia	10 hours	n=58 M age=9.0	Affection for others (AOQ) (PARENT)	Assess purpose of affection (WFT) (PARENT) Excessive or inadequate affection (GAQ) (PARENT)

Emotion management

The trial of the Junior Detective multimodal intervention also included measures of the management of emotions through the telling of two stories relating to coping with bullying and anxiety. A treatment effect was demonstrated suggesting that participation in the social skills group aided understanding of emotion management in these contexts (see **Table 2.7**).

Empathy

Empathy was assessed in one study of uncertain quality evaluating PEERS in young adults [30]. A positive treatment effect was found suggesting social skills groups may improve empathy (see **Table 2.7**).

Affection

A single study [10] of uncertain quality targeting affection considered three affection-related outcome measures (see **Table 2.7**). A treatment effect was found for one parent-reported outcome relating to giving appropriate affection to people outside the family (AOQ). Analyses of subscales revealed a significant difference for “giving affection,” but not for receiving affection, or communicating empathy. There were no significant differences for the two other affection measures: understanding the purpose of affection, or judging the appropriate level (excessive or inadequate) of affection.

Quality of play/interaction

All five PEERS/CFT evaluations included a version of a questionnaire aiming to measure the quality of get-togethers or “play-dates” (QSQ, QPQ). These asked parents or their children/young adults to report how many get-togethers were hosted or they were invited to, and where a gathering/s had occurred, whether there was conflict, and what sort of interaction occurred. Whilst there was no treatment effect for reduced conflict in any study, there were inconsistent effects between studies, and parents’ ratings often differed to childrens’ ratings within studies (see **Table 2.8**). Whether significant effects on this variable were found or not did not appear to relate to study quality or sample characteristics.

Programme specific skills/knowledge

Six of the 10 appraised primary studies evaluated knowledge relating to social skills taught in the programme. Improvements in programme knowledge of the taught skills was demonstrated for all studies including PEERS interventions [13, 30, 32, 35], and, through two separate measures, the two intensive Skillstreaming evaluations (see **Table 2.9**).

Table 2.8: Characteristics and results of included primary studies for quality of play during get-togethers.

Reference	Quality, country	Intensity	N, mean age (years)	Quality of play during get-togethers (assessment tool) (informant)	
				Significant improvement	No significant improvement
PEERS/Children's Friendship Training					
Laugeson et al (2009) [32]	Quality: ? US	18 hours	n=33 M age=14.6	Hosted get-togethers (QPQ) (ADOLESCENT)	Invited, conflict, get-togethers (QPQ) (ADOLESCENT) Hosted, invited, conflict, get-togethers (QPQ) (PARENT)
Frankel et al (2010) [25]	Quality: + US	18 hours	n=68 M age=8.4	Disengaged activity (QPQ) (PARENT) Hosted get-togethers (QPQ) (PARENT)	Invited, conflict, engagement in, get-togethers (QPQ) (PARENT)
Gantmen et al (2012) [30]	Quality: ? US	14 hours	n=17 M age=20.4	Hosted, invited, get-togethers (QSQ) (PARENT)	Hosted, invited to get-togethers (QSQ) (ADULT)
Schohl et al (2014) [13]	Quality: + US	21 hours	n=58 M age=13.6	Hosted, or invited, get-togethers (QSQ) (ADOLESCENT)	Hosted, invited, conflict at get-togethers (QSQ) (PARENT) Conflict at get-togethers (QSQ) (ADOLESCENT)
Yoo et al (2014) [35]	Quality: + South Korea	21 hours	n=47 M age=14.0		Hosted, invited, conflict at get-togethers (QPQ) (PARENT) Hosted, invited, conflict at get-togethers (QPQ) (ADOLESCENT)

Table 2.9: Characteristics and results of included primary studies for knowledge of programme-specific skills.

Reference	Quality, country	Intensity	N, mean age (years)	Social competence (assessment tool) (informant)	
				Significant improvement	No significant improvement
PEERS/Children's Friendship Training					
Laugeson et al (2009) [32]	Quality: ? US	18 hours	n=33 M age=14.6	PEERS social skills knowledge (TASSK) (ADOLESCENT)	
Gantmen et al (2012) [30]	Quality: ? US	14 hours	n=17 M age=20.4	PEERS social skills knowledge (TYASK) (ADULT)	
Schohl et al (2014) [13]	Quality: + US	21 hours	n=58 M age=13.6	PEERS social skills knowledge (TASSK) (ADOLESCENT)	
Yoo et al (2014) [35]	Quality: + South Korea	21 hours	n=47 M age=14.0	PEERS social skills knowledge (TASSK) (ADOLESCENT)	
Intensive Skillstreaming					
Lopata et al (2010) [33]	Quality: ? US	46 hours	n=36 M age=9.5	Skills taught in programme (ASC) (PARENT) Skillstreaming programme knowledge (SKA) (ASSESSOR)	
Thomeer et al (2012) [34]	Quality: + US	46 hours	n=35 M age=9.3	Skillstreaming programme knowledge (SKA) (ASSESSOR)	

Maintenance

Only 4 of the 10 primary studies assessed whether treatment effects were maintained at longer follow-up (6 weeks to 5 months) beyond the initial post-test assessment. Results were positive for two trials. For the study evaluating PEERS in South Korea [35], treatment effects were maintained for outcomes after 3 months with the exception of knowledge of PEERS skills, although these were still higher than at baseline. The trial of the affection-targeting social skills intervention [10] found no difference between post-test and 3 month follow-up scores suggesting treatment effects were generally maintained, although attrition at follow-up was 21%.

More mixed results were found for the multimodal Junior Detective intervention [29]. Treatment effects were maintained at 6 weeks and 5 months post baseline for social competence, and social skills, but not for emotion recognition and emotion management knowledge. Also variable were the findings for one of the intensive Skillstreaming evaluations [34] where parent-reported taught skills, and social skills, were maintained after 2-3 months, but the post-test treatment effects for parent-reported ASD social impairments, and social withdrawal, were not maintained at extended followup.

In addition to these studies, a follow-up study [36] of 24 of 66 study participants from the Children's Friendship Training evaluation [25] investigated maintenance of treatment effects 1-5 years post intervention (mean age of 12.6 years). Whilst not formally appraised as the data is uncontrolled, follow up data is of interest. Compared to baseline, at longer follow-up the participants were reported by their parents to have been invited on significantly more play dates and showed less conflict during them, and the children reported having increased social skills with fewer problem behaviours. However the study is marred by considerable attrition (many participants did not consent to followup at baseline), and the fact that these outcomes are uncontrolled and may improve over time with maturation.

2.6 Limitations and future research directions

Introduction

A systematic review by White and colleagues [9] identified 14 studies evaluating group-based social skills training programmes for school-aged children and adolescents with ASD published between 1985 and 2005. The review concluded that the empirical support for the approach based on several small initial efficacy studies was incomplete and provided detailed suggestions for designing future effectiveness RCTs. Notably, whilst appraising older studies, none would have met inclusion criteria for the current review due to methodological limitations including lack of control group, lack of randomisation, and/or small sample sizes.

The current review identified 10 primary studies which met a greater threshold of methodological criteria. The current evidence base represents a significant advance over the last decade, taking up many of the suggestions of the White et al review [9]. This includes employing randomised controlled designs, the increasing manualised of interventions, multi-site replication studies using the same intervention, increased use of independent outcome assessors, and follow-up assessment to assess maintenance. However limitations continue to limit the evidence for effectiveness and generalisability of findings. Key issues will be outlined below with suggestions on directions for future research.

Sample size

The 10 primary studies appraised in this review considered 443 participants with ASD. The studies were all relatively small, ranging from 17-68, particularly once randomised to treatment and control groups. The smaller sizes may not yield sufficient statistical power to consistently detect smaller treatment effects [33] and unstable or significant changes may go undetected [24]. Future research should include randomised experimental designs with adequate power (larger samples) to detect clinically significant effects. Multi-site approaches are likely to be needed in order to achieve the larger samples required to evaluate treatment effects and conduct sensitivity analyses robustly.

Sample characteristics and recruitment

The studies included in the current review predominantly concerned higher functioning, verbal children and young adults with ASD. Inclusion criteria for all appraised studies included that participants demonstrated cognitive and/or language functioning above specified thresholds (typically, composite IQ scores above 70) as assessed using standardised scales. More research is needed into the effectiveness of social skills groups for individuals with lower functioning cognitive and verbal levels.

The samples were predominantly primary school aged-children (mean age 8-10 years) in six of the included trials, or adolescents (mean age range 14-15 years) in three trials, and in one trial adapting PEERs for young adults, participants were aged between 18-23 years (mean age of 20 years). In initial scoping for the review, no eligible studies were identified investigating the impact of social skills groups in older adults and so study eligibility was restricted to children and young people as a comparable population.

Samples were predominantly male (range: 70-96%), broadly reflecting the 4:1 male to female gender distribution observed in the condition at a population level. As there may be differences in the way social skills deficits present in girls compared to boys, research including larger samples of female participants may be useful.

In four of the seven US based trials, the high majority of the sample were Caucasian [13, 31, 33, 34], but a more varied spread of ethnicities was represented in the other 3

US studies [25, 30, 32] including Caucasian (42-66%), Asian, Hispanic, and others. Ethnicity was not reported for the 3 studies undertaken outside the United States, including Australia [10, 29] and South Korea [35]. No studies were undertaken in New Zealand.

The sample characteristics need to be considered in applying the results of this review to populations with different characteristics as the findings may not generalise to different groups and countries with different social norms. The effectiveness of social skills group curriculum is likely to depend on it being culturally relevant and appropriate for the target audience. In the Korean trial of PEERS [35], examples of social situations and suggested phrases needed to be modified to make them relevant to the students. Interventions may similarly need careful adaptation when applied to cultural and ethnic populations found in New Zealand.

In future research, samples need to be broader to include pre-schoolers, adults, people using AAC, people with intellectual disability, and people from different cultures to permit understanding of the generalisability of the findings to different populations. Cross-cultural adaptation and validation is required for interventions which have shown some efficacy in order for them to be established as effective in broader social and cultural contexts.

The review was open to research in any context, and did not formally exclude non-manualised programmes. However, the evidence base was dominated by research in clinical and research centre settings. For several studies, methods of recruitment were related to research centres and clinics already known for developing social skills group training. Such approaches may be more likely to attract families who are particularly motivated to seek help in this area, and those of higher socio-economic status willing and able to participate in research, including having a parent who is able to volunteer significant periods of time to bring their child to multiple sessions, attend sessions themselves, and complete screening and outcome assessments. The feasibility and social validity of community based settings for social skills groups should be further explored, including schools where the training may be provided by special education teachers, guidance counsellors or speech-language therapists [23].

Some studies also required a degree of social skills deficits as a requirement for inclusion. Parent-reported social problems was an inclusion criteria for four PEERS studies [13, 30, 32, 35], and care-giver reported difficulties with affection was a requirement for participants in the affection-focused intervention study [29].

By contrast, participants eligible for the PEERS-related Children's Friendship Training trial [25] needed to demonstrate minimum social skills (whilst also having significant social challenges). They needed to be able to switch topics in a conversation, have capacity for joint attention and basic social reciprocity, and have knowledge of rules of at least common board and outdoor games. This approach was employed as the researchers' experience suggested that the intervention had more success with children who had a basic grasp of certain social skills to begin with. These

requirements may limit the ability of the findings to be applied to a broader population of young people with ASD.

Diagnosis of ASD

Most studies independently verified reported ASD diagnosis through the use of gold-standard autism diagnostic instruments. Three studies [30, 32, 33] relied on parent-reports of previous clinical diagnoses, though in one of these [33], documentation of diagnosis was available for most participants. Historic reporting of diagnoses may be inaccurate, no longer current, or not based on clinician interview without a full diagnostic assessment using standardised screening tools.

Control groups

All studies appraised in this review used wait list controls where treatment is deferred, although “no intervention” or “usual care” were eligible comparators. The ethical benefit of this approach is that participants in the control group are not deprived of the potential benefit of an intervention, especially for studies validating interventions previously shown to be promising. The wait list control approach also offers controls a reason to stay enrolled in the study and not drop out, and an incentive to complete the many time-consuming baseline and outcome measures. A further benefit is that the control group receive baseline and follow-up assessments at an equivalent time to the treatment group controlling for any time-relevant or maturation effects, as well as practice or other effects of having undertaken baseline measures that bring attention to social problems.

However there are down-sides to the use of wait list controls. Firstly, the control group during comparison “post” assessments have received no additional intervention and therefore do not receive “expectancy” effects [33]; that is, the possible effect of having any active treatment, no matter what the content, is not controlled for (as it is in a placebo-controlled trial). Such performance biases can also affect parent-completed outcomes. It is also not possible to determine whether indirect aspects of the intervention, such as gathering together in a group of similar children on a regular basis, may be contributing to social improvements apart from the programme content itself. Studies which provide alternative interventions where groups meet for a similar intensity but do not receive social skills training attempt to control for such biases [25].

Outcome measures

The 10 studies included in this review measured a large array of socialisation and quality of life outcomes. However there is little consensus on appropriate, reliable and valid outcome measures for social skills groups that are practical to apply and sensitive to change [9]. Ideally, development of an agreed set battery of assessment tools is needed [24]. These should include multiple instruments that measure both general areas of social skills (such as social competence and ASD symptomatology) as well as

more specific skills targeted by the intervention [33]. Some outcome measures employed by included studies were not designed for the ASD population (e.g., the SSRS) or for young adults [30], which may reduce their ability to identify change on relevant factors.

There has been minimal examination of the degree to which learned skills generalise into the “real world”. Beyond the learning of a specific taught skill, outcomes need to establish the degree to which a person uses, extends and adapts skills into new environments [9]. Researchers have called for greater use of direct behavioural assessments made by independent observers blinded to group assignment and study hypotheses [9]. This would involve coding of observed socially appropriate behaviours with peers in naturalistic contexts (e.g., with classmates in the playground). These provide ideal proximal indicators of taught skills and have a high degree of ecological validity [24].

Assessing social skills across settings using ratings from multiple informants such as the child, parent, and teacher should become the standard in this field [9]. The design of group-based social skill studies is such that third party raters such as teachers and independent clinicians/assessors are needed to enable assessment blind to group allocation and reduce the threat of detection and performance biases. Some studies in the current review included teachers who were blinded to group assignment [13, 25, 32], although unfortunately response rates were often poor for these informants [13, 29, 32].

Although some measures may be theorised as precursors or intermediary factors to behaviour change, they may be weakly associated with desired social outcomes in themselves. For example, one would expect an increase in hosted get-togethers to follow from an intervention where parents are actively encouraged to organise these as part of the training curriculum and set homework tasks [25]. However whether organised play dates in themselves lead to improvements in the quality of social exchanges needs to be established to demonstrate whether this outcome is meaningful.

Assessment and maintenance

For all 10 primary studies, outcome assessments were measured at baseline (pre-test) and at post-test, either upon completion of the last social skills training session or for the control group, after an equivalent time post baseline. Such timing is likely to maximise any short-term effects of the intervention and may exaggerate some indicators due to recency effects, such as programme knowledge.

Any longer term follow up for the treatment group was not able to be compared with control group assessments at the same time as the wait list control had begun their social skills treatment in that period so as to reduce the ethical risks of delaying any further a potentially useful treatment. However longer follow-up was able to suggest whether any treatment effects identified at post-test were maintained over time. Such maintenance follow-up was conducted for only 4 of the 10 studies [10, 29, 34, 35].

Longer term follow-up for one PEERS study [25] was undertaken 1-5 years after treatment [36]. There were mixed results, and the study suffered from high attrition, partly due to lack of informed consent initially when the followup study had not been envisioned.

Study design

The current review's objective was to establish the effectiveness of social skills groups for children and young people with ASD. Well conducted randomised controlled trials (RCTs) are considered the most robust test of intervention effectiveness consistent with the hierarchy of evidence, and the review was restricted to including studies of this design as offering the "best evidence" to answer the research question. Further, the intervention being provided in a group format would seem to make a group analysis particularly appropriate.

However, the group design, sample size, and cost of RCTs make them arguably less amenable to exploring why an intervention works, who it works best for, and what components are necessary. By contrast, single case experimental design (SCED) studies can be invaluable in the development, fine-tuning, adaption, and improvement of an intervention. Future research should explore these issues through the use of such designs in addition to the use of statistical and sub-group approaches within RCTs.

