Systematic Review of the Properties of Tools Used to Measure Outcomes in Anxiety Intervention Studies for Children with Autism Spectrum Disorders

Sarah Wigham, Helen McConachie*

Institute of Health & Society, Newcastle University, Newcastle, United Kingdom

Abstract

Background: Evidence about relevant outcomes is required in the evaluation of clinical interventions for children with autism spectrum disorders (ASD). However, to date, the variety of outcome measurement tools being used, and lack of knowledge about the measurement properties of some, compromise conclusions regarding the most effective interventions.

Objectives: This two-stage systematic review aimed to identify the tools used in studies evaluating interventions for anxiety for high-functioning children with ASD in middle childhood, and then to evaluate the tools for their appropriateness and measurement properties.

Methods: Electronic databases including Medline, PsychInfo, Embase, and the Cochrane database and registers were searched for anxiety intervention studies for children with ASD in middle childhood. Articles examining the measurement properties of the tools used were then searched for using a methodological filter in PubMed, and the quality of the papers evaluated using the COSMIN checklist.

Results: Ten intervention studies were identified in which six tools measuring anxiety and one of overall symptom change were used as primary outcomes. One further tool was included as it is recommended for standard use in UK children’s mental health services. Sixty three articles on the properties of the tools were evaluated for the quality of evidence, and the quality of the measurement properties of each tool was summarised.

Conclusions: Overall three questionnaires were found robust in their measurement properties, the Spence Children’s Anxiety Scale, its revised version – the Revised Children’s Anxiety and Depression Scale, and also the Screen for Child Anxiety Related Emotional Disorders. Crucially the articles on measurement properties provided almost no evidence on responsiveness to change, nor on the validity of use of the tools for evaluation of interventions for children with ASD.

PROSPERO Registration number: CRD42012002684.

Introduction

The choice of relevant outcomes, and of robust tools to measure those, is a vital stage in the design of evaluation of clinical interventions for children. Where tools are reliable and valid, and outcomes important to children and families, the findings can inform parents, clinicians, researchers, service providers and policy makers about which interventions are most effective. However, to date the outcome measures used for intervention trials for children with autism spectrum disorder (ASD) are too varied to allow sensible decisions about what interventions might be most effective [1,2].

Meta-analyses can increase the power of findings by pooling data from individual studies. For example, a meta-analysis of the Revised Children’s Manifest Anxiety Scale across 43 studies has found evidence of validity and responsiveness to treatment [3]. Cross-study syntheses of outcome evidence such as this are much needed in the field of ASD, because individual trials are in the main very small and include broad age groups [4,5]. There have been discussions of these problems and suggestions of which outcome measures to use [6,7], but no widespread uptake in ASD studies.

The focus of the current review is on how to choose appropriate and robust tools to measure outcomes of interventions for a common problem encountered by high-functioning children with ASD – how to cope with symptoms of anxiety in the period of middle childhood. With around 40 per cent having symptoms at the severity of an anxiety disorder [8], and the prevalence of ASD...
being around 1 per cent [9], this is an important public health problem. In the UK, a government initiative titled ‘Increasing Access to Psychological Therapies’ (IAPT) [10] has since 2012 been extended to children’s mental health services, with cognitive behaviour therapy (CBT) for problems such as anxiety and depression as one of the core strands. Outcome monitoring is embedded in the programme.

It is important to have a choice of reliable measurement tools for a particular health condition in order to capture relevant outcomes, and different points of view including patient reported outcomes [11]. A choice of measurement tools also facilitates answering a range of research questions, tailored to the objectives of the intervention, ideally meeting the needs of particular developmental stages [12], and allowing different tools to be used for study outcome evaluation and for selection criteria [13]. Without choice of appropriate tools the benefits of an intervention may be missed or inflated [11,14].

In this systematic review, the tools used to measure outcomes in evaluations of clinical interventions for anxiety in children with high-functioning ASD in middle childhood are identified and their quality assessed. Middle childhood is defined here as 8 to 14 years of age during which time children will be entering puberty, beginning some level of personal independence from their parents, and experiencing transition between primary and secondary school. We focus on high-functioning ASD as the children are likely to be able to participate in verbally-loaded interventions such as CBT, even although the prevalence of comorbid psychiatric conditions is similar across IQ and levels of adaptive behaviour [15]. This systematic review will facilitate recommendations of robust tools for use in anxiety intervention trials for children with high-functioning ASD in middle childhood.

The review was conducted in two stages. In stage 1, identification of tools was done by systematic search for literature describing studies of treatment interventions for anxiety in ASD in middle childhood. Then in stage 2, searches focused on the tools used to measure primary outcomes, and articles about these tools were examined for evidence of appropriateness and measurement properties.

### Table 1. Stage 1 review search terms.

1. (ASC or ASD or Asperger$^*$ or Autism$^*$ or high functioning or communicat$^*$ or Kanner$^*$ or language delay$^*$ or pervasive developmental disorder$^*$ or language disorder$^*$ or HFA or autistic disorder or child development disorders).tiab
2. (child$^*$ or school$^*$ or pediatric$^*$ or paediatric$^*$ or special needs or teenager$^*$ or adolescent$^*$ or youth$^*$).tiab
3. (behaviour$^*$ or intervention$^*$ or program$^*$ or rehabilitation$^*$ or therapy$^*$ or train$^*$ or treatment$^*$ or verbal or cognitive behaviour$^*$ therapy or CBT or pervasive therapy or outcome assessment or treatment outcome).tiab
4. (worry$^*$ or stress$^*$ or phobia$^*$ or anxiety or phobic disorders).tiab
5. randomi$^*$ or ed controlled trial.tiab
6. random$^*$
tiab
7. comparative stud$^*$
tw
8. prospective stud$^*$
tw
9. treatment effectiveness evaluation/
10. intervention$^*$
tiab
11. 5 or 6 or 7 or 8 or 9 or 10
12. 1 and 2 and 3 and 4 and 11
13. limit 12 to English language; 1992 – current; age 8 to 14 years

doi:10.1371/journal.pone.0085268.t001

### Review Methods: Stage 1

The review protocol was registered online with the International Prospective Register of Systematic Reviews (Registration number: CRD42012002684) and can be accessed at [http://www.crd.york.ac.uk/PROSPERO/prospero.asp](http://www.crd.york.ac.uk/PROSPERO/prospero.asp). The protocol also pertains to social skills interventions, though only the anxiety interventions and outcome tools are reported here. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards are followed in this report (see Checklist S1).

### Search Strategy

The following electronic bibliographic databases were searched: MEDLINE, EMBASE, ERIC, PsycINFO, The Cochrane Library (Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), and Cochrane Methodology Register). The search strategy included the terms shown in Table 1 which were combined using database-specific filters, where these were available. The search was restricted to articles in English, and those published between 1992 and February 2013, the date when the last searches were run. The term Asperger Syndrome was first included as a separate diagnosis in the WHO International Classification of Diseases in 1992 [16] so we expected separate identification of groups of children with ability in the average range to be more frequent and consistent in studies after this date.

### Selection Criteria

Anxiety was clinically defined as in the International Classification of Diseases [16] and the Diagnostic and Statistical Manual of Mental Disorders [17]. The interventions included cognitive and behavioural approaches, and excluded drug trials, physiological interventions (e.g., biofeedback) and purely physical interventions (e.g., massage). Intervention studies where a broad range of skills were the target (e.g., social skills, or drama classes) were excluded. The interventions included were ameliorative, preventative or educational, aimed at managing and regulating emotional reactions which may be precursors to anxiety disorder.

Studies were included when over 50% of participants were aged 8 to 14 years old, or the mean age of the ASD sample was within
this range, so that measures were likely to be appropriate across the target age range for the review. Where child participants had a range of differing diagnoses, the study was included if ASD outcome data were presented separately, and if half or more of participants have ASD.

Group studies with designs including before-and-after, controlled trials, quasi-experimental, and randomised controlled trials (RCT) were included. Studies which used only observational methods of recording outcomes (e.g. event recording) were excluded. The review was restricted to articles published in English.

One reviewer (SW) screened the titles and abstracts of articles; where there was doubt whether an article met the inclusion criteria it was included. Full text sifting was by one reviewer (SW); any ambiguous papers were discussed with the second reviewer (HM) to reach consensus. The references of the selected articles were searched.

**Data extraction**

Data extraction was performed by one reviewer (SW) using a previously tested data extraction form. The following information was noted: participant characteristics, focus of intervention, outcome tools used, domains captured, and by whom the tool was reported/measured.

**Results: Stage 1**

The searches retrieved 750 articles from which 10 articles were retained [18–27]. See Figure 1 for search strategy flow diagram.

Nine articles report on seven RCTs of adapted CBT for anxiety delivered to high-functioning children with ASD in middle childhood. These studies varied in sample size from 22 to 71 participants, used varied approaches and materials, and included from 6 group sessions [26] to 16 group [23] or individual sessions [19–21]. The before-and-after study included 6 participants in 16 group sessions of CBT [25]. All but two [23;25] included training for parents. Taken together the studies provide encouraging evidence that CBT can be efficacious for children with ASD and anxiety disorder.

Seven different primary outcome tools were used in these studies. The Anxiety Disorders Interview Schedule (ADIS) [28] is a clinician-administered interview. Five are parent and self-report child anxiety questionnaires [29–33]. One further tool is a clinician or researcher rating of overall improvement, the Clinical Global Impressions – Improvement (CGI-I) [34]. No intervention studies meeting our inclusion criteria used the Revised Children’s Anxiety and Depression Scale (RCADS) [35]; however as it is an IAPT recommended outcome tool it was also included in stage 2 (Table 2). None of the tools was developed specifically for children with ASD. All of the tools were developed in English (though at stage 2 some articles evaluating the measurement properties of the SCARED were on revised versions developed in Dutch).

**Review Methods: Stage 2**

In order to assess the measurement properties of the tools, a comprehensive search was conducted using a methodological
Table 2. Characteristics of the tools used in included studies.

| Measure                                      | Version | Year published | Aim of tool                                      | Number of items | Subscales                                                                 | Response options | Format                          | Used in references |
|----------------------------------------------|---------|----------------|-------------------------------------------------|-----------------|---------------------------------------------------------------------------|------------------|---------------------------------|--------------------|
| Anxiety Disorders Interview Schedule (ADIS)  | C       | 1996           | Measure of anxiety and related disorders.        | (depends on entry questions) | Separation anxiety; social phobia; specific phobia; panic disorder; agoraphobia; generalised anxiety; obsessive-compulsive; post-traumatic stress disorder | 9 point severity scale | Semi structured diagnostic interview | 19, 20             |
|                                              | P       |                |                                                 |                 |                                                                           |                  |                                 | 18, 19, 20, 21, 22 |
| Multidimensional Anxiety Scale for Children (MASC) | C + P  | 1997           | Measure of child anxiety                        | 39              | Physical symptoms; harm avoidance; social anxiety; separation anxiety.   | 4 point scale   | Questionnaire                   | 19                 |
| Revised Children's Anxiety and Depression Scale (RCADS) | C       | 2000           | Measure of child anxiety and depression         | 47              | Separation anxiety; social phobia; generalized anxiety; panic; obsessive-compulsive; major depressive disorder. | 4 point scale   | Questionnaire                   | **                 |
|                                              | P       | 2010           |                                                 |                 |                                                                           |                  |                                 |                    |
| Short                                        |         | 2012           | Anxiety; depression                             | 25              |                                                                           |                  |                                 |                    |
| Screen for Child Anxiety Related Emotional Disorders (SCARED) | C + P  | 1997           | Measure of child anxiety disorders              | 38              | Panic; generalized anxiety; separation anxiety; social phobia; school phobia. | 3 point scale   | Questionnaire                   | 18, 27             |
|                                              |         | 1999           |                                                 |                 |                                                                           |                  |                                 |                    |
|                                              | *71 C + P | 2009           |                                                 | 71              | Panic; generalized anxiety; social phobia; separation anxiety; obsessive-compulsive; post-traumatic stress; specific phobias (animal); blood/injury; situational. |                  |                                 |                    |
|                                              | *Revised C | 1999           |                                                 | 66              |                                                                           |                  |                                 |                    |
| Social Worries Questionnaire (SWQ)           | C       | 1995           | Behavioural indicators of social anxiety in children | 13              | none                                                                      | 3 point scale   | Questionnaire                   | 26                 |
|                                              | P       | 2010           |                                                 |                 |                                                                           |                  |                                 |                    |
Data extraction method

Once identified, the methodological quality of each article was examined using the COSMIN checklist (COnsensus based Standards for the selection of health based Measurement INstruments). The checklist considers 9 properties of measurement, each with multiple items rated on a 4 point scale: internal consistency, reliability, measurement error, content validity, structural validity, hypothesis testing, criterion validity, responsiveness to change (and cross-cultural validity, not considered in the present review). For each article, the properties addressed are given an overall rating of excellent, good, fair, poor based on the lowest item rating awarded [37]. The checklists were completed by one reviewer (SW) with frequent discussion of ratings with a second reviewer (HM) to reach consensus. To check reliability the second reviewer independently rated 10% of the articles using the checklist. Agreement on final rating of each property was 71.5%.

Evidence Synthesis

The quantitative findings in each study were then given a quality rating of positive, indeterminate or negative for each measurement property examined [38]. For example, internal consistency is considered positive where Cronbach’s alpha is equal to or greater than 0.70; criterion validity is considered positive where there are convincing arguments that the gold standard is ‘gold’ and correlation is equal to or greater than 0.70.

Finally the quality ratings for the findings were considered in conjunction with the quality rating for the level of evidence in the articles about each tool [38]. This synthesis records strong evidence (+++ or ++++) where several methodologically good articles, or one excellent article, find consistent evidence for or against a measurement property; moderate evidence (++ or ++) for several methodologically fair, or one good study; a rating of limited (+ or −) for one study of fair quality; and otherwise a rating of conflicting evidence (+/−) or unknown (?) evidence [38].