Moderators and mediators of treatment response

There are a number of factors that may moderate and mediate a treatment response that should be measured and reported, and ideally, experimentally manipulated in a controlled manner to determine their potential influence.

Programme standardisation and fidelity

Whilst the review did not formally exclude non-manualised programmes, in an improvement over earlier research [24], all programmes evaluated in the studies included in this review followed manualised procedures and were led by trained instructors/therapists, though this was not a criteria for study eligibility. The fidelity of programmes in providing scheduled content was also recorded in all appraised studies through checklists. However the degree to which sessions adhered to the manual was only reported for three studies: the two evaluations of intensive Skillstreaming [33, 34] and that of the social skills groups involving peer tutors [31] where fidelity to programme curricula was found to be moderate to high (75%-97%). Reliability of records of fidelity was rarely investigated. Future researchers should ensure fidelity is not only monitored but adherence to programme manuals reported reliably.

Manualised programmes attempt to standardise the programme format and content which allows for more consistent transfer of the intervention into a new setting and/or for a new population. Standardisation does not necessarily mean uniformity or inflexibility. Programmes can be multi-component, and can permit variation in aspects

such as choice of scenarios that are modelled, language employed, and social behaviours rehearsed. Some programmes evaluated in this review were adapted for use with an ASD population, for older youth, or for a different ethnic group.

That all programmes in this review were manualised shouldn't be taken to mean that non-manualised approaches are ineffective, or should deter practitioners from developing their own social skills groups programmes. Such attempts are likely to be informed by the components identified in the manualised programmes appraised in the current review.

Programme intensity

The intensity of an intervention is indicated by how long each session is, how often, and over how many weeks. Intensity may be associated with treatment effects indicating a dose-response relationship. In the current review, intensity was one of the prominent variations between intervention types. The most intensive intervention was for Skillstreaming [33, 34] involving five 70-minute sessions per day over 5 weeks (total of 46 hours). Moderate intensity interventions of 14-21 hours included PEERS/CFT [13, 25, 30, 32, 35], peer assisted groups [31] and the multimodal Junior Detective intervention [29]. The programme exhibiting the least intensity (10 hours) was the affection-focussed intervention involving 2 hourly sessions held weekly over only 5 weeks [10]. As the intensity of these interventions also varied characteristics of their content and approach, it is difficult to disentangle the impact of intensity from content in treatment response.

Programme components

Comparisons of different programmes can determine whether some programmes or programme components are more effective than others. It would be useful to isolate and analyse components from more comprehensive programmes to determine specific contributions to overall efficacy [34]. Such components include parental involvement, use of peer tutors, and different media for delivering content.

In all but the social skills intervention including peer tutors [31], parents received separate, and usually concurrent, information sessions to assist them in supporting their children's social skills development in the home. It appears that parental involvement is considered a key component of most social skills group interventions so that participants can be encouraged to practise their new skills at home and complete homework tasks.

Subject characterisation

Sociodemographic, diagnostic, and social skills of participants at baseline may relate to treatment outcomes in predictable ways, characterising treatment responders and non-responders. Identifying characteristics of individuals for whom the intervention is most likely to work will assist the tailoring and targeting of interventions.

One study attempted to examine the predictors of positive social skills following participation of individuals in social skills groups [37]. Sixty adolescents with ASD (mean age of 14.7 years) enrolled in previous PEERS efficacy studies between 2008 and 2011 (source studies not cited) were included. Multiple regressions found two significant baseline predictors accounting for 63% of the variance in individual's social skills scores after the intervention. The results revealed that adolescents with higher parent-reported social skills (on SSRS Social Skills, particularly responsibility and self control subscales) and lower self-reported perceived social functioning (Piers-Harris Popularity scale) at baseline demonstrated greater improvement in social skills following participating in PEERS. Notably, baseline age, IQ and Vineland-II Communication subscales scores were not significant predictors.

The results suggest that teenagers who have some fundamental social skills and awareness of their social deficits prior to the social skills group intervention benefit most. The authors suggest that being able to respond to authority figures and to self-regulate during conflicts may facilitate effectiveness of the PEERS programme, particularly for those motivated and committed to learning new skills. More research with a greater range of potential predictor variables and social skills group interventions is needed to further explore predictors of treatment success, and lack of success, and permit better targeting of interventions.

Consideration of a range of statistical and sub-group analyses of social skills group studies is needed. As social skills are taught in a group format, it is important to take into account intra-group correlation [12]. For example, hierarchical linear modeling (HLM) techniques model the similarity of participants within each individual treatment group. Such approaches require larger samples to permit the analysis of several groups receiving a programme [24].

Additional strategies

Severe behavioural issues, emotion regulation and mood difficulties may hinder the acquisition of new skills [10]. Strategies aimed at intervening on such factors may be helpful if integrated into social skills group training. For example, programmes could be augmented by approaches to overcome social anxiety, such as using distraction when having intrusive/repetitive thoughts [25]. Functional analysis of interfering behaviours could be employed to identify maintaining factors and incorporated into a group teaching approach [9]. Further research is needed to determine whether such new components would increase effectiveness of social skill group interventions.

Variability in results

Before considering the summary of results from the appraised studies in the current review and drawing conclusions, it is worth considering the meaning of variability in results between and within studies of the same outcome construct and the potential value of finding such variation.

Traditionally, particularly in psychopharmacological research, there is an expectation that an effective intervention should lead to movement of all indices assessing the

same construct (including those reported by different informants) in an expected positive direction and to a similar degree. Any inconsistency is interpreted as either error variance or lack of treatment efficacy. However this view has been challenged by Koenig et al [12] in considering the efficacy of social skill group interventions. They argue that multidimensional constructs such as social competence/reciprocity are extraordinarily complex, and dynamic, affected by different factors and at different times. As such, the authors suggest there may be little utility in expecting uniform improvement across outcomes. It is hypothesised that small, incremental improvement in one outcome (such as speech prosody) may be associated with a large improvement (e.g., in peer acceptance).

Further, there is an argument against the common practice of reducing the target of an intervention to a single “primary” outcome measure [12]. Using such an approach researchers may not be able to capitalise on the depth of information that inconsistent information (variation in treatment effect) between outcomes and informants can offer as to how and when interventions work. In this regard, one may expect variation in reports of the same outcome domain to come from the self-report of an affected individual compared with their parent or teacher due to variations in their perspectives and the context (including time, place and culture) in which the behaviour is expressed or observed. Accepting this variability through the use of multiple methods, measures and informants will lead to the qualification of the results of any individual study whilst allowing for the exploration of variability in the response of people on the spectrum to social skills groups. Such a view accepts and examines the inherent complexity of this research area.

2.7 Summary and conclusions

Overview

This systematic review updates evidence for the New Zealand Autism Spectrum Disorder Guideline [1] with respect to the effectiveness of social skills groups in children and young people with ASD aged 6-21 years. Following a comprehensive database search and reference checking of primary studies and systematic reviews published since 2004, 12 studies met selection criteria for inclusion: 2 systematic reviews and 10 primary studies.

Included primary studies were randomised controlled trials comparing outcomes for individuals receiving social skills group training with those in a wait list control group. Samples were required to have a minimum of 10 participants receiving the intervention. Included studies represented 443 participants with ASD ranged from 17 to 68 participants in individual studies. The majority of the studies considered children (mean age 8-10 years), 3 concerned mid-teens (mean age 14-15 years, and one included young adults aged 18-23 years.

Most studies were undertaken in the United States with two in Australia and one in South Korea, and included mostly caucasian English-speaking male participants. A more mixed range of ethnicities was represented in three studies.

The social skills group interventions were generally undertaken in clinical settings and interventions were broadly categorised into five subtypes of varying intensity. There was the high intensity (46 hours) Skillstreaming approach involving multiple daily sessions (2 studies). A range of moderate intensity (14-21 hours) approaches included Program for the Education and Enrichment of Relational Skills (PEERS) or the closely related Children's Friendship Training (CFT) intervention (5 studies); a multimodal Junior Detective Training Programme (JDTP) including computer-based content delivery (1 study); and a peer-mediated intervention including neurotypical peer tutors (1 study). Finally, there was a low intensity (10 hours) cognitive-behavioural intervention focused on improving demonstration and reception of affection (1 study).

All but the peer mediated intervention included weekly parent sessions and most had a homework component. All programmes were manualised and included checklists to monitor fidelity, although this was reported in only 3 studies and found to be moderate to good.

The quality of the studies was variable. According to validated GRADE checklists, 5 studies were rated as being of uncertain quality, and 5 of good quality. Given the intervention, it was not possible to blind the group participants or their parents from group allocation, and therefore the reported outcomes from these sources which make up the majority of findings were open to performance and detection biases. Some teacher-reported measures were blinded but tended to suffer from poor response rates and evidence from one study of a response was biased in favour of less aggressive students at baseline [25].

Summary of main findings

Social competence

All studies included at least one measure of social competence, including generalised scales of this construct, as well as associated measures of behavioural characteristics of ASD and/or adaptive functioning. Social competence significantly improved following social skills group intervention for *at least one* of these parent or clinician-reported measures in 8 of 10 studies compared with wait list controls. Results were inconsistent in that 3 of these 8 studies also reported no treatment effects for other social competence outcome/s also measured.

The 2 studies finding no significant treatment effects in any measure of general social competence/ASD symptoms/adaptive functioning included a low intensity 10-hour intervention targeting affection, and a trial of Childhood Friendship Training which did not include generalised measures. In that study, subscale outcomes for social competence led to mixed results, with improvement only evident for self-control following the intervention.

There were no treatment effects for any of the teacher-reported outcomes assessed in 5 studies (although these were often hampered by poor response rates).

Overall, the majority of studies suggest improvement in social competence following participation in a social skills group as reported by parents or clinician.

Quality of life

It is not possible to draw conclusions from the small and inconsistent evidence base of 3 studies reporting on the impact of social skills groups on anxiety and depression for both children/young people and parents.

From a small evidence base evaluating PEERS interventions, there was some evidence of the benefit of PEERS in reducing loneliness (in 2 studies), and mixed evidence from 3 studies of the effect of social skills groups on friendship quality or perceived popularity (2 finding positive treatment effects, one not).

Social communication

Only one dimension of social communication was considered, a measure of the use of idiomatic language, which found an increase following participation in the intensive Skillstreaming social skills groups relative to controls in 2 studies.

Problem behaviours

Regarding problem behaviours, there were too few studies (n=3), too much variability in findings, and a lack of direct behavioural assessment to make any firm conclusions about whether challenging behaviours may be impacted by attendance of social skills groups. However evidence of treatment effect from one good quality PEERS study of adolescents suggests this may be a variable worthy of further investigation.

Emotion recognition, management and empathy

Three studies included performance measures of emotion recognition of children with ASD. Two evaluations of intensive Skillstreaming interventions, and an evaluation of the Junior Detective multimodal intervention, found no difference in emotion recognition in faces and/or posture.

A single study evaluating the trial Junior Detective multimodal intervention found that participation in the social skills group aided understanding of emotion management (coping with bullying and anxiety).

Empathy was found to improve following PEERS in young adults in one study. Whilst promising, this finding should be taken with caution given the lack of research, and direct behavioural assessment was not undertaken.

Affection

Affection was targeted and measured in 3 parent-reported outcomes in a single, low intensity study of uncertain quality. Of these, a treatment effect was found for only one

outcome relating to giving appropriate affection to people outside the family (AOQ), and specifically in the subscale of “giving affection.” Understanding the purpose of affection, or judging the appropriate level (excessive or inadequate) of affection were not affected. The study suggests a possible improvement in giving affection to those outside the family from social skills group training. This outcome is worthy of investigating in studies of greater intensity.

Quality of play

The 5 studies evaluating PEERS/CFT interventions included a measure of quality of play relating to the number of quality of get-togethers hosted and invited to. It is not possible to draw conclusions given the inconsistency in treatment effects found between studies, as well as within studies between raters (parent and child/young person).

Programme specific skills/knowledge

All six studies which measured knowledge of social skills taught in the programme found improvements following the intervention compared to controls. Four considered PEERS interventions and two evaluated intensive Skillstreaming interventions. These consistent findings suggest that social skills groups are effective at increasing knowledge of specific skills that have been targeted in the intervention’s curriculum.

Maintenance

Outcomes were followed up beyond the initial post-test assessment for the intervention group in 4 of the 10 primary studies appraised to assess maintainance over time, ranging from 6 weeks to 5 months post intervention.

Significant treatment effects were generally maintained 3 months post intervention for two studies. For a third study, treatment effects were maintained for social competence and social skills, but not emotion management, after 5 months followup. Mixed results were found for a fourth study after 2-3 months extended followup.

In addition to appraised studies, a follow-up after 1-5 years for the CFT evaluation [36] found maintained treatment effects for a number of outcomes. However there was considerable attrition from baseline.

Together, these results suggest the possibility of longer term maintenance of improvements to relevant outcomes following participation in social skills groups.

Systematic reviews

The Cochrane collaboration’s systematic review [26] of 5 RCT’s (with 196 participants) found “emerging evidence for effectiveness” of social skills group interventions for some young people with ASD. Modest treatment gains were reported for social competence, friendship quality, and decreased loneliness, with no treatment effects found for emotional recognition, understanding of idioms, or child or parental depression.

More recently, a systematic review [24] was conducted of 48 studies of varying study designs, including 13 described as RCTs. The review provided methodological critique of the evaluation of social skills groups but did not attempt to provide an efficacy review given the variations in study designs, targeted skills, methods and outcomes of the evidence base.

Conclusions

The current systematic review identified 10 RCTs evaluating social skills groups for children and young people aged 16-21 years with ASD published since 2004. It also considered two recently published systematic reviews which were considered as providing background to the current review.

Key findings were:

- Overall, there appears to be reasonable evidence from the majority of studies of improvement in social competence following participation in a social skills group as reported by parents or clinician. However this conclusion is tempered by the appraised research occurring mainly in clinical/academic settings, and generally being evaluated without direct behavioural observation from independent, blinded assessors in naturalistic environments.
- There was consistent evidence that knowledge of specific social skills taught in the social skills group programme increased following the intervention compared to controls (in 6 studies).
- There was evidence from a small number of studies that participation in social skills groups may increase use of idiomatic language (in 2 studies of intensive interventions extending over 46 hours), and reduce loneliness (in 2 studies).
- Results from single studies suggest that social skills groups can lead to increased understanding of emotion management (coping with bullying and anxiety), empathy, perception of popularity, and increased giving of affection to those outside the family.
- No differences were found between social skills groups and wait list control groups for emotional recognition (in 3 studies), or in single studies for understanding the purpose of affection, or judging the appropriate level of affection.
- The evidence base was too small and/or results too inconsistent to allow conclusions regarding the impact of social skills groups on anxiety and depression (of child/young person or parents), friendship quality, problem behaviours, or quality of play.
- Maintenance of treatment effects were investigated between 6 weeks and 5 months post completion of the intervention in 4 studies. There is some evidence to support the possibility of longer term maintenance of improvements to relevant outcomes from social skills group. However reduced response rates, and the lack of control for possible maturation and learning effects suggest

these findings from a small number of interventions of varying content should be treated with caution.

- These findings are broadly consistent with the conclusions of previous systematic reviews of social skills groups. However whilst the Cochrane review reported improvement in friendship quality from both of the 2 RCTs including it as an outcome, the current review found no treatment effect for friendship quality in an additional trial leading to a more mixed evidence base for this outcome.

Many of the included studies covered a relatively narrow age range, gender, cognitive profile, and ethnicity, making generalisability to pre-schoolers, young adults, girls/women, people with below average cognitive abilities, and New Zealand-relevant cultures uncertain.

The current evidence base suggests the potential for social skills groups to broadly enhance social competence, increase the knowledge of specific social skills that have been directly taught, and possibly increase the use of idiomatic language, empathy, giving of affection beyond the family, emotion management, and perceived popularity, and possibly reduce loneliness.