Results: Stage 2

The search in PubMed produced 1096 articles from which 63 were retained for data extraction (Figure 2). The study population characteristics for these articles are shown in Table 3. Only four articles assessing measurement properties included an ASD sample, reporting on use of five of the tools (i.e. not RCADS, RCMAS or CGI). The majority of the studies were carried out in the USA. The methodological quality of each article is presented in Table 4. None of the articles had looked at measurement error, so this property is not included in the table. Only one article reported responsiveness to change. The synthesised evidence on the quality of the measurement properties of the individual tools is shown in Table 5. To aid interpretability [39], it is important to have evidence on differences in scores between subgroups (including normative data) and this was available in many of the articles; however, no article reported on levels of minimal important change, nor on floor and ceiling effects.

The ADIS is a clinical interview, with entry-level questions which determine which areas of anxiety disorder are explored. The recommended procedure is that parent and child are interviewed separately, and then the interviewer determines the disorder diagnoses and clinical severity rating. When the separate interviews are compared, agreement is low both at the level of whether a disorder is indicated and at symptom level (though one study [42] found the latter to be higher). As a clinical interview,
some measurement properties such as internal consistency and content validity have not been studied, with the latter presumably assumed because the measure was developed from the Diagnostic and Statistical Manual of Mental Disorders [17]. In many studies ADIS is used as the ‘gold standard’ against which questionnaire measures are compared. Its strengths lie in inter-rater reliability, and evidence also of test-retest reliability.

Turning to the questionnaire measures, evidence for internal consistency of the parent and child versions of the SCAS was strong for total and subscale scores, apart from the fear of physical injuries subscale [33;40;94;95] and generalised anxiety disorder (GAD) subscale [40]. Test-retest reliability for child report was $r = .60$ [33] at 6 months, and $r = .63$ at 3 months [94] which seems acceptable (the COSMIN criterion of $r \geq 0.80$ may be set unduly high for a subjective measure of feelings).

Evidence for the structural validity of the six factor structure for the SCAS child version was strong [33;91;94], though lower for the parent version the confirmatory factor analysis finding only acceptable evidence of fit [98] (root mean square error of approximation (RMSEA) = .075) [92]. Criterion validity of the SCAS was supported by significantly higher scores in a clinical than a non-clinical group [33;92;95], more than 80% of those with an anxiety disorder correctly classified, and discrimination between disorders good apart from GAD and panic-agoraphobia [92]. Convergent and divergent validity were demonstrated by significantly higher correlations between SCAS parent and the Child Behavior Checklist (CBCL) internalising than externalising scales [92], and higher between SCAS child and the Strengths and Difficulties Questionnaire emotional subscale than with the conduct or hyperactivity subscales [91].

Findings for parent-child agreement for the SCAS depended on the analysis conducted. Using ANOVA, it was found that parents rated significantly higher than children on all subscales apart from OCD and panic-agoraphobia [93]. In contrast, studies reporting correlations [40;95;96] consistently found $r > .50$ on total and subscale scores, apart from on GAD [40].

The RCADS was developed as a revision of the SCAS, in order to correspond to dimensions of several DSM-IV anxiety disorders and also to include major depression. In particular, it was intended to refine the measurement of GAD to reflect core aspects of ‘worry’. Internal consistency was found to be good for subscales, and also for the shortened Anxiety 15 item version. In the original study [35] one week test-retest reliability ranged from $r = .65$ to .80. The total variance explained by the factor analysis was less than 50%; however, subsequent confirmatory factor analyses have reported good fit to the 6 factor solution [52;62;66] for the child scale, and acceptable for the parent scale [63;64]. Convergent and divergent validity have been shown convincingly, as has criterion validity with diagnoses based on standardised clinical psychiatric interview.

The MASC has well-established strengths in internal consistency (except for the subscale Harm Avoidance in [55]) and in test-retest reliability. The latter has been shown at 3 weeks [56;57] and
Table 3. Characteristics of study populations in articles on measurement properties.

| Measure | Article | Ref No. | Study Population/sample | Mean age (SD) years; range | N   | Male % (child) | Country |
|---------|---------|---------|--------------------------|----------------------------|-----|----------------|---------|
| ADIS    | Brown-Jacobsen et al (2011) | 40 | Anxiety diagnostic clinic | 12.32(3.3); 7–18 | 88  | 55.7           | USA     |
|         | Canavera et al (2009) | 41 | OCD | 12.17 (X); 6–17 | 93  | 46.2           | USA     |
|         | Comer & Kendal (2004) | 42 | Anxiety disorder clinic | 10.2 (1.8); 7.5–14 | 98  | 54            | USA     |
|         | Grills & Ollendick (2003) | 43 | Psychology clinic | 10.69 (2.66); X | 165 | 63.6           | USA     |
|         | Higa-McMillan (2008) | 44 | CBT centre | 12.3 (3.3); 6–18.2 | 289 | 68.2           | Hawaii  |
|         | Lyneham et al (2007) | 45 | Anxiety disorder clinic | 11.25 (2.58); 7–16 | 153 | 50.3           | Australia |
|         | Lyneham & Rapee (2005) | 46 | Anxiety clinic and community | 9.2 (1.9); 6–12 | 73  | 67            | Australia |
|         | Silverman et al (2001) | 47 | Psychosocial research centre | 10.15 (X); 7–16 | 62  | 39.6           | USA     |
|         | Storch et al (2012) | 48 | ASD | 10.34 (2.21); 7–17 | 85  | 76.5           | USA     |
|         | Wood et al (2002) | 49 | Anxiety disorder clinic | 11.71 (2.64); 8–17 | 186 | 53.8           | USA     |
| MASC    | Anderson et al (2009) | 50 | School | 14.5 (X); 13–17 | 372 | 45.7           | USA     |
|         | Baldwin & Dadds (2007) | 51 | Community | 11.36 (1.21); 9.16–14.42 | 452 | 63            | Australia |
|         | Brown et al (2012)* | 52 | Paediatric primary care clinic | 12.3 (2.7); 7–17 | 229 | 36.6           | USA     |
|         | Dierker et al (2001) | 53 | School | 14 (X); 13–15 | 632 | 45            | USA     |
|         | Grills-Taquechel et al (2008) | 54 | Mixed clinical diagnoses | 10.44 (2.64); 7–17 | 262 | 65           | USA     |
|         | Langer et al (2010) | 55 | Anxiety clinic | 11.61 (2.64); 6–17 | 174 | 54            | USA     |
|         | March et al (1997) | 56 | Clinical | 11.6 (X); 8–18 (M) | 24  | 75            | USA     |
|         |                     |                 |                          | 11.8 (X); 8–16 (F) |     |               |         |
|         | March et al (1999) | 57 | School | 13.98 (2.6); 8–18 | 142 | 35            | USA     |
|         | Ross et al (2007) | 58 | Asthma | 14.66 (1.75); 12–18 | 53  | 50.9           | Canada  |
|         | Rynn et al (2006) | 59 | Anxiety disorder | 12.2 (3.3); X | 116 | 57.8           | USA     |
|         | Thaler et al (2010) | 60 | Learning disability | 13.75 (1.81); 11–17 | 41  | 78.6           | USA     |
|         | White et al (2012) | 61 | HFASD | 14.58 (1.67); 12–17 | 30  | 76.67          | USA     |
|         | Wood et al (2002) | 62 | Outpatient anxiety clinic | 11.71 (2.64); 8–17 | 186 | 53.8           | USA     |
| RCADS   | Brown et al (2012)* | 52 | Pediatric primary care clinic sample | 12.13 (2.7); 7–17 | 229 | 36.6           | USA     |
|         | Chorpita et al (2000) | 35 | Schools | 12.87 (2.82); 6.17–18.92 | 1641 | 45.6          | Hawaii  |
|         | Chorpita et al (2005) | 62 | Clinical (CBT centre) | 12.9 (2.7); 7.5–17.9 | 513 | 67.4          | Hawaii  |
|         | Ebesutani et al (2010) | 63 | Clinical (CBT centre) | 11.5 (2.5); 6.55–18.97 | 490 | 67.8          | Hawaii  |
|         | Ebesutani et al (2011) | 64 | School | 13.3 (2.98); 8–18 | 967 | 45.1          | Hawaii  |
|         | Ebesutani et al (2012) | 65 | School | 10.3 (1.7); 7–14 | 1060 | 39.3          | Hawaii  |
|         |                     |                 |                          | Clinic | 303 | 67.3          |         |
|         | Trent et al (2012) | 66 | School | Grades 2 to 12 | 12 659 | 49          | USA     |
| RCMAS   | Cole et al (2000) | 67 | School | 11.9 (0.5); 10.7–13.5 | 562 (T1) | 49          | USA     |
|         |                     |                 |                          | 630 (T2) |     |               |         |
|         | Dadds et al (1998) | 68 | School | X (X); 7–14 | 1786 | 40.9          | Australia |
|         | Dierker et al (2001) | 53 | School | 14 (X); 13–15 | 632 | 45           | USA     |
|         | Kenny & Faust (1997) | 69 | Mental health centre | 10 (2.87); 5–16 | 54  | 64.9          | USA     |
|         | Nelson & Renzenbrink (1995) | 70 | Psychiatric hospital | 14.28 (1.57); 12–17 | 25  | 36           | USA     |
|         | Olatunji & Cole (2009) | 71 | School | 8.96 (0.61); 8–12 | 787 | 48           | USA     |
| Measure | Article | Ref No. | Study Population/sample | Mean age (SD) years; range | N   | Male % (child) | Country |
|---------|---------|---------|-------------------------|-----------------------------|-----|----------------|---------|
|        | Paget & Reynolds (1984) | 72 | Learning disability | 12.8 (1.2); 6–17 | 106 | 68.9 | USA |
|        | Perrin & Last (1992) | 73 | Clinic and community | X (X); 5–17 | 213 | 100 | USA |
|        | Pina et al (2009) | 74 | Anxiety clinic | 10.21 (2.78); 6–16 | 677 | 52.7 | USA |
|        | Reynolds & Paget (1981) | 75 | School | X (X); 6–19 | 4972 | 50.2 | USA |
|        | Reynolds & Richmond (1987) | 30 | School | X (X); 6–18 (group 1) | 329 | X | USA |
|        | White & Farrell (2001)** | 76 | School | 11.6 (0.6); 10.8–14.1 | 898 | X | USA |
|        | Wisniewski et al (1987) | 77 | School | 12.10(0.92); 11–14 | 161 | 45 | USA |
|        | Gonzalez et al (2012)** | 81 | Drawn from families seeking outpatient mental health services | 13.96 (1.89); 11–18 | 374 | C | USA |
|        | Bodden et al (2009) | 80 | Clinically anxious | 12.5 (2.7); 8–18 | 138 | 40 | Netherlands |
|        | Gen pop control | 12.4(2.6) | 8–18 | 38 | 37 | |
|        | Gonzalez et al (2012)** | 81 | Drawn from families seeking outpatient mental health services | 13.96 (1.89); 11–18 | 374 | C | USA |
|        | Jastrowski et al (2012) | 82 | Pediatric chronic pain clinic | 14.21 (2.54); 8–18 | 349 | 31 | USA |
|        | Monga et al (2000) | 83 | Mood/anxiety disorders outpatient clinic | 14.4 (2.3); 9–18.9 | 295 | 43 | USA |
|        | Muris and Steerneman (2001) | 84 | Anxiety disorder or disruptive disorder | 12.8 (2.4); 8–17 | 48 | 54.2 | Netherlands |
|        | Muris et al (1999) | 85 | School | 12.2 (0.5); 11–14 | 101 | 47 | Netherlands |
|        | | | 9.6 (1.1); 8–12 | 71 | 50.1 | |
|        | | | 10 (1.2); 8–12 | 88 | 61.2 | |
|        | Muris et al (2004) | 86 | Anxiety outpatient | 12 (2.9); 7–17 | 242 | 57.9 | Netherlands |
|        | Muris & Mayer et al (2001) | 87 | Clinically anxious | 9.9 (1.4); 8–13 | 36 | 25 | Netherlands |
|        | Simon & Bogels (2009) | 88 | School | High anxious | 9.92 (1.23); 8–13 | 188 | 45 | Netherlands |
|        | | | Median anxious | 10.22 (1.13); 8–13 | 82 | 54 | |
|        | Van Steensel (2012) | 89 | ASD and anxiety disorder | 11.37 (X); 7–18 | 115 | 78.3 | Netherlands |
|        | | | Anxiety disorder | 12.79; 7–18 | 122 | 50.8 | |
|        | Wren et al (2007) | 90 | Pediatric Primary care | 10.5 (1.4); 8–13 | 515 | 49.1 | USA |
|        | Brown-Jacobsen et al (2011) | 40 | Clinically anxious | 12.32 (3.3);7–18 | 88 | 55.7 | USA |
|        | Essua et al (2011)*** | 91 | General population | 14.44 (1.7); 12–17 | 469 | 46.5 | England |
|        | Nauta et al (2004) | 92 | Clinically anxious; general population | 10.8 (2.4); 6–18 | 745 | 52 | Australia; Netherlands |
|        | Russell & Sofronoff (2005) | 93 | AS; clinically anxious and normative data | X (X); 10–13 | 65 (AS) | 84.6 | Australia |
|        | Spence (1998) | 33 | General population; clinical (separation and social phobia) | 10.11 (1.25); 8–12 | 2052 | 42 | Australia |
3 months [56]; indeed [60] have shown stability estimates around \( r = 0.50 \) for child-report and higher (.56 to .70) for parent-report in a community sample. The original factor analysis [56] for the MASC explained only 39.4\% of the total variance; however, subsequent confirmatory factor analyses have found good fit for the four factor solution [51;52;54] though one study [39] found only acceptable evidence of fit (e.g. RMSEA = 0.73) [98]. Correlations with other measures of anxiety are high, with discriminant validity established (usually by lower agreement with scales measuring depression). Findings for criterion validity have been variable, showing high levels of agreement with diagnostic groupings (diagnostic interview or ADIS as ‘gold standard’) in some community [50;54] and clinical studies [59;60;49– except for generalised anxiety disorder] though one study [59] found only acceptable evidence of fit (e.g. RMSEA = 0.73) [98].