However the complexity of the evidence base makes evaluating effectiveness of social skills groups challenging given the variability in programme content, programme approach and intensity; imprecision in the large range of outcomes measures for similar outcomes; the range of outcome variables considered; and the variability in findings for some outcomes. Given these issues, it is not currently possible to offer clear conclusions about the *necessary content, approach and intensity* of social skills groups. Future research is needed to systematically shed light on these factors, particularly when delivered in the community, in order to improve and maximise the benefits of social skills groups for a range of individuals on the spectrum. As a priority, researchers should also investigate the degree to which new skills generalise into the “real world”, and are extended, adapted and maintained in new situations.

3 Recommendation development

The Living Guideline Group was tasked with considering the systematically updated evidence on social skills groups reported above in terms of its implications for the ASD Guideline [1]. Specifically, the LGG considered whether the new evidence required revisions of existing recommendations as well as the development of any new recommendations. Both text of recommendations and their graded “strength of evidence” (see **Appendix 1, Table A1.2**) were revised/developed and considered at an all day face-to-face meeting. The LGG’s decisions for recommendation development and grading are presented below. Revised or new recommendations are accompanied by a brief rationale which highlights any particular issues that the LGG took into account while formulating the recommendations.

Preamble

In considering the evidence identified and synthesised in the systematic review update, the Living Guideline Group acknowledge the challenges and limitations of evaluating behavioural interventions. In particular, the LGG are aware that lack of consistency within and between outcomes measured across different studies does not necessary reflect a flawed evidence base. Rather, the variability reflects the complex, dynamic and multidimensional nature of social competence indicators and the richness of the research data (see **Section 2.6**, under “Variability of results” for further discussion of these issues).

Revision of existing recommendations

One recommendation in the ASD Guideline [1] was considered for revision by the Living Guideline Group.

- **Original Recommendation 4.2.1:** “The development of social skills and community support groups for young people and adults should be undertaken to minimise and avoid problems.” (**Grade C**)
- **Unchanged**

Rationale: This Recommendation was unchanged. It is broader in scope than the current research which focused on structured and facilitated social skills groups.

New recommendation

A new recommendation was developed by the LGG (see **Table 3.1**).

- **New Recommendation 4.2.1a:** “Facilitated and structured social skills groups should be considered for high functioning children and young people with ASD”. (Grade B).

Additional text: Social skills groups as referred to in this new Recommendation are defined as interventions which provide structured sessions in social skills training in small groups of people of a similar age group and with similar social problems. A session typically includes teaching a specific skill, demonstration of the skill through role playing, practice of the skill, and individualised feedback. Groups meet on a regular basis, typically for 1-2 hours, for several weeks, facilitated by at least one trained instructor/therapist. Parents are typically provided training in concurrent sessions.

Whilst it is acknowledged that the term “high functioning” is not universally favoured in the context of autism, in this Recommendation the term “high functioning” is used to refer to people with higher cognitive functioning either as established by intelligence tests (generally indicated by full IQ scores of 70 or above), or through the diagnosis of “high-functioning autism” or Asperger syndrome (under DSM-IV criteria).

There is insufficient evidence relating to people who are not verbally fluent, people who use augmentative and alternative communication (AAC), people with intellectual disability, adults, and pre-schoolers.

It is not currently possible to offer clear conclusions about the necessary content, approach and intensity of social skills groups. Further research is needed, particularly considering social skills interventions conducted and evaluated in naturalistic settings, and into the generalisation and maintenance of acquired social competence and other salient outcomes.

There is insufficient evidence relating to the relative benefits of social skills groups versus other group-based interventions, or individual social skills interventions.

Rationale: Whilst there is overall evidence of benefit from participating in social skills groups, the Recommendation is graded B to reflect the variability in programme content, approach and intensity, large range of outcome variables and measures, and uncertainty about the applicability and generalisability to the New Zealand context.

Participants in the included studies were required to demonstrate a level of cognitive and/or verbal ability above a stated threshold as determined from standardised IQ tests (generally indicated by full IQ scores of 70 or above).

Two new good practice points were developed by the LGG (see **Table 3.2**).

- **New Good Practice Point 4.2.1b:** “Social skills groups approaches in New Zealand need to be responsive to the cultural and linguistic diversity of the group participants.” ✓

Rationale: Further research is needed to explore the applicability and effectiveness of social skills groups in the New Zealand context.

- **New Good Practice Point 4.2.1c:** “Decisions about participating in social skills groups should be guided by whether a person with ASD values it, and whether they are expected to benefit from it.” ✓

Rationale: Interest in participating in a social skills group was a common inclusion criteria for participants in the research trials appraised. Other factors such as the ability to attend, and the ability to participate in the group, may also be considered by the person with ASD, their family, and the programme instructors, in deciding whether a social skills group is suitable.

It is noted that the new Recommendation (4.2.1a) and Good Practice Points (4.2.1b and 4.2.1c) could equally sit in Chapter 3, Section 3.2 (pg 106) of the NZ ASD Guideline, after Recommendation 3.2.2.4.

Table 3.1: New recommendations relevant to social skills groups in children and young people with ASD.

Reference	New recommendations	Grade
4.2.1a	Facilitated and structured social skills groups should be considered for high functioning children and young people with ASD.	B

Table 3.2: New good practice points relevant to social skills groups in children and young people with ASD.

Reference	New Good Practice Points	Grade
4.2.1b	Social skills groups approaches in New Zealand need to be responsive to the cultural and linguistic diversity of the group participants.	✓
4.2.1c	Decisions about participating in social skills groups should be guided by whether a person with ASD values it, and whether they are expected to benefit from it.	✓

Appendix 1: Methods

A1.1 Contributors

Living Guideline Group members

Matt Frost (Chair)²

Senior Advisor, Office for Disability Issues, Ministry of Social Development

Matt Eggleston (Deputy Chair, Acting Chair)³

Child and Adolescent Psychiatrist, Clinical Head, Child and Family Specialty Service, Canterbury DHB

Tanya Breen

Consultant Clinical Psychologist, Specialist in Autism Spectrum Disorder and Disability Issues, Clinical Consultant for Altogether Autism, Hamilton

Jill Bevan-Brown

Director, Inclusive Education Research Centre, Institute of Education, Massey University

Sally Clendon⁴

Senior Lecturer, Speech Language Therapy Programme, School of Education at Albany, College of Education, Massey University

Elizabeth Doell⁵

Senior Lecturer, Speech Language Therapy Programme, School of Education at Albany, College of Education, Massey University

Debbie Fewtrell

General Practitioner (special interest in autism spectrum disorder), Kerikeri

Andrew Marshall

Developmental Paediatrician, Child Development Team at Puketiro Centre, Porirua and Clinical Leader, Child Health, Wellington Hospital

Ex-officio LGG members

Natasha Gartner

ASD Project Manager, Family and Community, Disability Support Services, National Services Purchasing, National Health Board, Ministry of Health

² Whilst the Chair of the LGG, Mr Frost was only able to attend part of the LGG meeting considering the evidence on social skills groups held in November 2014 and so did not chair the meeting.

³ Dr Eggleston chaired the LGG face-to-face meeting as Acting Chair.

⁴ Dr Clendon was unable to attend the LGG's meeting. Dr Clendon participated fully as member of the LGG for topic prioritisation processes.

⁵ As Dr Clendon was unable to attend the LGG's meeting, her colleague Dr Doell was co-opted to participate in her place for this update topic.

Anna Kelly (until July 2014)⁶

Team Leader – Research, Autism Spectrum Disorder (ASD) National Team, Special Education, Ministry of Education

INSIGHT Research

Marita Broadstock

Director, INSIGHT Research Ltd, Living Guideline Group Project Manager and lead researcher

Declarations of competing interest

None

Acknowledgements

INSIGHT Research thanks the Ministry of Education Library staff for their assistance in retrieval of articles pertinent to this review.

A1.2 Review scope

The current review updates evidence for the ASD Guideline [1] on group-based social skills training for children and young people aged 6-21 years with autism spectrum disorder.

The original searching for the ASD Guideline [1] was performed in July 2004. For the original Guideline, papers published before the search dates and in some cases after the completion of searching were suggested by members of all workstreams and incorporated in the text and evidence tables, where appropriate.

In the current update, the search was limited to articles published in the English language on or beyond January 1 2004. Given the overlap in search periods in 2004, and the inclusion of papers outside the date range in the original Guideline, papers identified in the current search strategy which were already appraised in the original ASD Guideline [1] were excluded.

Publications were considered where evaluating the effectiveness of group-based social skills training for children and young people (aged 6-21 years) with Autism Spectrum Disorder. Eligible studies were randomised controlled trials evaluating group-based interventions among children and young people with ASD where social skills are the focus of the training and of the outcomes measured. Comparators were wait list control groups, or those receiving usual care and/or no treatment.

⁶ This position is currently being filled. Warwick Philips, Manager Professional Practice, Ministry of Education, is acting as Ministry sponsor in the interim.

A1.3 Research question

The Living Guideline Group identified social skills groups in children and young people with ASD as a priority topic to update. The lead researcher prepared the research questions in the PECO format (which identifies the Patient, Exposure, Comparison, and Outcomes of interest) to ensure effective and focused searches and reviews could be undertaken. The research question was:

- RESEARCH QUESTION: How effective are social skills groups for assisting children and young people with ASD?

A1.4 Search strategy

Search strategies were limited to publications from January 1 2004 onwards. Database searches were conducted over May-June 2014.

The INSIGHT Research lead researcher set the inclusion and exclusion criteria for the review in consultation with the Ministry of Health. Systematic database searching was designed and conducted by the INSIGHT lead researcher. Full search strategies are available upon request.

Search databases

Bibliographic, health technology assessment and guideline databases were included in the search strategy, listed below.

- Medline
- Cinahl
- Embase
- Social Sciences Citation Index (SSCI)
- PsycInfo
- Cochrane Database of Systematic Reviews (CDSR)
- Central Register of Controlled Trials (CRCT)
- Database of Abstracts of Reviews of Effects (DARE)
- Health Technology Assessment Database (HTA Database)

Search terms were adapted for different databases. The following search for EMBASE includes a search string relating to social skills groups adapted from that used by the Cochrane review of Reichnow et al [26].

- child development disorders, pervasive/ or asperger syndrome/ or autistic disorder/ OR (pdd* or asd*).tw. OR autis*.tw. OR asperger*.tw. OR kanner*.tw.

AND

- social behavior/ OR interpersonal communication/ OR social competence/ OR social interaction/ OR (interpersonal adj3 (behav\$ or communication\$ or competenc\$ or relation\$ or skill\$)).tw. OR (social adj3 (behav\$ or communication\$ or competenc\$ or relation\$ or skill\$)).tw.

AND

- behavior therapy/ OR social adaptation/ OR (educat\$ or train\$ or program\$ or therap\$ or intervention\$).tw.

AND

- English, human only, 2004-current

Cross-checking of references from retrieved studies was conducted to identify additional references.

A1.5 Appraisal of studies

For this review, a single researcher performed study selection, critical appraisal and synthesis. The following steps were followed in appraising the evidence.

Assigning a level of evidence

Following the completion of searches, retrieved studies meeting the selection criteria were assigned a “level of evidence”. The level of evidence indicates how well the study eliminates bias based on its design. INSIGHT Research uses a published evidence hierarchy, designed by the National Health and Medical Research Council of Australia (NHMRC) [28]. These describe research designs which are broadly associated with particular methodological strengths and limitations so as to rank them in terms of quality, from I (systematic reviews of level II studies) to IV (case series).

For intervention studies where an intervention can be allocated experimentally, randomised controlled trials (level II studies) are considered the most robust way of determining a true association.

Table A1.1: NHMRC levels of evidence [28]

Level	Intervention
I	A systematic review of level II studies
II	A randomised controlled trial
III-1	A pseudo-randomised controlled trial (ie, alternate allocation or some other method)
III-2	A comparative study with concurrent controls: Non-randomised, experimental trial Cohort study Case-control study Interrupted time series with a control group
III-3	A comparative study without concurrent controls: Historical control study Two or more single arm study Interrupted time series without a parallel control group
IV	Case series with either post-test or pre-test/post-test outcomes

In the hierarchy of evidence employed (see **Table A1.1**), systematic reviews which included level II studies are ranked as level I evidence whereas systematic reviews of lower order evidence rank at the same level as that order of evidence.

Appraising the quality of included studies

Completing evidence tables

Evidence tables (**Appendix 3**) were completed for each appraised study. Evidence tables present the key characteristics of each of the appraised studies including sample characteristics, methodology, results, the level of evidence, and the summary codes of study quality.

Appraisal of primary and secondary studies

Studies were appraised using adapted versions of the GATE (Graphic Appraisal Tool for Epidemiology) Frame tools (designed by the University of Auckland's School of Population Health) appropriate to study design (systematic reviews, and randomised controlled trials). The adapted GATE has been validated by the New Zealand Guidelines Group (NZGG).

In brief, the GATE checklists are comprised of slightly different criteria depending on the study design but all broadly address each part of the PECO framework. The case is slightly different for systematic reviews and meta-analyses where additional criteria are included to assess the appropriateness of combining and analysing multiple studies. In general however, the checklists help the researcher to assess study quality in three main areas:

- study validity (steps made to minimise bias)
- study results (size of effect and precision)
- study relevance (applicability and generalisability).

For each checklist item, the reviewer codes whether the criterion for quality has been met (+), is unmet (x) or, where there is not enough information to make a judgement, is unknown (?). Reviewers then assign the same quality codes to each of three summary sections which assess the accuracy, relevance and applicability of the findings. Here, the reviewer indicates whether the study has any major flaws that could affect the validity of the findings and whether the study is relevant to clinical practice. The three summary sections include:

1. internal validity – potential sources of bias
2. precision of results
3. applicability of results/external validity – relevance to key questions and clinical practice.

Finally, reviewers assign an overall assessment of quality for the study as a whole based on a consideration of all checklist criteria; codes used are:

+ good

x not ok, poor

? unclear

Codes for each of the three summary domains, and an overall study quality code are presented in the bottom row of the evidence tables for each study (**Appendix 3**).

A1.6 Preparing recommendations

Developing recommendations

A one-day face-to-face meeting was held on 11 November 2014 where the Living Guideline Group considered the findings of the current systematic review and developed new recommendations or revised those of the original ASD Guideline [1]. Using their collective professional judgement and experience, the LGG discussed the body of evidence with respect to the research questions and the applicability of the evidence within New Zealand.

Developing recommendations involves consideration of the whole evidence base for the research question. The quality and consistency of the evidence and the clinical implications of the evidence within a New Zealand context is weighed up by all the LGG members. The recommendations were agreed by consensus during the meeting.

Table A1.2: Guide to grading recommendations [1]

Recommendations	Grade
The recommendation is supported by good evidence (based on a number of studies that are valid, consistent, applicable and clinically relevant)	A
The recommendation is supported by fair evidence (based on studies that are valid, but there are some concerns about the volume, consistency, applicability and clinical relevance of the evidence that may cause some uncertainty but are not likely to be overturned by other evidence)	B
The recommendation is supported by international expert opinion	C
The evidence is insufficient, evidence is lacking, of poor quality or opinions conflicting, the balance of benefits and harms cannot be determined	I

Note: Grades indicate the strength of the supporting evidence rather than the importance of the evidence.

Good practice point	Grade
Where no evidence is available, best practice recommendations are made based on the experience of the Living Guideline Group or feedback from consultation within New Zealand.	✓

Note: Good practice points are the opinion of the Living Guideline Group, or developed from feedback from consultation within New Zealand where no evidence is available

Each recommendation is assigned a grade to indicate the overall “strength of the evidence” upon which it is based. Strength of the body of evidence is determined by three domains [28]:

- quality (the extent to which bias was minimised as determined by study design and the conduct of the study)
- quantity (magnitude of effect, numbers of studies, sample size or power)
- consistency (the extent to which similar findings are reported).

It should be noted that systematic reviews and meta analyses (secondary studies) considered drawing on publications over an overlapping timeframe could report on (some of) the same studies. For this reason it is important to be aware that the results from secondary studies should not be summated as independent sources of evidence as this would misrepresent the quantity of studies and give shared primary studies undue weight.

The grades of recommendations used by the Living Guideline Group, and also used in the original ASD Guideline [1], are presented in **Table A1.2**.

A1.7 Consultation

Seeking comments from stakeholders is vital for peer-review and quality assurance processes in developing the report. In a focused consultation 9 key stakeholder organisations/individuals were approached for feedback on a late draft of the report. Particular attention was sought regarding the relevance of the report to New Zealand

services and needs, clarity and ease of use of the report, and implementability of the revised or new recommendations.