| Measure | Article | Ref No. | Study Population/sample | Mean age (SD) years; range | N | Male % (child) | Country |
|---------|---------|---------|-------------------------|---------------------------|---|---------------|---------|
| Spence (2003) | 94 | General population | 13.51 (0.51); 13–14 | 875 | 54 | Australia |
| Whiteside & Brown (2008) | 95 | Community | 12.74 (2.35); 9–18 | 85 | 55 | USA |
| Whiteside et al (2012) | 96 | Community | 10.42 (1.5); 8–13 | 420 | 49 | USA |
| Anxiety disorders | 93 | AS; clinically anxious and normative data | X (X); 10–13 | 65 (AS) | 84.6 | Australia |
| SWQ | Bailey et al (2006) | 78 | Primary care | 14.5 (1.27); 7–18 | 196 | 52 | USA |
| Russell & Sofronoff (2005) | 93 | OCD | 12.2 (2.5); X | 71 | 36.6 | USA |

ADIS: Anxiety Disorders Interview Schedule. MASC: Manifest Anxiety Scale for Children. RCADS: Revised Child Anxiety and Depression Scale. RCMAS: Revised Child Manifest Anxiety Scale. SCARED: Screen for Child Anxiety and Related Emotional Disorders. SCAS: Spence Children’s Anxiety Scale. SWQ: Social Worries Questionnaire. CGI: Clinical Global Impressions.

X = No details. M = male; F = female. C: child; P: parent. AS: Asperger Syndrome. ASD: autism spectrum disorder. CBT: cognitive behaviour therapy. HFASD: high-functioning ASD. OCD: obsessive-compulsive disorder. T1 = Time 1; T2 = Time 2.

*looked at cross ethnic validity (all participants African American); ** looked at cross ethnic validity (all participants African American or non Hispanic white); *** Data collected from 5 European countries – details are for UK sample.

doi:10.1371/journal.pone.0085268.t003

The two studies of criterion validity against diagnostic interview produced conflicting results; the clinic study supported criterion validity [73] but the community study concluded that the RCMAS was less successful than the MASC in identifying anxiety and depression [53].

There are a number of versions of the SCARED. The original 38 and 41 item tools have good content validity being derived from DSM [31], some evidence of test-retest reliability for total and subscale scores on both parent and child versions [31], plus consistently good internal reliability. Good structural validity was found [82] though evidence for measurement invariance was not as strong (RMSEA > .06) [81]. Criterion validity was good [78;81]; clinically anxious children scored significantly higher on the child SCARED than non-anxious, depressed and disruptive groups on total and subscale scores [79;83], and by examining area under the curve (AUC) against clinical interview [78;83].

The SCARED-Revised is a 66 item measure with nine subscales. Internal consistency was found to be good though the quality of the articles varied. The total scores and most of the subscales had good internal consistency, except OCD (parent and child versions), blood/injection/injury (child) and environmental/neurological (situational (parent) [85] and specific phobias [84]. Test-retest reliability of the child total score was positive \( r > .30 \) within the subscales approaching this level apart from GAD, separation, OCD and traumatic stress \( r < .70 \) [85]. Correlations across time with the State Trait Anxiety Inventory for Children demonstrated responsiveness to change though the quality of the evidence was limited [87]. Significantly higher SCARED-R scores were predictive of those with anxiety disorders, demonstrating criterion validity, though the GAD, specific phobias and separation anxiety subscales performed less well in the child version [86]. Correlations between parent and child were mixed with both high [86] and low [85] agreement found.