Responses were received from 8 organisations/individuals, including: Altogether Autism, Autism New Zealand, IDEA, Ministry of Education, Ministry of Health (two respondents), New Zealand Psychological Society, and the Paediatric Society of New Zealand.

The lead researcher (INSIGHT Research) collated feedback and drafted revisions for the LGG to consider. Amendments were finalised by group consensus. Suggestions identified in the consultation led to several improvements to the final report. INSIGHT Research and the LGG are grateful to those individuals and organisations who participated in the consultation process.

Appendix 2: Abbreviations and glossary

A2.1 Abbreviations and acronyms

Miscellaneous Terms

ADHD	attention-deficit/hyperactivity disorder
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AS	Asperger syndrome
ASD	Autism Spectrum Disorder
CG	control group
CI	confidence interval
ES	effect size
HFA	high functioning autism
HLM	hierarchical linear modeling
HTA	Health Technology Assessment
ID	intellectual disability
IQ	intelligence quotient
INSIGHT Research	Independent Network of Specialists in Guidelines & Health Technology Research
LGG	Living Guideline Group
M	mean
MANOVA	multivariate analysis of variance
N (or n)	number (usually, sample size)
NHMRC	National Health and Medical Research Council (Australia)
NICHD	National Institute of Child Health and Human Development
NIH	National Institute of Health (US)
NIMH	National Institute of Mental Health (US)
NINDS	National Institute of Neurological Disorders and Stroke
NZ	New Zealand
NZGG	New Zealand Guidelines Group
PDD	Pervasive Developmental Disorder
PDD-NOS	Pervasive Developmental Disorder – Not Otherwise Specified
PECO	Patient, Exposure, Comparison, Outcome
RCT	Randomised controlled trial
SD	Standard deviation
SES	socio-economic status
SR	Systematic review
SST	social skills training
TG	treatment group
UCLA	University of California Los Angeles
UK	United Kingdom
US	United States of America
vs	versus

Tests, scales and measures

ABC	University of California Los Angeles
ADI-R	Aberrant Behavior Checklist
ADOS	Autism Diagnostic Observation Schedule
AOQ	Affection for Others Questionnaire
AQ	Autism Spectrum Quotient
ASC	Adaptive Skillstreaming Checklist
ASDI	Asperger Syndrome Diagnostic Interview
ASDS	Asperger Syndrome Diagnostic Scale
ASSQ	High Functioning Autism Spectrum Screening Questionnaire
BASC-2	Behavior Assessment System for Children – 2nd edition
BDI	Beck Depression Inventory
CASL	Comprehensive Assessment of Spoken Language
CAST	Childhood Asperger Syndrome Test
CBCL	Child Behavior Checklist
CDI	Child Depression Inventory
CFT	Children's Friendship Training
CGI-I	Clinical Global Impressions Scale - Improvement
CSI	Children's Symptom Inventory;
DANVA2	Diagnostic Analysis of Nonverbal Accuracy – 2nd edition
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders - IV (text revision)
DSM5	Diagnostic and Statistical Manual of Mental Disorders – 5 th edition
EQ	empathy quotient
ERSSQ	Emotion Regulation and Social Skills Questionnaire
FQS	Friendship Qualities Scale
GAQ	General Affection Questionnaire
GATE	Graphic Appraisal Tool for Epidemiology
JDTP	Junior Detective Training Program
PDDBI	Pervasive Developmental Disorder Behavior Inventory
PEERS	Program for the Education and Enrichment of Relational Skills
PEERS	Program for the Education and Enrichment of Relational Skills
PEI	Pupil Evaluation Inventory
PHS	Piers-Harris Self-Concept Scale
PSI	Pro-Social Index
PSS	Parent Satisfaction Survey
QPQ	Quality of Socialization Questionnaire
SCAS	Spence Child Anxiety Scale
SCI	Social Competence Inventory
SCPQ-P	Social Competence with Peers Questionnaire
SCQ	Social Communication Questionnaire
SELSA	Social and Emotional Loneliness Scale for Adults
SIAS	Social Interaction Anxiety Scale
SKA	Skillstreaming Knowledge Assessment
SRS	Social Responsiveness Scale
SSI	Social Skills Inventory
SSIS	Social Skills Improvement System-Rating Scale
SSQ	Social Skills Questionnaire
SSRS	Social Skills Rating Scale
STAI-T/STAI-S	State and Trait Anxiety Inventory, trait and state versions

STAIC-T/STAIC-S	State and Trait Anxiety Inventory for Children, trait and state versions
TASSK	Test of Adolescent Social Skills Knowledge
TASSK-R	Test of Adolescent Social Skills Knowledge-Revised
TYASK	Test of Young Adult Social Skills Knowledge
VABS	Vineland Adaptive Behavior Scale
WASI	Weschler Abbreviated Scale of Intelligence
WFT	Walk in the Forest Test
WISC-III/IV	Weschler Intelligence Scale for Children-3rd edition and 4th editions

Databases

CENTRAL	Cochrane Central Register of Controlled Trials
CINAHL	Cumulative Index to Nursing and Allied Health Literature
DARE	Database of Abstracts of Reviews of Effects
Embase	Excerpta Medica Database
ERIC	Education Resources Information Centre
HTA database	Health Technology Assessment Database
Medline	Medical Literature Analysis and Retrieval System Online
OCLC WorldCat	Online Computer Library Center World Catalogue
PsycINFO	Psychology Information Database
SSCI	Social Sciences Citation Index

A2.2 Glossary

Epidemiological and statistical terms

Bias

Bias is a systematic deviation of a measurement from the “true” value leading to either an over- or under-estimation of the treatment effect. Bias can originate from many different sources, such as allocation of patients, measurement, interpretation, publication and review of data

Bonferroni’s correction

In statistics, the Bonferroni correction is a method used to counteract the problem of multiple comparisons increasing the likelihood of chance effects being interpreted as significant. The correction increases the p value accepted as denoting a statistically significant difference or effect.

Case series

Case series are collections of individual case reports, which may occur within a fairly short period of time. Cases consist of either only the exposed people with the outcomes, or people with the outcome regardless of the exposure. In neither of these examples can the risk for the outcome be determined

Case-control study

Patients with a certain outcome or disease and an appropriate group of controls without the outcome or disease are selected (usually with careful consideration of appropriate choice of controls, matching, etc.) and then information is obtained on whether the subjects have been exposed to the factor under investigation.

Cohort study

Subsets of a defined population can be identified who are, have been, or in the future may be exposed or not exposed in different degrees, to a risk factor or factors hypothesised to influence the probability of occurrence of a given disease or other outcome. Subjects are followed from a well-described starting point to determine whether the outcome/disease occurs (either retrospectively, or prospectively). The control group of people not exposed to the risk factor can be identified within the population-based cohort, and be matched by confounders known to be associated with the outcome (e.g., age, sex), or can be obtained from an historical cohort. Studies usually involve the observation of a large population, for a prolonged period (years).

Comorbid condition

One that exists at the same time as another condition in the same individual. The two conditions are usually independent of each other. For example a child who has autism might also develop leukaemia. That the child has autism complicates treating the leukaemia, but the two conditions are independent of each other.

Co-morbidities

Conditions which occur in association with another condition (e.g., ASD) more commonly than in the general population

Cross-sectional study

A study that examines the relationship between exposures (e.g., risk factor) and outcomes (e.g., disease) as they exist in a defined population, at a particular time.

Detection bias

Detection bias refers to systematic differences between groups in how outcomes are determined. Awareness by outcome assessors/respondents of whether an intervention was received or not (i.e., they are not blind to allocated condition) may increase the risk of their measurements/ratings/reports being affected by detection bias.

Effect size

A quantitative measure of the strength of a phenomenon, a standardised measure of the size of the difference between two groups.

Effectiveness

A measure of the extent to which a specific intervention, procedure, regimen, or service, when deployed in the field in routine circumstances, does what it is intended to do for a specified population.

Generalisability

Applicability of the results to other populations.

High functioning

Whilst it is acknowledged that the term “high functioning” is not universally favoured, in this report, the term “high functioning” is used to refer to people with higher cognitive functioning either as established by intelligence tests (generally indicated by full IQ scores of 70 or above), or through the diagnosis of “high-functioning autism” or Asperger syndrome (under DSM-IV criteria). In light of the removal of Asperger syndrome as a separate diagnostic classification in DSM-5, these distinctions may no longer be used clinically.

Idiomatic language

Idiomatic language is the use of expressions such as metaphors and idioms that don't mean the same as their literal meaning (e.g., its raining cats and dogs).

Level of evidence

A hierarchy of study evidence that indicates the degree to which bias has been eliminated in the study design.

Matched controls

Matching is a method used to ensure that two study groups are similar with regards to "nuisance" factors that might distort or confound a relationship that is being studied (e.g., age, sex)

Mean

Calculated by adding all the individual values in the group and dividing by the number of values in the group.

Observational studies

Also known as epidemiological studies. These are usually undertaken by investigators who are not involved in the clinical care of the patients being studied, and who are not using the technology under investigation.

Performance bias

Performance bias refers to systematic differences between groups in the care that is provided, or in exposure to factors other than the interventions of interest. After enrolment into the study, blinding (or masking) of study participants and personnel may reduce the risk that knowledge of which intervention was received, rather than the intervention itself, affects outcomes. Effective blinding can also ensure that the compared groups receive a similar amount of attention, ancillary treatment and diagnostic investigations. Blinding is not always possible, however.

Power

The probability that a statistical test or study will detect a defined pattern in data and declare the extent of the pattern as showing statistical significance.

Quality of evidence

Degree to which bias has been prevented through the design and conduct of research from which evidence is derived.

Randomised controlled trial (RCT)

An epidemiological experiment in which subjects in a population are randomly allocated into groups to receive or not receive an experimental preventive or therapeutic procedure, manoeuvre, or intervention. The groups are compared prospectively.

Secondary study

An analysis or synthesis of research data reported elsewhere, including systematic reviews, meta analyses and guidelines.

Selection bias

Error due to systematic differences in characteristics between those who are selected for inclusion in a study and those who are not (or between those compared within a study and those who are not).

Strength of evidence

The strength of evidence for an intervention effect includes the level (type of studies), quality (how well the studies were designed and performed to eliminate bias) and statistical precision (P-value and confidence interval).

Systematic review (SR)

A literature review reporting a systematic method to search for, identify and appraise a number of independent studies.

Topic Specific Terms

Neurotypical

An abbreviation of neurologically typical, a term coined in the autistic community as a label for people who are not on the autism spectrum.

Prosody

In linguistics, prosody is the rhythm, stress, and intonation of speech.

Social competence

The ability to achieve personal goals in social interaction while simultaneously maintaining positive relationships with others over time and across situations [21]

Social skills groups

Social skills groups provide structured sessions in social skills training in small groups of people of a similar age group with similar social problems. A session typically includes teaching a specific skill, demonstration of the skill through role playing, practice of the skill, and individualised feedback. Groups meet on a regular basis, typically for 1-2 hours, for several weeks, facilitated by at least one trained instructor/therapist. Parents are typically provided training in concurrent sessions to encourage their children to practice newly learned skills at home.

Appendix 3: Evidence Tables of included studies

Studies are ordered using the following hierarchy: study type (systematic reviews then primary studies), level of evidence (highest first), year of publication (most recently published last), first author's surname (alphabetical order).

Systematic reviews

Kaat & Lecavalier (2014) [24]					
Country, study type, aims	Review scope	Participants and search method	Inclusion and exclusion criteria	Results	Conclusions
<p>Country: US</p> <p>Study type: systematic review of group-based social skills training (SST)</p> <p>Evidence level: I (systematic review includes level II primary studies)</p>	<p>Review scope: group based SST for children with ASD, using a range of study designs, published between 2000 and 2012. Focus is on review of methodology rather than efficacy <i>per se</i>.</p>	<p>Participants: children and adolescents aged 18 years or younger with ASD</p> <p>Search method: PsycINFO and PubMed databases searched using narrow scope of search terms in English, peer reviewed journals, as well as reference checking.</p> <p>Appraisal: Studies were coded according to quality criteria including: study design criteria, sample characteristics, manualisation of intervention, programme fidelity, treatment intensity, outcome variables, rigour of statistical analyses, generalisation or maintenance effects, funding source.</p>	<p>Inclusion: children aged 18 years or younger with ASD; group-based (n=3 or more, at least 2 with ASD) SST intervention; original data on treatment outcomes.</p> <p>Exclusion: music and art therapies unless social skills primarily targeted and social gains outcomes measured.</p>	<p>48 eligible studies identified, 7 were single case experimental design studies, 20 were uncontrolled trials, and 21 were controlled studies, 13 of which (27% of whole sample) were randomised controlled trials (RCTs).</p> <p>Key findings:</p> <p>Considering all 48 studies:</p> <ul style="list-style-type: none"> - very few studies independently confirmed IQ or ASD diagnoses - 4 (8%) did not exclude people with intellectual disability (ID) - 18 (38%) had sample of 10 or fewer in treatment group - most (71%) used or adapted manualised treatments - most (91%) had validated outcome measures. <p>Limitations of evidence base</p> <p>Methodological limitations were discussed at length in comparison to quality criteria.</p> <p>Recommendations provided for future trials with respect to study design, expanding participant characterisation (e.g., people with ID), increasing sample sizes, statistical methods and analyses, and outcome measures.</p>	<p>Author conclusions:</p> <p><i>“Methodological rigor has been improving as the field grows.”</i></p> <p>Argues it is <i>“presumptive to conduct an efficacy review across all SST interventions given the differences in program design, targeted skills, instructional methods, and the various outcomes reported.”</i></p> <p>Reviewer's comments: Fairly restricted search strategy, useful qualitative appraisal of study quality, useful tables of study characteristics organised around study designs. Two of the 13 studies identified as RCT's (Lopata 2008; Owens 2008) were not fully randomised in group allocation.</p> <p>Source of funding: not stated, authors are academic university-based researchers</p>
Study quality: Internal validity: ? Precision: ? Applicability: ?					Overall Score: ?

Key: ASD=autism spectrum disorder; ID=intellectual disability; IQ=Intelligence Quotient; RCT=randomised controlled trial; SST=social skills training; US=United States of America

Reichow et al (2012) [26]					
Country, study type, aims	Review scope	Participants and search method	Inclusion and exclusion criteria	Results	Conclusions
<p>Country: US</p> <p>Study type: systematic review of social skills groups for young people with ASD</p> <p>Evidence level: I (systematic review includes level II primary studies)</p>	<p>Review scope: effectiveness of social skills groups for improving social competence, social communication and quality of life in people aged 6-21 years with ASD.</p>	<p>Participants: children and young people aged 6-21 years with ASD</p> <p>Search method: Databases: CENTRAL, MEDLINE, EMBASE, PsycINFO, CINAHL, ERIC, Sociological Abstracts, OCLC WorldCat, Social Citation Index, Controlled Trials, searched from inception to 2011 using broad search strategies, and reference checking of published papers.</p> <p>Appraisal: Two reviewers independently selected studies and assessed risk of bias using validated checklists. Effect sizes and random effects meta analyses conducted.</p>	<p>Inclusion: young people aged 6-21 years with ASD; group-based (n=2 or more participants with ASD per group) social skills intervention; RCT design; compared with no intervention, wait list or usual treatment; evaluated social competence as primary outcome.</p> <p>Exclusion: studies evaluating support groups and psychodynamic group therapies.</p>	<p>5 RCTs with 196 participants were included.</p> <p>Key findings: Social skills group vs control group had significantly increased:</p> <ul style="list-style-type: none"> - social competence (ES=0.47, 95% CI=0.16 to 0.78, p=0.003), and - friendship quality (ES=0.41, 95% CI=0.02 to 0.81, p=0.04). <p>No treatment effects found for emotional recognition (2 studies) or social communication as related to understanding of idioms (1 study).</p> <p>Single studies found treatment effect of decreased loneliness (ES=-0.66, 95% CI=-1.15 to -0.17, p=0.008), but no effect on child or parental depression. No adverse effects reported.</p> <p>Limitations of evidence base</p> <p>A small number of studies included. A high risk for bias due to parents knowing that their child was in the intervention group or not. Limited generalisability. Studies were all conducted in the US, mainly focused on children aged 7-12 years, and participants were of average or above average intelligence.</p>	<p>Author conclusions: There is emerging evidence of effectiveness of social skills group interventions for young people with ASD. <i>“Participants in social groups may make modest gains in social competence, have better friendships, and experience less loneliness... More research is needed to draw more robust conclusions, especially with respect to improvements in quality of life.”</i></p> <p>Reviewer’s comments: Comprehensive search strategy, robust appraisal and analysis of study quality, detailed tables of study characteristics and results.</p> <p>Source of funding: internal funding from Yale Child Center and Yale University School of Medicine. No external funding.</p>
Study quality: Internal validity: + Precision: + Applicability: +					Overall Score: +

Key: ASD=autism spectrum disorder; CENTRAL=Cochrane Central Register of Controlled Trials; CI=confidence interval; CINAHL=Cumulative Index to Nursing and Allied Health Literature; Embase=Excerpta Medica Database; ERIC=Education Resources Information Centre; ES=effect size; Medline=Medical Literature Analysis and Retrieval System Online; OCLC Worldcat=Online Computer Library Center World Catalogue; PsycINFO= Psychology Information Database; RCT=randomised controlled trial; US=United States of America, vs=versus

Primary studies

Beaumont & Sofronoff, 2008 [29]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: Australia Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Intervention delivered in Queensland, Australia. Participants recruited from newspaper notices, Asperger syndrome (AS) support group newsletter, and practitioner referrals.</p> <p>Participants: 49 children with AS aged 7-11 years (M age=9.7 years); 44 male (90%), 5 female. No ethnicity information provided.</p> <p>Treatment group (TG): N=26 Control group (CG); N=23</p>	<p>Inclusion: aged 7 1/2 -11 years; reported AS diagnosis confirmed by a pediatrician (with parent-completed CAST at recruitment); WISC-III prorated IQ of ≥ 85 or above.</p> <p>Exclusion: none reported</p> <p>Follow-up: pre-test, post-test 7 weeks post baseline (immediately after treatment for TG) to determine treatment effects, then 6 weeks, and 5 months, after post-test to assess maintenance.</p> <p>Fidelity: fidelity check sheets completed during sessions and verified by raters from sample of video-taped sessions.</p> <p>Make-up sessions scheduled for children who missed sessions (not clear if these only included one child). Completion rate appears to be 100% but not explicitly stated.</p>	<p>Treatment (TG): Junior Detective Training Program (JDTP): manualised multicomponent social skills intervention (with computer game time for part of first 4 sessions). Each small group (3 children) received eight, 2-hourly, weekly sessions. Parents attended separate concurrent sessions. A token economy was used to reward children for task and homework completion. Teacher handouts were provided.</p> <p>Control (CG): wait list control</p> <p>Outcomes (completed by):</p> <ul style="list-style-type: none"> - SSQ: social competence (parent, teacher) - ERSSQ; social skills measure (parent, teacher) - Emotion recognition: facial expression, and posture cues expression (assessor) - Emotion management knowledge: "Dylan is being teased" scenario about coping with anxiety; and "James and the maths test" scenario about coping with bullying (assessor) 	<p>There were no significant differences between groups at baseline in age, IQ, ASD symptoms (CAST), and outcome variables.</p> <p>Key findings:</p> <p><u>Parent reported social skills:</u> repeated measures mixed-model Group (Treatment vs Control) X Time (pre vs post) MANOVA found significant Group X Time interaction: $F(2,45)=13.14, p<0.001, \eta^2=.37$. Follow-up analyses confirmed TG improved in SSQ (parent) ($p<0.001, \eta^2=.54$), and ERSSQ (parent) ($p<0.001, \eta^2=.57$), but not the control group. These differences brought post-test results into the normal range suggesting clinical significance.</p> <p><u>Teacher reported social skills:</u> insufficient response rate (only returned for 19 of 49 students) to permit investigation of group effects. Whole group data not reported here.</p> <p><u>Emotion recognition</u></p> <p>Mixed model MANOVA found no significant Group effect or Group X Time interaction (both TG and CG made significant improvements over time).</p> <p style="text-align: right;">...continued overpage</p>	<p>Author conclusions: The Junior Detective Training Program appeared effective in enhancing the social skills and emotional understanding of children with AS.</p> <p>Reviewer's comments: ASD diagnoses not independently verified by researchers. Programme fidelity recorded but not reported. No direct observational data collected. Parents and teachers not blind to allocation and open to bias in assessments. Teacher-reported social skills not analysed due to poor response rate. Lack of significant findings for emotion recognition may be due to ceiling effects as task was simple, and practice effects as task repeated from baseline (improvement evident in both TG and CG). As multimodal intervention, not clear what components of JDTP were effective/necessary.</p> <p>Source of funding: Not reported; university based academics.</p>

Key: AS=Asperger Syndrome; ASD=autism spectrum disorder; CAST=Childhood Asperger Syndrome Test; CG=control group; ERSSQ=Emotion Regulation and Social Skills Questionnaire; IQ=intelligence quotient; JDTP=Junior Detective Training Program; M=mean; MANOVA=multivariate analysis of variance; RCT=randomised controlled trial; SSQ:Social Skills Questionnaire; TG=treatment group; WISC-III=Weschler Intelligence Scale for Children-3rd edition.

Beaumont & Sofronoff, 2008 [29] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
				<p>...continued</p> <p><u>Emotion management knowledge:</u> Mixed-model MANOVA found significant Group X Time interaction: $F(2,45)=9.61, p<0.001, \eta^2=.30$. Follow-up analyses confirmed TG improved ($p<0.001, \eta^2>.35$), but not the control group.</p> <p><u>Maintenance</u> Improvements for TG for SSQ and ERSSQ maintained at 6-week and 5-month follow-up after post-test ($p<0.001$). There were trends for treatment gains for Dylan and James scenario measures ($p<0.02$) but these didn't reach significance using Bonferroni-adjusted alpha of 0.008.</p>	
Study quality: Internal validity: ? Precision: + Applicability: +					Overall Score: +

Key: AS=Asperger Syndrome; ASD=autism spectrum disorder; CAST=Childhood Asperger Syndrome Test; CG=control group; ERSSQ=Emotion Regulation and Social Skills Questionnaire; IQ=intelligence quotient; JDTP=Junior Detective Training Program; M=mean; MANOVA=multivariate analysis of variance; RCT=randomised controlled trial; SSQ:Social Skills Questionnaire; TG=treatment group; WISC-III=Weschler Intelligence Scale for Children-3rd edition.

Laugeson et al, 2009 [32]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Intervention delivered in the community in Southern California. Recruited from regional centres and school, and UCLA outpatient clinics. Participants: 33 adolescents (of 36 randomised) aged 13-17 years (M age=14.6 years); 28 male (85%), 5 female; 23 with high functioning autism (HFA), 9 with Asperger Syndrome (AS), 1 with PDD-NOS; 14 identified as Caucasian, 6 as Hispanic/Latino, 3 as African American, 4 as Asian, 3 as Middle Eastern, 2 as mixed ethnicities; 7 were prescribed psychoactive medications; 17 were in mainstream classes, 8 were in special education classes, 2 received partial special education; 3 were home schooled, and 3 in other educational arrangements. Treatment group (TG): N=17 Control group (CG): N=16</p>	<p>Inclusion: aged 13-17 years; has parent-reported social problems; previous diagnosis of ASD; English fluency; parental English fluency; verbally fluent with verbal IQ of ≥ 70 or above; adolescent interest in joining group. Exclusion: history of major mental illness; or hearing, visual, or physical impairments. Follow-up: pre-test, post-test 12 weeks post baseline (immediately after treatment for TG). Maintenance not measured. 92% completion rate: 3 dropouts in TG. Fidelity: weekly fidelity check sheets from research assistants.</p>	<p>Treatment (TG): Manualised PEERS. 5 groups of approximately 7 adolescents received PEERS over 12, 90-minute, weekly sessions. Parents attended separate concurrent sessions. Control (CG): wait list control Outcomes (completed by): - TAASK: knowledge of PEERS skills (adolescent) - QPQ: number of get-togethers (i) hosted, (ii) invited to, (iii) with conflict (adolescent, parent) - SSRS: (i) social skills, (ii) problem behaviour (parent, teacher) - FQS: friendship quality (adolescent)</p>	<p>There were no significant differences between groups at baseline in socio-demographic, VABS, and outcome variables. Key findings: <u>Parent and adolescent variables:</u> repeated measures Group (Treatment vs Control) X Time (pre vs post) MANOVA found significant Group X Time interaction: Wilks Lambda=0.36; $F(10,22)=3.99$, $p<0.005$. Univariate analyses found significant Group X Time interactions for 4 outcomes: - knowledge of PEERS: TASSK (adolescent) ($p<0.0001$); - hosted get-togethers: QPQ (adolescent) ($p<0.025$); - friendship quality: FQS (adolescent) ($p<0.05$) - social skills: SSRS (parent) ($p<0.05$). Post hoc Newman Kuels tests confirmed TG improved in: adolescent-reported PEERS knowledge (TASSK) & hosted get-togethers (QPQ), & parent-reported social skills (SSRS). ...continued overpage</p>	<p>Author conclusions: In comparison with the control group, the treatment (PEERS) group significantly improved their knowledge of social skills, frequency of hosted get-togethers, and overall social skills as reported by parents. Reviewer's comments: ASD diagnoses not independently verified. No direct observational data collected. Parents not blind to allocation and open to bias. Teachers blinded to allocation but poor response rate (only returned for 13 of 33 students) and may have led analyses to be underpowered to detect differences. Maintenance not measured. Significant differences for 4 or 12 outcome measures investigated. Source of funding: Research grants from NIH and NIMH.</p>

Key: AS=Asperger syndrome; ASD=autism spectrum disorder; FQS=Friendship Qualities Scale; CG=control group; HFA=high functioning autism; IQ=intelligence quotient; M=mean; NIH=National Institute of Health; NIMH=National Mental Institute of Health; MANOVA=multivariate analysis of variance; PDD-NOS=pervasive developmental disorder not otherwise specified; PEERS=Program for the Education and Enrichment of Relational Skills; QPQ=Quality of Play Questionnaire; RCT=randomised controlled trial; SSRS=Social Skills Rating Scale; TASSK=Test of Adolescent Social Skills Knowledge; TG=treatment group; UCLA=University of California Los Angeles; US=United States of America; VABS=Vineland Adaptive Behavior Scale

Laugesan et al, 2009 [32] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
				<p>...continued</p> <p>However the effect for friendship quality (FQS) was found to be due to a decrease in friendship quality for the CG, with no significant change for the TG from baseline.</p> <p><u>Teacher variables:</u> repeated measures Group X Time MANOVA found no significant interaction for teacher-reported SSRS measures.</p> <p><u>No significant treatment effects</u> for adolescent-reported problem behaviour (SSRS), or invited or conflict at get-togethers (QPQ); parent-reported ratings of get-togethers (QPQ); or parent- or teacher-rated problem behaviour, or teacher-rated social skills (SSRS).</p>	
Study quality: Internal validity: ? Precision: + Applicability: ?					Overall Score: ?

Key: AS=Asperger syndrome; ASD=autism spectrum disorder; CG=control group; FQS=Friendship Qualities Scale; HFA=high functioning autism; IQ=intelligence quotient; M=mean; NIH=National Institute of Health; NMIH=National Mental Institute of Health; MANOVA=multivariate analysis of variance; PDD-NOS=pervasive developmental disorder not otherwise specified; PEERS=Program for the Education and Enrichment of Relational Skills; QPQ=Quality of Play Questionnaire; RCT=randomised controlled trial; SSRS=Social Skills Rating Scale; TASSK=Test of Adolescent Social Skills Knowledge; TG=treatment group; UCLA=University of California Los Angeles; US=United States of America; VABS=Vineland Adaptive Behavior Scale

Frankel et al, 2010 [25]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Intervention delivered in Southern California. Recruited from UCLA outpatient clinics referred by community practitioners.</p> <p>Participants: 68 children with high functioning ASD (of 76 randomised), M age=8.4 years; 58 male (85%), 10 female; 45 identified as Caucasian, 10 as Asian, 7 as African American, 4 Hispanic, 1 Pacific Islander and 1 Native American; 62 were completely in mainstream classes, and 6 were in special education classes some of the time.</p> <p>Other group participants not evaluated had ADHD (46%), Adjustment disorder and/or Oppositional Defiant Disorder (22%), anxiety disorder (5%), Fetal Alcohol Syndrome Disorder (1%), mood disorder (1%), learning disability (1%), or no diagnosis (25%).</p> <p>Treatment group (TG): N=35 Control group (CG): N=33</p>	<p>Inclusion: attended 2nd – 5th grade most of school day without a closely supervising adult; “high functioning ASD” verified by ADOS, ADI-R, ASSQ; verbal IQ of >60; child mental status examination established that child was able to switch topics in a conversation, had capacity for joint attention and basic social reciprocity, and had knowledge of rules of at least 2 common board games (e.g., chess) and of common school yard games (e.g., handball).</p> <p>Exclusion: currently prescribed psychotropic medication; history of thought disorder, clinical seizure disorder, gross neurological disease, or other major medical disorder.</p>	<p>Treatment (TG): Manualised parent-assisted PEERS-like social skills intervention, Children’s Friendship Training (CFT). CFT given in groups of 10 children (<5 with ASD) over 12, 60-minute, weekly sessions. Parents (mothers) attended separate concurrent sessions.</p> <p>Control (CG): wait list/delayed treatment control</p> <p>Outcomes (completed by):</p> <ul style="list-style-type: none"> - LS: Loneliness (child) - PHS: self-esteem (popularity subscale) (child) - QPQ: quality of get-togethers, subscales (i) hosted, (ii) invited to as guests, (iii) with conflict (iv) engage: interactive activities, (v) disengage: minimally interactive activities (e.g., technology, television) (parent) <p>...continued overpage</p>	<p>There were no significant differences between groups at baseline in socio-demographic, IQ, VABS, and outcome variables, or in number of sessions attended.</p> <p>Key findings:</p> <p>ANCOVA performed at group (TG vs CG) level to predict post-test scores, with baseline scores as covariate. 5 of 13 analyses statistically significant such that TG (CFT) compared with CG improved:</p> <ul style="list-style-type: none"> - in perceived loneliness (child) (p<0.025); - in perceived popularity (child) (p<0.025); - hosted get-togethers: QPQ (parent) (p<0.0001); - decreased disengaged get-together activity: QPG (parent) (p<0.0001) - self-control: SSRS (parent) (p<0.05). <p>...continued overpage</p>	<p>Author conclusions: The programme was feasible and cost effective. In comparison with the control group, the treatment (CFT) group was significantly superior on parent measures of social skill and play date behaviour, and child measures of perceived popularity and loneliness.</p> <p>Reviewer’s comments: ASD diagnoses were independently and rigorously verified. Sample were “high functioning” with some social skills required to be eligible. No direct observational data collected. Parents not blind to allocation and open to reporting bias as active implementers of treatment. Teachers blinded to allocation but some missing data (not obtained for between 11-15% of children). Also, analyses revealed that teachers were less likely to report on children rated as having more externalising (intrusive, aggressive) behaviours and so may be biased.</p> <p>Significant differences for 5 or 13 outcome measures investigated.</p> <p>Source of funding: Research grant from NIH, NIMH, NICHD, NIDCD, NINDS.</p>

Key: ADI-R=Autism Diagnostic Interview-Revised; ADOS=Autism Diagnostic Observation Schedule; ANCOVA=analysis of covariance; ASSQ=High Functioning Autism Spectrum Screening Questionnaire; ASD=autism spectrum disorder; CG=control group; CFT=Children’s Friendship Training; IQ=intelligence quotient; LS=Loneliness Scale; M=mean; NICHD=National Institute of Child Health and Human Development; NIDCD= National Institute of Deafness and other Communication; NINDS=National Institute of Neurological Disorders and Stroke; NIH=National Institute of Health; NIMH=National Mental Institute of Health; PEI=Pupil Evaluation Inventory; PHS=Piers-Harris Self-Concept Scale; QPQ=Quality of Play Questionnaire; RCT=randomised controlled trial; SSRS=Social Skills Rating Scale; TG=treatment group; UCLA=University of California Los Angeles; US=United States of America; VABS=Vineland Adaptive Behavior Scale