The SCARED-71 is a version adding five further social phobia items to the SCARED-R. Internal consistency was positive in parent and child versions for total and all but one subscale scores (OCD, child report) [89]. Criterion validity in terms of predictability of diagnosis by corresponding subscale was good except for...
| Measure | Article                        | Ref No. | Version | Internal Consistency | Reliability | Content Validity | Structural Validity | Hypothesis Testing | Criterion Validity | Responsiveness |
|---------|-------------------------------|---------|---------|----------------------|-------------|------------------|---------------------|--------------------|-------------------|-----------------|
| ADIS    | Brown-Jacobsen et al (2011)   | 40      | C + P   | -                    | -           | -                | -                   | Good               | -                 | -               |
|         | Canavera et al (2009)         | 41      | C + P   | -                    | -           | -                | -                   | Good               | -                 | -               |
|         | Comer & Kendal (2004)         | 42      | C + P   | Good                 | -           | -                | -                   | Good               | -                 | -               |
|         | Grills & Ollendick (2003)     | 43      | C + P   | Good                 | -           | -                | -                   | Good               | -                 | -               |
|         | Higa-McMillan (2008)          | 44      | C       | -                    | -           | -                | -                   | Good               | -                 | -               |
|         | Lyneham et al (2007)          | 45      | C + P   | -                    | Good        | -                | -                   | -                  | -                 | -               |
|         | Lyneham & Rapee (2005)        | 46      | C + P   | -                    | Good        | -                | -                   | -                  | -                 | -               |
|         | Silverman et al (2001)        | 47      | C + P   | -                    | Good        | -                | -                   | Good               | -                 | -               |
|         | Storch et al (2012)           | 48      | C + P   | -                    | -           | -                | -                   | Good               | -                 | -               |
|         | Wood et al (2002)             | 49      | C + P   | -                    | -           | -                | -                   | Good               | -                 | -               |
| MASC    | Anderson et al (2009)*        | 50      | C       | good                 | -           | -                | -                   | excellent          | excellent         | -               |
|         | Baldwin & Dadds (2007)        | 51      | C + P   | excellent            | -           | -                | -                   | excellent          | good              | -               |
|         | Brown et al (2012)            | 52      | C       | good                 | -           | -                | -                   | good               | -                 | -               |
|         | Dierker et al (2001)          | 53      | C       | fair                 | -           | -                | -                   | -                  | good              | -               |
|         | Grills-Taquechel et al (2008) | 54      | C       | excellent            | -           | -                | -                   | good               | -                 | excellent        |
|         | Langer et al (2010)           | 55      | C + P   | fair                 | -           | -                | -                   | -                  | good              | -               |
|         | March et al (1997)            | 56      | C + P   | excellent            | poor α      | -                | -                   | excellent          | poor α            | -               |
|         | March & Sullivan (1999)       | 57      | C       | -                    | -           | -                | -                   | -                  | -                 | -               |
|         | Ross et al (2007)             | 58      | C       | -                    | -           | -                | -                   | -                  | -                 | -               |
|         | Rynn et al (2006)             | 59      | C       | good                 | -           | -                | -                   | good               | -                 | excellent        |
|         | Thaler et al (2010)           | 60      | C + P   | fair                 | -           | -                | -                   | fair               | -                 | fair            |
|         | White et al (2012)            | 61      | C + P   | fair                 | -           | -                | -                   | fair               | -                 | -               |
|         | Wood et al (2002)             | 62      | C + P   | -                    | -           | -                | -                   | -                  | excellent         | -               |
| RCADS   | Brown et al (2012)            | 63      | C       | -                    | -           | -                | -                   | -                  | -                 | -               |
|         | Chorpita et al (2000)         | 64      | C       | -                    | -           | -                | -                   | -                  | -                 | -               |
|         | Chorpita et al (2005)         | 65      | C       | -                    | -           | -                | -                   | -                  | -                 | -               |
|         | Ebesutani et al (2010)        | 66      | P       | excellent            | -           | -                | -                   | excellent          | good              | -               |
|         | Ebesutani et al (2011)        | 67      | P       | excellent            | -           | -                | -                   | excellent          | excellent         | -               |
|         | Ebesutani et al (2012)        | 68      | C Short  | excellent            | -           | -                | -                   | -                  | -                 | -               |
| Measure | Article | Ref No. | Version | Internal Consistency | Reliability | Content Validity | Structural Validity | Hypothesis Testing | Criterion Validity | Responsiveness |
|---------|---------|---------|---------|----------------------|-------------|------------------|---------------------|---------------------|-------------------|-----------------|
|         | Trent et al (2012) | 66 | C | excellent | - | - | excellent | - | - |
| RCMAS   | Cole et al (2000)** | 67 | C + P | poor β | - | - | excellent | - | - |
|         | Dadds et al (1998)*** | 68 | C | good | - | - | good | - | - |
|         | Dierker et al (2001) | 53 | C | fair | - | - | good | good | - |
|         | Kenny & Faust (1997) | 69 | C | - | - | - | good | - | - |
|         | Nelson & Renzenbrink (1995) | 70 | C | poor | - | - | poor x | - | - |
|         | Olatunji & Cole** (2009) | 71 | C | good | excellent | - | excellent | - | - |
| SCARED  | Paget & Reynolds (1984) | 72 | C | poor β | - | - | poor x | - | - |
|         | Perrin & Last (1992) | 73 | C | - | - | - | - | fair |
|         | Pina et al (2009) | 74 | C | - | - | - | excellent | fair | - |
|         | Reynolds & Paget (1981) | 75 | C | - | - | - | excellent | - | - |
|         | Reynolds & Richmond (1997) | 30 | C | poor | - | fair | - | fair | - |
|         | White & Farrell (2001) | 76 | C | - | - | excellent | excellent | - | - |
|         | Wisniewski et al (1987) | 77 | C | - | - | good | - | - |
| SCARED  | Bailey et al (2006)* | 78 | C + P | - | - | - | - | excellent | - |
|         | Birmaher et al (1997) | 31 | C + P | excellent | fair | good | good | fair | excellent | - |
|         | Birmaher et al (1999) | 79 | C + P | poor x | - | - | poor x | fair | fair | - |
|         | Bodden et al (2009) | 80 | C + P | good SCARED-71 | - | - | - | excellent | - |
|         | Gonzalez et al (2012) | 81 | C + P | good | - | - | excellent | - | excellent (P) |
|         | Jastrowski et al (2012) | 82 | C + P | excellent | - | - | excellent | good | - |
|         | Monga et al (2000) | 83 | C + P | - | - | - | - | excellent | excellent | - |
|         | Muris et al (1999) | 85 | C + P | good SCARED-R | good | - | - | good | - |
|         | Muris et al (2004) | 86 | C + P | poor SCARED-R | - | - | - | fair | excellent | - |
|         | Muris & Mayer et al (2001) | 87 | C SCARED-71 R | - | - | - | fair | fair | - |
|         | Muris and Steermeman (2001) | 84 | C SCARED-71 R | - | - | - | fair | fair | - |
|         | Simon & Bogels (2009) | 88 | C SCARED-71 | - | - | - | - | excellent | - |
|         | Van Steensel (2012) | 89 | C + P | good SCARED-71 | - | - | - | good | excellent | - |
|         | Wren et al (2007) | 90 | C + P | - | - | - | excellent | good | - |
| SCAS    | Brown-Jacobsen et al (2011) | 40 | C + P | good | - | - | good | good | - |
though judgements tended to converge by 14 weeks of treatment
children than by therapists and the independent evaluator,
improvement was reported significantly sooner by parents and
parent though most of the correlations were
therapist-parent, therapist-child, and independent evaluator-
anxiety differed [93].

good evidence of criterion validity with agreement for social
self-report; P = parent report. R = Revised.
*looked at particular subscale (Social Anxiety/Social Phobia); ** created continuous data by altering response format; *** looked at particular subscale (total score and
- - - - good - -
Agreement for social anxiety differed. An additional property
described was that of content validity. The parent version of
the Social Worries Questionnaire has demonstrated that children
with Asperger syndrome were significantly more anxious than
typically developing children, on a par with a clinically anxious sample. As
predicted by Russell and Sofronoff, parent and child reports of
anxiety differed [93].
The CGI-I showed inter-rater agreement for parent-child,
therapist-parent, therapist-child, and independent evaluator-
parent though most of the correlations were <.70. Across time,
improvement was reported significantly sooner by parents and
children than by therapists and the independent evaluator,
though judgements tended to converge by 14 weeks of treatment
for OCD.

Discussion
Principal Findings
In this systematic review, eight tools were found which had been
used to measure primary outcomes in anxiety intervention trials
for children with high-functioning ASD in middle childhood. A
second systematic search of literature found sixty-three articles
studying children and examining the measurement properties of
the eight tools.

There was limited or no evidence for three of the eight
properties of measurement tools rated in this review using the
COSMIN checklist: measurement error, content validity and
responsiveness to change. In terms of the primary purpose of the
review – to inform the choice of tool to measure outcomes of
intervention trials for anxiety in children with ASD – these are
serious limitations in the evidence.

Only four articles included children with ASD, and none of
these considered content validity. Indeed, the field is hampered by
lack of a definitive conceptualisation of anxiety in ASD, and the
means to capture features of anxiety as a clinical disorder separate
from ASD [15-99–102]. Anxiety interacts with core symptoms
(such as poor social skills and repetitive thoughts) and so differs in
several ways from anxiety seen in typically developing children.
For example, a child with ASD who is reluctant to go to school is
more likely to be experiencing social anxiety rather than
separation anxiety. However, until basic psychometric work
including content analysis is carried out, outcome measures
developed with typically developing children will continue to be
utilised with children with ASD [100;101].

The lack of evidence about responsiveness to change of tools is
also a limitation for the purpose of the review. The CGI
Improvement rating explicitly focuses on change, and was
utilised by three of the ASD intervention studies, indicating
treatment effects. It has been used widely in autism medication
trials [6], has comparable effect sizes to other rating scales in
adult anxiety intervention trials [103], and has the advantage
that it can be rated blind to group and time point. Therefore it is
likely to continue to be used in intervention trials for children,
even though evidence for its measurement properties was sparse
in this review.