Frankel et al, 2010 [25] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
		<p>...continued Follow-up: pre-test, post-test 12 weeks post baseline (immediately after treatment for TG). CG also assessed at 12 weeks post delayed CFT (that is, making the CG an uncontrolled treatment group) but results not reported here as not controlled. 89% completion rate (to post-test): 5 dropouts in TG. Fidelity: weekly fidelity check sheets from research assistants</p>	<p>...continued - SSRS: subscales (i) assertion: making friends, playing well, (ii) self-control: appropriate response to provocation, (iii) externalising: intrusive and aggressive behaviour, (iv) internalising: social withdrawal (parent) - PEI: subscales: withdrawal: shyness, sadness (ii) aggression (teacher)</p>	<p>...continued <u>No significant treatment effects</u> for parent reported ratings of get-togethers invited to, conflict during, engagement in (QPQ); social skill subscales: assertion, externalising, and internalising (SSRS); or teacher-reported pupil withdrawal, or aggression (PEI). - post hoc analyses of reliable change suggested substantially more children in TG than CG demonstrated reliable change post-test.</p>	
Study quality: Internal validity: + Precision: + Applicability: ?					Overall Score: +

Key: ADI-R=Autism Diagnostic Interview-Revised; ADOS=Autism Diagnostic Observation Schedule; ANCOVA=analysis of covariance; ASSQ=High Functioning Autism Spectrum Screening Questionnaire; ASD=autism spectrum disorder; CG=control group; CFT=Children's Friendship Training; IQ=intelligence quotient; LS=Loneliness Scale; M=mean; NICHD=National Institute of Child Health and Human Development; NIDCD= National Institute of Deafness and other Communication; NINDS=National Institute of Neurological Disorders and Stroke; NIH=National Institute of Health; NMIH=National Mental Institute of Health; PEI=Pupil Evaluation Inventory; PHS=Piers-Harris Self-Concept Scale; QPQ=Quality of Play Questionnaire; RCT=randomised controlled trial; SSRS=Social Skills Rating Scale; TG=treatment group; UCLA=University of California Los Angeles; US=United States of America; VABS=Vineland Adaptive Behavior Scale

Koenig et al, 2010 [31]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Recruited from University outpatient clinic in Connecticut. Peer tutors recruited from local school. Participants: 42 children (of 44 randomised) with PDD including 11 with PDD-NOS, 7 with autism, 6 with Asperger syndrome, M age=9.2 years); 33 male (77%), 10 female; all but one identified as Caucasian, one was African American. Treatment group (TG): N=24 Control group (CG); N=18</p>	<p>Inclusion: PDD verified by ADOS, SCQ, PDDBI; full IQ of ≥ 70; aged 8-11 years. Exclusion: psychiatric problems screened by ABC and the CSI including severe aggression, self injury, oppositional behaviour. Peer tutors screened for behavioural and psychological problems through CBCL. Follow-up: pre-test, post-test 16 weeks post baseline (immediately after treatment for TG). Maintenance not investigated. 95% completion rate (to post-test): 1 dropout in TG, 1 in CG. Only one child missed more than two sessions. Fidelity: fidelity checklists sheets completed for each group and 25% of sessions independently reviewed from videotape. All groups achieved at least 75% programme fidelity.</p>	<p>Treatment (TG): Manualised social skills intervention (given in groups of 4-5 participants plus 2 peer tutors) over 16, 75-minute, weekly sessions. Control (CG): wait list/delayed treatment control Outcomes (completed by): - CGI-I: overall symptomatic change in social functioning with respect to two target behaviours identified at baseline (blinded rater interviewing unblinded parent) - SCI: social competence, <i>pro-social</i> index, and <i>social initiation</i> index (parent) - PSS: parent's satisfaction with intervention (parent, post – test only)</p>	<p>There were no significant differences between groups at baseline in socio-demographic, IQ, SCQ, ADOS, and outcome variables. However, TG had marginally fewer children receiving medication than CG (6 vs 10), $p=0.04$. Key findings: <i>Treatment responders</i> were those much improved or very much improved on CGI-I with significantly greater rate of responding for TG (70% or 16/23) than CG (0% or 0/18; $p<0.01$). ANCOVA performed at group (TG vs CG) level to predict post-test scores on SCI, with baseline scores as covariate. No significant differences on either of two indexes. Parent satisfaction: 33 of 42 families returned survey (78%). Over 90% of respondents said children liked the group, 13% reported communication from researchers was unsatisfactory (wanted information about session content to follow-up at home).</p>	<p>Author conclusions: Between groups comparisons showed that children in treatment were rated as improved on the primary outcome measure (change on two target behaviours) but not on the secondary outcome (social competence). Parents reported a high level of satisfaction with the intervention. Reviewer's comments: ASD diagnoses were independently and rigorously verified. No direct observational data collected. No parent involvement in training. Parents not blind to allocation and open to reporting bias. Whilst raters were blinded to allocation, their ratings were based on reports from parents who were unblinded and may have had biased perceptions of improvement. IQ scores were based on unverified past records from school or clinic and were missing for 5 participants (who were all in mainstream classes). Analyses were re-run without these 5 with the same results. Significant differences for 1 of 2 key outcome measures, in addition to reported broad satisfaction of parents with intervention. Maintenance not investigated. Source of funding: Organization for Autism Research; Yale School of Nursing Clinical Initiatives Fund; Research Units on Pediatric Psychopharmacology; NIMH.</p>
Study quality: Internal validity: ? Precision: + Applicability: ?					Overall Score: ?

Key: ABC=Aberrant Behavior Checklist; ADOS=Autism Diagnostic Observation Schedule; ANCOVA=analysis of covariance; ASD=autism spectrum disorder; CBCL=Child Behavior Checklist; CG=control group; CGI-I=Clinical Global Impressions Scale - Improvement; CSI=Children's Symptom Inventory; IQ=intelligence quotient; M=mean; NIMH=National Institute of Mental Health; PDD=Pervasive Developmental Disorder; PDD-NOS= Pervasive Developmental Disorder - Not Otherwise Specified; PDDBI= Pervasive Developmental Disorder Behavior Inventory; PSI=Pro-Social Index; PSS=Parent Satisfaction Survey; RCT=randomised controlled trial; SCI=Social Competence Inventory; SCQ=Social Communication Questionnaire; TG=treatment group; US=United States of America

Lopata et al, 2010 [33]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Intervention conducted on College campus. Recruited using public announcements. Participants: 36 children with higher functioning ASD aged 7-12 years (M age=9.5 years); 34 male (94%), 2 female; 26 with Asperger Syndrome (AS), 9 with PDD-NOS, 1 with HFA; 32 identified as Caucasian, 2 as African American, 1 as Asian-American, 1 as other. Children stratified by age (7-8, 9-10, 11-12 years) before randomisation to group. Treatment group (TG): N=18 Control group (CG); N=18</p>	<p>Inclusion: aged 7-12 years; with a written diagnosis of ASD; minimum cognitive and language function (IQ of >70 on WISC-IV, ≥80 on Verbal Comprehensions Index and Perceptual Reasoning Index of WISC-IV, and ≥80 on CASL); Exclusion: none reported Follow-up: pre-test, post-test 5 weeks post baseline (immediately after treatment for TG). Maintenance not measured. No dropouts reported Fidelity: fidelity check sheets from research assistants and subset (14%) analysed. All groups achieved at least 94% programme fidelity. Reliability of fidelity assessment also checked (97.2% agreement).</p>	<p>Treatment (TG): Manualised “Skillstreaming” programme including activities targeting social skills, face-emotion recognition, interest expansion, and interpretation of non-literal language. 3 treatment groups of different age ranges (7-8, 9-10, 11-12 years), each including approximately 6 children (3 staff). Skillstreaming over 5 weeks, with five 70-minute treatment cycles per day, each cycle included 20 minutes of intensive instruction, and 50 minutes of therapeutic activity to practice and reinforce skills. Parents attended weekly 90-minute sessions covering programme and training to encourage child’s generalisation of skills at home. Included response-cost program to reduce problem behaviours and foster skills acquisition. Control (CG): wait list/delayed treatment control Outcomes (completed by): - ASC: reported skills taught in skillstreaming programme (parent)*</p> <p>...continued overpage</p>	<p>There were no significant differences between groups at baseline in socio-demographic, IQ, and language variables. Key findings: <u>Child & parent report outcomes</u> Between group (TG vs CG) ANCOVA conducted to predict post-test scores, with baseline scores as covariate (Bonferroni correction applied resulting in adjusted critical $p=0.0071$). Effect sizes calculated for mean differences: Cohen’s <i>d</i>. Significant improvements in TG compared with CG in: - ASD symptoms (SRS) (parent) ($p=0.003$, $d=.625$); - skills taught in programme (ASC) (parent) ($p=0.006$, $d=.584$); - Withdrawal (BASC-2) (parent) ($p<0.001$, $d=1.055$); - SKA programme knowledge (child) ($p<0.001$) - idiomatic language (CASL) (child) ($p<0.001$). <u>Clinical significance:</u> For TG, 50% had post-test ASD symptoms (SRS) that decreased from severe to mild-to-moderate, or mild-to-moderate to normal range vs in CG where none changed to a less severe score.</p> <p>...continued overpage</p>	<p>Author conclusions: Standardised effect size estimates were in the medium and large ranges and favoured the treatment group. High levels of parent, child and staff satisfaction reported, along with high levels of treatment fidelity, supporting the feasibility and social validity of the programme. Reviewer’s comments: ASD diagnoses not independently verified however for 29 of 36 participants, historical ADI-R results were made available to support ASD diagnoses. Sample largely male and Caucasian. No direct observational data collected. Parents not blind to allocation and open to bias. Staff assessments made only of treatment group (not reported here as not controlled). Norms only available for CASL for those aged >10 years and so raw scores used. Maintenance not measured. Significant differences for 5 or 7 group differences investigated. Source of funding: Not reported but authors from University of Buffalo, Summit Educational Resources, Carnisius College’s Schol for Autism Research.</p>

Key: AS=Asperger syndrome; ADOS=Autism Diagnostic Observation Schedule; ASC=Adaptive Skillstreaming Checklist; ASD=autism spectrum disorder; BASC-2=Behavior Assessment System for Children – 2nd edition; CASL=Comprehensive Assessment of Spoken Language; CG=control group; DANVA2=Diagnostic Analysis of Nonverbal Accuracy – 2nd edition; HFA=high functioning autism; IQ=intelligence quotient; M=mean; PDD-NOS=pervasive developmental disorder not otherwise specified; RCT=randomised controlled trial; SKA=Skillstreaming Knowledge Assessment; SRS=Social Responsiveness Scale; TG=treatment group; US=United States of America; WISC-IV=Weschler Intelligence Scale for Children-4th edition

Lopata et al, 2010 [33] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
			<p>...continued</p> <ul style="list-style-type: none"> - SRS: ASD symptoms (parent)* - BASC-2: behaviour assessment, Withdrawal, and Social Skills subscales (parent)* - SKA: knowledge of Skillstreaming skills steps (child) - DANVA2: Child Faces subtest assesses ability to identify a simple emotion (child, via computer) - CASL: idiomatic language subtest (child) - Satisfaction surveys (child)(parent)(staff) for TG only, not reported here as not controlled outcomes. <p>* Staff also rated these outcomes for TG only, not reported here as not controlled outcomes.</p>	<p>...continued</p> <p>No significant treatment effects for Social Skills (BASC-2) (parent reported) and emotion recognition (DANVA-2) at Bonferonni adjusted <i>p</i>.</p>	
Study quality: Internal validity: ? Precision: + Applicability: ?					Overall Score: ?

Key: AS=Asperger syndrome; ADOS=Autism Diagnostic Observation Schedule; ASC=Adaptive Skillstreaming Checklist; ASD=autism spectrum disorder; BASC-2=Behavior Assessment System for Children – 2nd edition; CASL=Comprehensive Assessment of Spoken Language; CG=control group; DANVA2=Diagnostic Analysis of Nonverbal Accuracy – 2nd edition; HFA=high functioning autism; IQ=intelligence quotient; M=mean; PDD-NOS=pervasive developmental disorder not otherwise specified; RCT=randomised controlled trial; SKA=Skillstreaming Knowledge Assessment; SRS=Social Responsiveness Scale; TG=treatment group; US=United States of America; WISC-IV=Weschler Intelligence Scale for Children-4th edition

Gantman et al, 2012 [30]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Recruited from a UCLA Help Group (Community mental health agency), regional centres, colleges, universities throughout Southern California, and online research announcements. Participants: 17 young adults (of 19 randomised) aged 18-23 years (M age=20.4 years, SD=1.6); 12 male (70%), 5 female; 4 with autistic disorder, 11 with Asperger Syndrome (AS), 2 with PDD-NOS; 10 identified as Caucasian, 5 as Asian, 2 as Hispanic/Latino; all were in college at least part-time, all lived with their caregivers except one who was in daily interaction with them. Treatment group (TG): N=10 Control group (CG): N=7</p>	<p>Inclusion: aged 18-23 years; previous clinician diagnosis of ASD; has caregiver-reported social problems; English fluency; caregiver English fluency; motivated to join group; composite IQ of ≥ 70 or above; score ≥ 26 on AQ; ≥ 65 on SRS; ≥ 85 on Adaptive Behaviour Composite of VABS. Exclusion: history of major mental illness. Follow-up: pre-test, post-test 14 weeks post baseline (immediately after treatment for TG). Maintenance not investigated. 89% completion rate Fidelity: weekly fidelity check sheets from research assistants.</p>	<p>Treatment (TG): Manualised UCLA PEERS for Young Adults provided in groups of approximately 9-10, over 14, 90-minute, weekly sessions. Parents attended separate concurrent sessions. Control (CG): wait list control Outcomes (completed by): - TYASK: knowledge of PEERS skills (young adult) - SELSA: romantic, social, and family loneliness (young adult) - SRS: ASD symptoms (parent) - SSRS: social skills (parent) - EQ: empathy (parent) - QSQ: number of get-togethers (i) hosted, (ii) invited to (young adult, parent) - SSI: Social skills (young adult) Some qualitative data reported for TG only.</p>	<p>There were no significant differences between groups at baseline in socio-demographic, IQ, VABS, and outcome variables. Key findings: Outcome scores converted to difference scores (post-test–baseline) MANOVA found main effect for group differences: Wilks Lambda=0.34; $F(1,17)=4.27$, $p<0.02$. Univariate analyses found significant improvements in TG compared with CG for young adult self-report outcomes: - knowledge of PEERS (TYASK) ($p<0.01$); and - social loneliness (SELSA) ($p<0.05$) and parent report outcomes: - ASD symptoms (SRS) ($p<0.05$), - social skills (SSRS) ($p<0.01$) - empathy (EQ) ($p<0.05$) - hosted get-togethers: QSQ ($p<0.05$) and invited get-togethers ($p<0.05$) (caregiver report, using non-parametric tests).</p>	<p>Author conclusions: In comparison with the control group, the treatment group reported significantly less loneliness and improved social skills knowledge, while parents/caregivers reported significant improvements in young adult's overall social skills, social responsiveness, empathy, and frequency of get-togethers. Reviewer's comments: ASD diagnoses not independently verified. No direct observational data collected. Parents not blind to allocation. No independent unblinded assessments collected. No maintenance data collected. Only significant results reported and not clear what non-significant tests were performed. For example, results from QSQ completed by young adults and some subscale results not reported. No adjustment to alpha level to allow for multiple tests suggesting some positive results may be chance effects, especially as most of the treatment effects were only significant at alpha of 0.05. Power to detect significant results may have been limited by the small sample size. Fidelity recorded but not reported. Source of funding: Research grants from NIH Training Grant, Organization for Autism Research Grant, and Philip and Aida Siff Educational Foundation.</p>
Study quality: Internal validity: ? Precision: ? Applicability: ?					Overall Score: ?