One further property included in the COSMIN checklist was
not included in the review, cross-cultural validity. However
support for the measurement properties of the SCARED across

Table 4. Cont.

| Measure | Article | Ref No. | Version | Internal Consistency | Reliability | Content Validity | Structural Validity | Hypothesis Testing | Criterion Validity | Responsiveness |
|---------|---------|--------|---------|----------------------|-------------|------------------|---------------------|-------------------|------------------|----------------|
| LGI     | Lewin et al (2012) | 97     | Improvement | - | good | - | - | - | - |
| SWQ     | Bailey et al (2006) | 78     | P | - | - | - | - | - | excellent |
|         | Russell & Sofronoff (2005) | 93     | C + P | - | - | - | - | - | good |

ADIS: Anxiety Disorders Interview Schedule. MASC: Manifest Anxiety Scale for Children. RCADS: Revised Child Anxiety and Depression Scale. RCMAS: Revised Child
Manifest Anxiety Scale. SCARED: Screen for Child Anxiety and Related Emotional Disorders. SCAS: Spence Children’s Anxiety Scale. SWQ: Social Worries Questionnaire.
CGI: Clinical Global Impressions.

Measurement Error was not evaluated in any article; Cross cultural validity was not included in the review.
C = child self report; P = parent report. R = Revised.
*looked at particular subscale (Social Anxiety/Social Phobia); ** created continuous data by altering response format; *** looked at particular subscale (total score and
- - - - good - -
Agreement for social anxiety differed. An additional property
described was that of content validity. The parent version of
the Social Worries Questionnaire has demonstrated that children
with Asperger syndrome were significantly more anxious than
typically developing children, on a par with a clinically anxious sample. As
predicted by Russell and Sofronoff, parent and child reports of
anxiety differed [93].
The CGI-I showed inter-rater agreement for parent-child,
therapist-parent, therapist-child, and independent evaluator-
parent though most of the correlations were <.70. Across time,
improvement was reported significantly sooner by parents and
children than by therapists and the independent evaluator,
though judgements tended to converge by 14 weeks of treatment
for OCD.

Discussion
Principal Findings
In this systematic review, eight tools were found which had been
used to measure primary outcomes in anxiety intervention trials
for children with high-functioning ASD in middle childhood. A
second systematic search of literature found sixty-three articles
studying children and examining the measurement properties of
the eight tools.

There was limited or no evidence for three of the eight
properties of measurement tools rated in this review using the
COSMIN checklist: measurement error, content validity and
responsiveness to change. In terms of the primary purpose of the
review – to inform the choice of tool to measure outcomes of
intervention trials for anxiety in children with ASD – these are
serious limitations in the evidence.

Only four articles included children with ASD, and none of
these considered content validity. Indeed, the field is hampered by
lack of a definitive conceptualisation of anxiety in ASD, and the
means to capture features of anxiety as a clinical disorder separate
from ASD [15-99–102]. Anxiety interacts with core symptoms
(such as poor social skills and repetitive thoughts) and so differs in
several ways from anxiety seen in typically developing children.
For example, a child with ASD who is reluctant to go to school is
more likely to be experiencing social anxiety rather than
separation anxiety. However, until basic psychometric work
including content analysis is carried out, outcome measures
developed with typically developing children will continue to be
utilised with children with ASD [100;101].

The lack of evidence about responsiveness to change of tools is
also a limitation for the purpose of the review. The CGI
Improvement rating explicitly focuses on change, and was
utilised by three of the ASD intervention studies, indicating
treatment effects. It has been used widely in autism medication
trials [6], has comparable effect sizes to other rating scales in
adult anxiety intervention trials [103], and has the advantage
that it can be rated blind to group and time point. Therefore it is
likely to continue to be used in intervention trials for children,
even though evidence for its measurement properties was sparse
in this review.

One further property included in the COSMIN checklist was
not included in the review, cross-cultural validity. However
support for the measurement properties of the SCARED across

Table 5. Quality of measurement properties of tools.

| Tool and Version | Measurement properties | Interpretability | Differences in scores between subgroups |
|------------------|------------------------|------------------|----------------------------------------|
|                  | Internal Consistency   | Reliability      | Content Validity | Structural Validity | Hypothesis Testing | Criterion Validity | Responsiveness | Par/ch ** | Conv/div *** |                      |
| Anxiety Disorders Interview Schedule (ADIS) | C | na | +++ * | na | ++ | – – | – | – | Na | Na | na | na |
| Multidimensional Anxiety Scale for Children (MASC) | P | na | +++ * | na | na | – – | – | – | Na | Na | na | na |
| Revised Children's Anxiety and Depression Scale (RCADS) | C | +++ | – – | ++ | +/- | na | ++ | +++ | na | Y |
| Revised Children's Manifest Anxiety Scale (RCMAS) | P | +++ | ++ | na | – – | na | +++ | +++ | na | Y |
|                  | C Short | +++ | Na | na | +++ | na | na | +++ | na | Y |
| Screen for Child Anxiety Related Emotional Disorders (SCARED) | P | ? | Na | na | ? | na | na | Na | na | na | na | Y |
| Social Worries Questionnaire (SWQ) | C | +++ | + | ++ | +/- | – – | na | +++ | na | Y |
| Social Worries Questionnaire (SWQ) | P | +++ | + | ++ | +/- | – – | na | +++ | na | Y |
| Social Worries Questionnaire (SWQ) | C - Rev | ++ | ++ | na | na | +/- | ++ | ++ | na | Y |
| Social Worries Questionnaire (SWQ) | P - Rev | ++ | na | na | na | +/- | ? | ++ | na | Y |
| Social Worries Questionnaire (SWQ) | C - 71 | +++ | na | na | na | ++ | ++ | +/- | na | Y |
| Social Worries Questionnaire (SWQ) | P - 71 | +++ | na | na | na | ++ | – – | +/- | na | Y |
| Spence Children's Anxiety Scale (SCAS) | C | +++ | – – | na | +++ | +++ | +++ | +++ | na | Y |
| Spence Children's Anxiety Scale (SCAS) | P | +++ | – – | na | ++ | +++ | +++ | +++ | na | Y |
several countries and cultures has been found in a meta analysis [104] and by Gonzalez and colleagues [81].

Overall the findings of the review suggest that the tools which are most robust in their measurement properties are the Spence Children’s Anxiety Scale, its revised version – the Revised Children’s Anxiety and Depression Scale, and also the Screen for Child Anxiety Related Emotional Disorders. The weakness of the measurement of GAD by the SCAS appears to have been improved in the RCADS. However, self-report and parent report generally have the limitation in RCTs that they are not ‘blinded’. In the four ASD intervention trials which used ADIS, participants were asked not to unblind the researcher as they described current events and behaviours in the clinical interview. Thus a combination of ways of measuring anxiety (feelings and behaviours) appears to be necessary to achieve robust measurement.

Clinical Implications

The review found a mixed picture in terms of the level of correlation between parent and child report. Agreement is not necessarily to be expected, with each individual reflecting different symptoms captured (for example, more observable behaviours being identified by the parent), and the possible influence of factors such as the parent’s own experiences affecting sensitivity to the child’s symptoms [105]. While the level of agreement between parents and their children with ASD may actually be higher than observed for typically developing groups [106], a number of researchers [e.g. 26;61;93] comment that children with high functioning ASD are likely to under-report anxiety symptoms, one reason being difficulty in identifying their own (and others’) emotions. Therefore a combination of perspectives is likely to give a more rounded picture.