Key: AS=Asperger syndrome; ASD=autism spectrum disorder; AQ=Autism Spectrum Quotient; CG=control group; EQ=empathy quotient; IQ=intelligence quotient; M=mean; NIH=National Institute of Health; MANOVA=multivariate analysis of variance; PDD-NOS=pervasive developmental disorder not otherwise specified; PEERS=Program for the Education and Enrichment of Relational Skills; QPQ=Quality of Socialization Questionnaire; RCT=randomised controlled trial; SELSA=Social and Emotional Loneliness Scale for Adults; SRS=Social Responsiveness Scale; SSI=Social Skills Inventory; SSRS=Social Skills Rating Scale; TYASK=Test of Young Adult Social Skills Knowledge; TG=treatment group; UCLA=University of California Los Angeles; US=United States of America; VABS=Vineland Adaptive Behavior Scale

Thomeer et al, 2012 [34]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Intervention conducted on College campus during Summer . Participants: 35 children with higher functioning ASD aged 7-12 years (M age=9.3 years); 30 male (86%), 5 female; 25 with Asperger Syndrome (AS), 9 with PDD-NOS, 1 with HFA; 28 identified as Caucasian, 2 as African American, 1 as Asian-American, 1 as Hispanic, 1 as other. Children stratified by age (7-8, 9-10, 11-12 years) before randomisation to group. Treatment group (TG): N=17 Control group (CG); N=18</p>	<p>Inclusion: aged 7-12 years; prior diagnosis and ADI-R verified ASD; minimum cognitive and language function (IQ of >70 on WISC-IV, ≥80 on Verbal Comprehensions Index and Perceptual Reasoning Index on WISC-IV, and ≥80 on CASL); Exclusion: none reported Follow-up: pre-test, post-test 5 weeks post baseline (immediately after treatment for TG) to determine treatment effects, then 2-3 months after post-test to assess maintenance for parent ratings only. No dropouts. Fidelity: fidelity check sheets from research assistants and random subset (31%) analysed. All groups achieved at least 95% programme fidelity. Reliability of fidelity assessment also checked (99% agreement).</p>	<p>Treatment (TG): Manualised “Skillstreaming” programme including activities targeting social skills, face-emotion recognition, interest expansion, and interpretation of non-literal language. 3 treatment groups of different age ranges (7-8, 9-10, 11-12 years), each including 6 children (3 staff) received Skillstreaming over 5 weeks, with five 70-minute treatment cycles per day. Parents attended weekly 90-minute sessions. Included response-cost program. Control (CG): wait list/delayed treatment control Outcomes (completed by): - ASC: reported skills taught in skillstreaming programme (parent)* - SRS: ASD symptoms (parent)* - BASC-2: behaviour assessment, Withdrawal, and Social Skills subscales (parent)* - SKA: knowledge of Skillstreaming skills steps (child) ...continued overpage</p>	<p>There were no significant differences between groups at baseline in socio-demographic, IQ, ADI and language variables. Key findings: <u>Child & parent report outcomes</u> Between group (TG vs CG) ANCOVA conducted to predict post-test scores, with baseline scores as covariate (Bonferroni correction applied resulting in adjusted critical $p=0.0125$). Effect sizes calculated for mean differences: Cohen’s d. Significant improvements in TG compared with CG in: - ASD symptoms (SRS) (parent) ($p=0.007$, $d=.67$); - skills taught in programme (ASC) (parent) ($p=0.001$, $d=.86$); - Social skills (BASC-2) (parent) ($p=0.011$, $d=.70$); - SKA programme knowledge (child) ($p<0.001$, $d=1.15$) - idiomatic language (CASL) (child) ($p<0.001$, $d=0.34$). No significant treatment effects for Withdrawal (BASC-2) (parent reported) and emotion recognition (DANVA-2) at Bonferonni adjusted p. ...continued overpage</p>	<p>Author conclusions: Replication of Lopata et al [33] addressing limitations of that study including verification of ASD diagnoses using ADI-R, and follow-up assessment of maintenance Standardised effect size estimates were predominantly in the medium and large ranges. Reviewer’s comments: ASD diagnoses were independently verified by ADI-R. No direct observational data collected. Parents not blind to allocation and open to bias. Staff assessments made only of treatment group (not reported here as not controlled). High treatment fidelity. Norms only available for CASL for those aged >10 years and so raw scores used. Maintenance was measured. Significant differences for 5 of 7 group differences investigated. Source of funding: John R. Osishei Foundation.</p>

Key: AS=Asperger syndrome; ADI-R=Autism Diagnostic Interview-Revised; ASC=Adaptive Skillstreaming Checklist; ASD=autism spectrum disorder; BASC-2=Behavior Assessment System for Children – 2nd edition; CASL=Comprehensive Assessment of Spoken Language; CG=control group; DANVA2=Diagnostic Analysis of Nonverbal Accuracy – 2nd edition; HFA=high functioning autism; IQ=intelligence quotient; M=mean; PDD=pervasive developmental disorder; RCT=randomised controlled trial; SKA= Skillstreaming Knowledge Assessment; SRS=Social Responsiveness Scale; TG=treatment group; US=United States of America; WISC-IV=Weschler Intelligence Scale for Children-4th edition

Thomeer et al, 2012 [34] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
			<p>...continued</p> <ul style="list-style-type: none"> - DANVA2: Child Faces subtest assesses ability to identify a simple emotion (child) - CASL: idiomatic language subtest (child) <p>* Staff also rated these outcomes for TG only, not reported here.</p>	<p>...continued</p> <p><u>Maintenance</u></p> <p>Improvements for TG at follow-up vs pre-test evident for parent-reported ASC ($p=0.006$, $d=0.47$) and BASC-2 Social skills ($p=0.004$, $d=0.68$). Non significant improvements for parent-reported SRS and BASC-2</p> <p>Withdrawal following Bonferroni-adjusted alpha of 0.008.</p>	
Study quality: Internal validity: + Precision: + Applicability: +					Overall Score: +

Key: AS=Asperger syndrome; ADI-R=Autism Diagnostic Interview-Revised; ASC=Adaptive Skillstreaming Checklist; ASD=autism spectrum disorder; BASC-2=Behavior Assessment System for Children – 2nd edition; CASL=Comprehensive Assessment of Spoken Language; CG=control group; DANVA2=Diagnostic Analysis of Nonverbal Accuracy – 2nd edition; HFA=high functioning autism; IQ=intelligence quotient; M=mean; PDD =pervasive developmental disorder; RCT=randomised controlled trial; SKA= Skillstreaming Knowledge Assessment; SRS=Social Responsiveness Scale; TG=treatment group; US=United States of America; WISC-IV=Weschler Intelligence Scale for Children-4th edition

Andrews, Attwood and Sofronoff, 2013 [10]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: Australia Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: Intervention conducted within postgraduate clinical psychology programme by probationary psychologists. Recruitment was via university, local support group and primary school newsletters/websites. Participants: 58 children with higher functioning ASD aged 7-12 years (M age=9.0 years); 52 male (90%), 6 female; 40 with Asperger Syndrome (AS), 9 with PDD-NOS, 9 with HFA. No ethnicity information provided. Children stratified by age and gender before randomisation to group in blocks of four. Treatment group (TG): N=29 Control group (CG): N=29</p>	<p>Inclusion: aged 7-12 years; ASD diagnosis confirmed by paediatrician as HFA or AS (ASDI verified), or PDD-NOS; minimum IQ of 79 on WASI; "difficulties with affection" indicated from parent report at screening interview. Exclusion: none reported Follow-up: pre-test, post-test 5 weeks post baseline (immediately after treatment for TG) to determine treatment effects, then 3 months after post-test to assess maintenance for TG only. No dropouts post baseline to post-test. Attrition of 21% (7/29) for maintenance follow-up in TG. Fidelity: fidelity check sheets completed by therapists and reviewed during supervision (adherence not reported).</p>	<p>Treatment (TG): affection-focussed cognitive behavioural programme including tuition, visual tools, social stories, role play, strategies, and rehearsal. 9 groups of 3-4 children (2-3 therapists) attended five, 120-minute, weekly sessions. Parents attended concurrent weekly sessions in one group. Control (CG): wait list/delayed treatment control Outcomes (completed by): - WFT: ability to determine purpose of affection (parent) - AOQ: capacity to engage in affectionate behaviour outside immediate family (parent) - GAQ: difficulty in general affection (excessive and inadequate affection) (parent) - SCPQ: social competence (parent) - SCAS: symptoms of child anxiety (parent)</p>	<p>No analyses comparing groups at baseline reported except the lack of differences in AOQ at baseline. Key findings: Repeated measures Group (Treatment vs Control) X Time (pre vs post) MANOVA found significant main effects & Group X Time interaction: $F(5,50)=2.75$, $p=0.029$, partial $\eta^2=.503$. Univariate analyses found significant Group X Time interactions for 1 of the 5 outcomes using Bonferroni adjusted alpha: - AOQ (appropriateness of affectionate behaviour outside of family) ($p=0.002$). Analyses of 3 subscales (giving affection, receiving affection, communicating empathy) found significant difference (at Bonferroni adjusted alpha) for Giving Affection subscale only ($p=0.0001$). <u>No significant treatment effects</u> for Understanding purpose of affection (WFT), difficulty in general affection (GAQ), social competence (SCPQ), or anxiety (SCAS) at Bonferroni adjusted p. Maintenance: No differences for TG between immediate post-test and 3 month follow-up suggesting results maintained.</p>	<p>Author conclusions: Overall this study indicates that children with AS can be taught to interact more appropriately. Reviewer's comments: Potential for bias at preselection interview where difficulties with affection informally discussed. ASD diagnoses independently verified by ASDI and/or paediatrician diagnoses. Ethnicity not reported. Fidelity recorded but adherence not reported. Whether treatment and control groups differed at baseline in sociodemographic or IQ variables (i.e., whether randomisation led to similar groups) was not reported. No direct observational data collected. Parents not blind to allocation and open to bias. Maintenance was measured but no control group for longer follow-up. Bonferroni corrected alpha level used to allow for multiple tests. Significant treatment effects (group differences) for only one of 5 outcomes investigated (AOQ: judging appropriateness of affectionate behaviour to people outside family members). No differences for other correlated and related affection indices. Source of funding: not reported but authors are University affiliated.</p>
<p>Study quality: Internal validity: ? Precision: + Applicability: ?</p>					Overall Score: ?

Key: AOQ=Affection for Others Questionnaire; AS=Asperger syndrome; ASD=autism spectrum disorder; ASDI=Asperger Syndrome Diagnostic Interview; CG=control group; GAQ=General Affection Questionnaire; HFA=high functioning autism; IQ=intelligence quotient; M=mean; MANOVA=multivariate analysis of variance; PDD-NOS=pervasive developmental disorder-not otherwise specified; RCT=randomised controlled trial; SCPQ-P=Social Competence with Peers Questionnaire; SCAS=Spence Child Anxiety Scale; TG=treatment group; WASI=Weschler Abbreviated Scale of Intelligence; WFT= Walk in the Forest Test

Schohl et al, 2014 [13]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: US Study type: randomised controlled trial (RCT) Evidence level: II</p>	<p>Setting: A medium-sized Midwestern city. University-based PEERS programme, led by trained graduates and undergraduates. Participants: 58 adolescents (of 63 randomised) with higher functioning ASD aged 11-16 years (M age=13.6 years, SD=1.5); 47 male (81%), 11 female; 52 identified as caucasian, 3 African American, 1 Asian. Treatment group (TG): N=29 Control group (CG); N=29</p>	<p>Inclusion: aged 11-16 years; has parent-reported social problems; English fluency for participant and parent; previous and current diagnosis of HFA, AS or PDD-NOS verified by ADOS; verbal IQ of 70 or above; interest in making friends and joining group. Exclusion: history of major mental illness, hearing, visual or physical impairments. If more than 2 homework assignments or two sessions missed, dismissed from group. Follow-up: pre-test, post-test 14 weeks post baseline (immediately after treatment for TG). No maintenance data collected. Completion rate post randomisation = 92% with 5 dropouts in CG. Fidelity: weekly fidelity check sheets from research assistants</p>	<p>Treatment (TG): Manualised Program for the Education and Enrichment of Relational Skills (PEERS) - small groups of up to 10 people met over 14, 90-minute, weekly sessions. \$30 given upon completion. Parents attended separate concurrent sessions. Control (CG): wait list control Outcomes (completed by): - knowledge of PEERS skills: TASSK (adolescents) - number of get-togethers (i) hosted, (ii) invited to, (iii) with conflict: QSQ-P (parents), QSQ-A (adolescents) - friendship quality: FQS (adolescents) - social anxiety: SIAS (adolescents) - ASD symptoms: SRS-P (parent) and SRS-T (teacher) - SSRS social skills & SSRS problem behaviours: SSRS-P (parent), SSRS-T (teacher)</p>	<p>There were no significant differences between groups at baseline in demographic variables or outcome measures. No outcome differences between group leader assignment. Key findings: <u>Parent and adolescent variables:</u> repeated measures Group (Treatment vs Control) X Time (pre vs post) MANOVA found significant main effects & Group X Time interaction: Wilks Lambda=0.19; $F(1,56)=13.54$, $p<0.001$. Univariate analyses found significant Group X Time interactions for 4 <i>adolescent outcomes</i>: - TASSK ($p<0.001$); - QSQ-A hosted get-togethers ($p<0.005$); - QSQ-A invited get-togethers ($p<0.01$); - SIAS ($p<0.01$); And 2 <i>parent outcomes</i>: - SRS ($p<0.010$); - SSRS-P Problem Behaviours ($p<0.05$). ...continued overpage</p>	<p>Author conclusions: This study replicates and extends the evaluation of PEERS by Laugesan et al, 2009 [32] (included in the current review). The PEERS group, compared to controls, significantly improved their knowledge of PEERS concepts and friendships skills, their amount of get-togethers, and decreased their social anxiety, core autistic symptoms, and problem behaviours. Reviewer's comments: Mainly white sample. Verification of ASD diagnoses and cognitive ability. Fidelity recorded but adherence not reported. Parents not blind to allocation. Teachers blind to group allocation. Large amount of missing teacher data which may have decreased power in analysis. No direct behavioural observations made. Significant group differences over time for 7 of 14 outcome variables. Bonferroni corrected alpha level used to allow for multiple tests. No maintenance data collected. Source of funding: Research grants from Marquette University, and the Autism Society of Southwestern Wisconsin. ...continued overpage</p>

Key: AS=Asperger syndrome; ADOS=Autism Diagnostic Observation Schedule-Generic; ANOVA= analysis of variance; ASD=autism spectrum disorder; CG=control group; FQS=Friendship Qualities Scale; HFA=high functioning autism; IQ=intelligence quotient; M=mean; MANOVA=multivariate analysis of variance; PDD-NOS=pervasive developmental disorder not otherwise specified; PEERS=Program for the Education and Enrichment of Relational Skills; QSQ=Quality of Socialization Questionnaire; RCT=randomised controlled trial; SD=standard deviation; SIAS=Social Interaction Anxiety Scale; SRS=Social Responsiveness Scale; SSRS=Social Skills Rating Scale; TASSK=Test of Adolescent Social Skills Knowledge; TG=treatment group; US=United States of America

Schohl et al, 2014 [13] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
				<p>...continued</p> <p>Post hoc paired t-tests confirmed TG significantly improved in adolescent-reported:</p> <ul style="list-style-type: none"> - PEERS knowledge (TASSK), - hosted & invited get-togethers (QSQ-A), - in decreased social anxiety (SIAS); - parent-reported decreases in ASD symptoms (SRS-P) and - problem behaviour (SSRS-P). <p><u>Teacher variables:</u></p> <p>2 X 2 mixed model repeated measures ANOVA's: significant Group X Time interaction for SSRS-T Problem behaviours, Wilks Lambda=0.91; F(1,39)=3.93, p<0.05. Post hoc paired t-test confirmed decrease in reported problem behaviours in TG, not CG.</p> <p><u>No significant treatment effects</u> for adolescent reported conflict at get-togethers (QSQ-A-R) or friendship quality (FQS); parent reported ratings of get-togethers (QSQ-P) or SSRS-P social skills, or teacher-rated SSRS-T social skills or ASD symptoms (SRS-T).</p>	
Study quality: Internal validity: + Precision: + Applicability: +					Overall Score: +

Key: AS=Asperger syndrome; ADOS=Autism Diagnostic Observation Schedule-Generic; ANOVA= analysis of variance; ASD=autism spectrum disorder; CG=control group; FQS=Friendship Qualities Scale; HFA=high functioning autism; IQ=intelligence quotient; M=mean; MANOVA=multivariate analysis of variance; PDD-NOS=pervasive developmental disorder not otherwise specified; PEERS=Program for the Education and Enrichment of Relational Skills; QSQ=Quality of Socialization Questionnaire; RCT=randomised controlled trial; SD=standard deviation; SIAS=Social Interaction Anxiety Scale; SRS=Social Responsiveness Scale; SSRS=Social Skills Rating Scale; TASSK=Test of Adolescent Social Skills Knowledge; TG=treatment group; US=United States of America

Yoo et al, 2014 [35]					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
<p>Country: South Korea</p> <p>Study type: randomised controlled trial (RCT)</p> <p>Evidence level: II</p>	<p>Setting: Recruited from child psychiatric clinics, promotion to psychiatrists, and support groups for parents of children with ASD.</p> <p>Participants: 47 (of 55 randomised) verbally fluent adolescents with ASD aged 12-18 years (M age=14 years); 45 male (96%), 2 female. Korean participants. Treatment group (TG): N=23 Control group (CG); N=24</p>	<p>Inclusion: aged 12-18 years; enrolled between 6th grade elementary to 3rd grade high school; has parent-reported social problems; previous diagnosis/suspicion of ASD verified by ADOS and ADI-R; verbal IQ of 65 or above; interest in joining group.</p> <p>Exclusion: history of major mental illness; hearing, visual, physical or neurological impairments; current problems with aggressive behaviour or severe oppositional tendency.</p> <p>Follow-up: pre-test, post-test 14 weeks post baseline (immediately after treatment for TG) to determine treatment effects, then 3 months after post-test to assess maintenance with the exception of ADOS and VABS.</p> <p>CG also assessed post delayed PEERS but results not reported here as not controlled.</p>	<p>Treatment (TG): Manualised Program for the Education and Enrichment of Relational Skills (PEERS) – translated, peer-reviewed, pre-tested, and modified to be culturally appropriate – taught in 3 small groups of 6-10 people over 14, 90-minute weekly sessions. Parents (mothers and in one case both parents) attended separate concurrent sessions.</p> <p>Control (CG): wait list/delayed treatment control</p> <p>Outcomes (Korean versions) (completed by):</p> <p><u>Relating to adolescent:</u></p> <ul style="list-style-type: none"> - ASD diagnostic symptoms: ADOS (assessor observation) - interpersonal relationship skills: socialisation domain of the VABS (parent) - knowledge of PEERS skills: TASSK-R (adolescent) - quality of get-togethers: QPQ (parent, and adolescents) - SSRS social behaviour: (adolescent) - ASD symptoms: SCQ (parent) <p>...continued overpage</p>	<p>There were no significant differences between groups at baseline in sociodemographic variables, ASD symptoms or some outcome measures except a difference on one sub-scale (imagination) on ADOS (increased for CG). No outcome differences between group leader assignment.</p> <p>Key findings:</p> <p><u>Parent and adolescent variables:</u></p> <p>Repeated measures Group (treatment & control) by Time (baseline & 14 weeks) ANOVAs suggest improvements for the TG in:</p> <ul style="list-style-type: none"> - communication and social interaction: ADOS (p<0.01) - interpersonal relationships and play/leisure time: VABS (p<0.01) - PEERS social skills knowledge: TASSK-R (p<0.01) - decreased depression: CDI (p<0.05) - decreased parental state anxiety: STAI-S (p<0.01). <p>Models constructed to control for likely covariates led to similar results. Model controlling for SES, teen age, sex, IQ, medication, and maternal age, education, and state anxiety remains significant (p<0.05).</p> <p>...continued overpage</p>	<p>Author conclusions: This study replicates and cross-culturally validates the evaluation of PEERS by Laugesan et al, 2009 [32] (included in the current review) in a Korean sample. Participants receiving PEERS, compared to controls, significantly improved in social skills knowledge, interpersonal skills, and play/leisure skills, as well as a decrease in depressive symptoms and ASD symptoms. The PEERS intervention appears efficacious for teens with ASD in Korea with modest cultural adjustment.</p> <p>Reviewer's comments: Verification of ASD diagnoses and cognitive ability. Fidelity recorded but adherence not reported. Parents not blind to allocation. Direct behavioural observations using diagnostic ADOS instrument by therapist. No adjustment to alpha level to allow for multiple ANOVAs suggesting some positive results may be chance effects. Not clear whether 3-month maintenance analyses included CG data post delayed treatment.</p> <p>Source of funding: research grants from Seoul University Bundang Hospital, Korea Healthcare Technology, and the Ministry of Health and Welfare. Some authors receive royalties from sale of PEERS manual.</p>

Key: AS=Asperger syndrome; ASDS=Asperger Syndrome Diagnostic Scale; ADOS=Autism Diagnostic Observation Schedule-Generc; ADI-R=Autism Diagnostic Interview-Revised; ANOVA=analysis of variance; ASD=autism spectrum disorder; BDI=Beck Depression Inventory; CBCL=Child Behavior Checklist; CDI=Child Depression Inventory (CD(I)); CG=control group; IQ=intelligence quotient; M=mean; PEERS=Program for the Education and Enrichment of Relational Skills; QPQ=Quality of Play Questionnaire; RCT=randomised controlled trial; SD=standard deviation; SCQ=Social Communication Questionnaire; SES=socio-economic status; SRS=Social Responsiveness Scale; SSRS=Social Skills Rating Scale; STAI-T/STAI-S=State and Trait Anxiety Inventory, trait and state versions; STAIC-T/STAIC-S=State and Trait Anxiety Inventory for Children, trait and state versions; TASSK-R=Test of Adolescent Social Skills Knowledge-Revised; TG=treatment group; VABS=Vineland Adaptive Behavior Scale

You et al, 2014 [35] <i>continued</i>					
Country, study type, aims	Participants	Inclusion and exclusion criteria	Exposure, comparison and outcome measures	Results	Conclusions, quality issues
		<p>...continued Completion rate post randomisation=85% with 5 dropouts in TG and 3 in CG. Homework completion=43% Fidelity: monitored by research assistants</p>	<p>...continued - ASD symptoms: SRS-P (parent) - Asperger syndrome behavioural characteristics (ASDS) (parent) - depression: CDI (adolescent) - adolescent's state or trait anxiety: STAI-S/T (adolescent) - adolescent emotional problems: CBCL (parent) <u>Relating to parent:</u> - parent's state anxiety: STAI-S (parent) - parent's trait anxiety: STAI-T (parent) - parent's depression: BDI (parent)</p>	<p>...continued However treatment effects for parent-rated quality of get-togethers (QPQ) and emotional problems (CBCL) appear confounded by group differences in adjusted means at baseline in this model. <u>No significant treatment effects</u> for adolescent rated get-together quality (QPQ); social behaviour (SSRS); ASD social impairments (SCQ, SRS); Asperger behaviour characteristics (ASDS); child's state or trait anxiety (STAI-S/T); or parent's trait anxiety (STAI-T) or depression (BDI). <u>Maintenance:</u> repeated measures ANOVAs found scores post PEERS were not significantly different 3 months later with the exception of TASSK scores which decreased at 3 months, though were still increased compared with baseline.</p>	
Study quality: Internal validity: + Precision: + Applicability: +					Overall Score: +

Key: AS=Asperger syndrome; ASDS=Asperger Syndrome Diagnostic Scale; ADOS=Autism Diagnostic Observation Schedule-Generic; ADI-R=Autism Diagnostic Interview-Revised; ANOVA=analysis of variance; ASD=autism spectrum disorder; BDI=Beck Depression Inventory; CBCL=Child Behavior Checklist; CDI=Child Depression Inventory (CD(I)); CG=control group; IQ=intelligence quotient; M=mean; PEERS=Program for the Education and Enrichment of Relational Skills; QPQ=Quality of Play Questionnaire; RCT=randomised controlled trial; SD=standard deviation; SCQ=Social Communication Questionnaire; SES=socio-economic status; SRS=Social Responsiveness Scale; SSRS=Social Skills Rating Scale; STAI-T/STAI-S=State and Trait Anxiety Inventory, trait and state versions; STAIC-T/STAIC-S=State and Trait Anxiety Inventory for Children, trait and state versions; TASSK-R=Test of Adolescent Social Skills Knowledge-Revised; TG=treatment group; VABS=Vineland Adaptive Behavior Scale

References

Studies included for appraisal in the current review are identified by an asterix (*)

- [1] Ministries of Health and Education. New Zealand autism spectrum disorder guideline. 2008. Available from <http://www.health.govt.nz/publication/new-zealand-autism-spectrum-disorder-guideline>
- [2] Broadstock M. New Zealand Autism Spectrum Disorder Guideline supplementary paper on implications of DSM-5 for the diagnosis of ASD. Christchurch: INSIGHT Research; 2014. Available from <http://www.health.govt.nz/our-work/disability-services/disability-projects-and-programmes/autism-spectrum-disorder-guideline/keeping-guideline-date>
- [3] Broadstock M. New Zealand Autism Spectrum Disorder Guideline supplementary paper on gastrointestinal problems in young people. Christchurch: INSIGHT Research; 2013. Available from <http://www.health.govt.nz/our-work/disability-services/disability-projects-and-programmes/autism-spectrum-disorder-guideline/keeping-guideline-date>
- [4] Broadstock, M. New Zealand Autism Spectrum Disorder Guideline supplementary paper on supported employment services. Wellington: New Zealand Guidelines Group; 2012. Available from <http://www.health.govt.nz/our-work/disability-services/disability-projects-and-programmes/autism-spectrum-disorder-guideline/keeping-guideline-date>
- [5] Broadstock M. New Zealand Autism Spectrum Disorder Guideline supplementary paper on three pharmacological interventions. Wellington: New Zealand Guidelines Group; 2011. Available from <http://www.health.govt.nz/our-work/disability-services/disability-projects-and-programmes/autism-spectrum-disorder-guideline/keeping-guideline-date>
- [6] Broadstock M. New Zealand Autism Spectrum Disorder Guideline supplementary paper on applied behaviour analysis. Wellington: New Zealand Guidelines Group; 2010. Available from <http://www.health.govt.nz/our-work/disability-services/disability-projects-and-programmes/autism-spectrum-disorder-guideline/keeping-guideline-date>
- [7] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC: APA Press; 2013.
- [8] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th edition (Text revision) ed. Washington, DC: APA Press; 2000.
- [9] White SW, Keonig K and Scahill L. Social skills development in children with autism spectrum disorders: A review of the intervention research. *Journal of Autism and Developmental Disorders* 2007;37(10):1858-68.
- *[10] Andrews L, Attwood T and Sofronoff K. Increasing the appropriate demonstration of affectionate behavior, in children with Asperger syndrome, high functioning autism, and PDD-NOS: A randomized controlled trial. *Research in Autism Spectrum Disorders* 2013;7(12):1568-78.
- [11] Baron Cohen S, Leslie A and Frith U. Does the autistic child have a 'theory of mind'? *Cognition* 1985;21:37-46.

- [12] Koenig K, De Los Reyes A, Cicchetti D, Scahill L and Klin A. Group Intervention to Promote Social Skills in School-age Children with Pervasive Developmental Disorders: Reconsidering Efficacy. *Journal of Autism and Developmental Disorders* 2009;39(8):1163-72.
- *[13] Schohl KA, Van Hecke AV, Carson AM, Dolan B, Karst J, et al. A replication and extension of the PEERS intervention: Examining effects on social skills and social anxiety in adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2014;44(3):532-45.
- [14] Weiss MJ. Teaching social skills to people with autism. *Behavior Modification* 2001;25(5):785-802.
- [15] Mitchell K, Regehr K, Reaume J and Feldman M. Group social skills training for adolescents with asperger syndrome or high functioning autism. *Journal of Developmental Disabilities* 2010;16(2):52-63.
- [16] Shtayermman O. Peer victimisation in adolescents and young adults diagnosed with Asperger's syndrome: A link to depressive symptomatology, anxiety symptomatology and suicidal ideation. *Issues in Comprehensive Pediatric Nursing* 2007;30:87-107.
- [17] Cappadocia MC and Weiss JA. Review of social skills training groups for youth with Asperger Syndrome and High Functioning Autism. *Research in Autism Spectrum Disorders* 2011;5(1):70-78.
- [18] Howlin P. Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism* 2000;4:63-83.
- [19] Shattuck PT, Orsmond GI, Wagner M and Cooper BP. Participation in social activities among adolescents with an autism spectrum disorder. *PloS one* 2011;6(11):e27176.
- [20] Sterling L, Dawson G, Estes A and Greenson J. Characteristics associated with presence of depressive symptoms in adults with autism spectrum disorders. *Journal of Autism & Developmental Disorders* 2008;37:1969-86.
- [21] Rubin KH and Rose-Krasnor L. Interpersonal problem solving. In: V. B. Van Hasselt and M. Hersen, editors. *Handbook of Social Development*. New York: Plenum; 1992
- [22] Katz LG and McClennan DE. *Fostering children's social competence*. Washington, D.C.: National Association for the Education of Young Children; 1997.
- [23] Reichow B and Volkmar FR. Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders* 2010;40(2):149-66.
- *[24] Kaat AJ and Lecavalier L. Group-based social skills treatment: A methodological review. *Research in Autism Spectrum Disorders* 2014;8(1):15-24.
- *[25] Frankel F, Myatt R, Sugar C, Whitham C, Gorospe CM, et al. A randomized controlled study of parent-assisted children's friendship training with children

- having autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2010;40(7):827-42.
- *[26] Reichow B, Steiner AM and Volkmar F. Cochrane review: social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Evidence-based child health : a Cochrane review journal* 2013;8(2):266-315.
- [27] Mastergeorge AM, Rogers SJ, Corbett BA and et al. Nonmedical interventions for autism spectrum disorders. In: S. Ozonoff, S. J. Rogers and R. L. Hendren, editors. *Autism Spectrum Disorders: a research review for practitioners*. Washington, DC: American Psychiatric Publishers; 2003
- [28] National Health and Medical Research Council. *NHMRC additional levels of evidence and grades for recommendaitons for developers of guidelines: pilot program 2005-2007*. Canberra, Australia: NHMRC; 2008.
- *[29] Beaumont R and Sofronoff K. 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* 2008;49(7):743-53.
- *[30] Gantman A, Kapp SK, Orenski K and Laugeson EA. Social skills training for young adults with high-functioning autism spectrum disorders: A randomized controlled pilot study. *Journal of Autism and Developmental Disorders* 2012;42(6):1094-103.
- *[31] Koenig K, White SW, Pachler M, Lau M, Lewis M, et al. Promoting social skill development in children with pervasive developmental disorders: a feasibility and efficacy study. *Journal of Autism & Developmental Disorders* 2010;40(10):1209-18.
- *[32] Laugeson EA, Frankel F, Mogil C and Dillon AR. Parent-assisted social skills training to improve friendships in teens with autism spectrum disorders. *Journal of autism and developmental disorders* 2009;39(4):596-606.
- *[33] Lopata C, Thomeer ML, Volker MA, Toomey JA, Nida RE, et al. RCT of a manualized social treatment for high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2010;40(11):1297-310.
- *[34] Thomeer ML, Lopata C, Volker MA, Toomey JA, Lee GK, et al. Randomized Clinical Trial Replication of a Psychosocial Treatment for Children with High-Functioning Autism Spectrum Disorders. *Psychology in the Schools* 2012;49(10):942-54.
- *[35] Yoo HJ, Bahn G, Cho IH, Kim EK, Kim JH, et al. A Randomized Controlled Trial of the Korean Version of the PEERS (R) Parent-Assisted Social Skills Training Program for Teens With ASD. *Autism Research* 2014;7(1):145-61.
- [36] Mandelberg J, Laugeson EA, Cunningham TD, Ellingsen R, Bates S, et al. Long-term treatment outcomes for parent-assisted social skills training for adolescents with autism spectrum disorders: The UCLA PEERS program. *Journal of Mental Health Research in Intellectual Disabilities* 2014;7(1):45-73.
- [37] Chang Y-C, Laugeson E, Gantman A, Ellingsen R, Frankel F, et al. Predicting treatment success in social skills training for adolescents with autism spectrum

disorders: The UCLA Program for the Education and Enrichment of Relational Skills. *Autism* 2014;18(4):467-70.