One further issue for the measurement of outcomes in intervention trials and clinical practice in ASD is a need to consider further what constitutes a successful outcome [13]. Necessarily, the tools reviewed here as primary outcomes focus on clinical symptoms; however, the goals of intervention are likely to include broader constructs such as participation and quality of life. The International Classification of Functioning, Disability and Health paradigm [107], which is the World Health Organisation recommended conceptual model for measuring health and disability and evaluating interventions, emphasizes that body functions, activity and participation may all be important indicators of intervention success. Children with ASD may have hypersensitivity to visual and auditory stimuli, which in turn may result in activity limitations (e.g. social anxiety) and restricted social participation (e.g. reluctance to go to new places). Effective interventions for anxiety would also expect to see change in socially valid outcomes for children such as new experiences and greater success in friendships, whatever the nature of the baseline anxiety.

Limitations

This systematic review had some limitations. Articles were accessed only in English as we lacked resources for translation. Data extraction was done only in part by two independent reviewers. Although the COSMIN manual and checklist is validated and well structured, there is still an element of subjectivity in the review process such that different decisions regarding ratings and synthesis might be made by other reviewers.

The focus was on children in middle childhood who are high-functioning, and anxiety measurement issues in other age and ability groups have not been considered. Nevertheless, children
with ASD are reported to be vulnerable to high anxiety across ages [8;108] and abilities [8;108;109] and so intervention and measurement issues require wider examination.

Conclusions

Though there appears to be a certain international practice concept developing in research groups undertaking trials of intervention for anxiety in children with ASD, the evidence for the measurement properties of the chosen tools is patchy. The review has allowed some conclusions to be drawn on what may be the psychometrically sound assessment tools. However, there requirements to be further consideration of how to achieve blinded outcome measurement in RCTs, and how to judge the appropriateness of tools developed to measure anxiety in typically developing children when applied with children who have ASD.

References

1. Reichow B, Wolery M (2009) Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA young autism project model. J Autism Dev Disord 39: 23–41.
2. Osipova MB, Seida JK, Clark B, Karkhanine M, Hartling L, et al (2008) Behavioral and developmental interventions for autism spectrum disorder: a clinical systematic review. PLoS ONE 3(11): e37535.
3. Seligman LD, Ollendick TH, Langley AK, Baldacci HB (2004) The utility of measures of child and adolescent anxiety: a meta-analytic review of the Revised Children’s Manifest Anxiety Scale, the State–Trait Anxiety Inventory for Children, and the Child Behavior Checklist. J Clin Child Psychol 33(3): 357–365.
4. Diggle T, McConachie H (2002) Parent-mediated early intervention for young children with autism spectrum disorder. Cochrane Database Syst Rev, Issue 2. Art. No.: CD003496. DOI: 10.1002/14651856.CD003496.
5. Williams K, Wheeler DM, Silove N, Hazell P (2010) Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). Cochrane Database Syst Rev, Issue 8. Art. No.: CD004677. DOI: 10.1002/14651858.CD004677.pub2.
6. Arnold LE, Aman MG, Martin A, Collier-Crespin A, Vitiello B, et al (2000) Assessment in multisite randomized clinical trials of patients with autistic disorder: the Autism RUPP Network. J Autism Dev Disord 30: 99–111.
7. Wolery M, Garfinke AN (2002) Measures in intervention research with young children who have autism. J Autism Dev Disord 32: 436–478.
8. van Steenfel EFA, Bogs SM, Perrin S (2011) Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis. Clin Child Fam Psych 14: 302–317.
9. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, et al (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the special needs and autism project (SNAP). Lancet 368(9531): 210–215.
10. Increasing Access to Psychological Therapies http://www.iapt.nhs.uk/cypiapt/routine-outcome-monitoring-as-part-of-iapt/Accesed 2013 August 26.
11. Simka L, Jones J, Smith RL, Williamsen PR (2008) A systematic review of studies that aim to determine which outcomes to measure in clinical trials in children. PLoS One 5(4): 0569–0578.
12. Honey E, Rodgers J, McConachie H (2012) Measurement of restricted and repetitive behaviour in children with autism spectrum disorder: selecting a questionnaire or interview. Res Autism Spectr Disord 6: 757–776.
13. Matson J (2007) Determining treatment outcomes in early intervention programs for ASD: a critical analysis of measurement issues in learning based interventions. Res Dev Disabil 28: 207–218.
14. Arnold LE, Violiello B, McDougle C, Scibill S, Shah B, et al (2005) Parent-defined target symptoms respond to Risperidone in RUPP Autism study: customer approach to clinical trials. J Am Acad Child Adolesc Psychiatry 44(2): 1462–1450.
15. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, et al (2008) Psychiatric disorders in children with Autism Spectrum Disorders: prevalence, comorbidity, and associated factors in a population-derived sample. J Am Acad Child Adolesc Psychiatry 47(8): 921–929.
16. World Health Organisation (1992) International Classification of Diseases (ICD 10). Geneva: World Health Organization.
17. American Psychiatric Association (2013) Diagnostic and Statistical Manual of Mental Disorders (5th edition) Washington, DC.
18. Reaven J, Blakeley-Smith A, Calhane-Sheburne K, Hephurn S (2011) Group cognitive behavior therapy for children with high-functioning autism spectrum disorders and anxiety: a randomized trial. J Child Psychol Psychiatry 53(4): 410–9.
19. Wood J, Drahotha A, Sze K, Har K, Chiu A, et al (2009) Cognitive behavioral therapy for anxiety in children with autism spectrum disorders: a randomized, controlled trial. J Child Psychol Psychiatry 50(3): 224–234.
20. Wood J, Drahotha A, Sze K, Van Dyke M, Decker K, et al (2009) Brief report: Effects of cognitive behavioral therapy on parent-reported autism symptoms in school-age children with high-functioning autism. J Autism Dev Disord 39(11): 1608–1612.
21. McCully Leek RE, Lincoln AJ, Brown MZ, Chavira DA (2012) The Coping Cat Program for children with anxiety and ASD: A pilot randomized controlled trial. J Autism Dev Disord DOI 10.1007/s10803-012-1541-9.
22. McConachie H, Malcaughlin E, Grahame V, Taylor H, Honey E, et al (2013) Group therapy for anxiety in children with autism spectrum disorder. Autism DOI 10.1177/1362361314538999.
23. Sung M, Ooi VP, Goh TJ, Pathy P, Fung DSS, et al (2011) Effects of Cognitive-Behavioral Therapy on anxiety in children with Autism Spectrum Disorders: a randomized controlled trial. Child Psychiatry Hum Dev 42: 634–649.
24. Chalpant AM, Rapee R, Carroll L (2006) Treating anxiety disorders in children with high functioning autism spectrum disorders: a controlled trial. J Autism Dev Disord DOI 10.1007/s10803-006-0318-4.
25. Ooi VP, Lam CM, Sung M, Tan WTS, Goh TJ, et al (2008) Effects of cognitive-behavioural therapy on children with high functioning autistic spectrum disorders. Singapore Med J 49(3): 215–220.
26. Sofronoff K, Atwood T, Hinton S (2003) A randomised controlled trial of a CBT intervention for anxiety in children with Asperger syndrome. J Child Psychol Psychiatry 46(1): 1152–60.
27. Reaven JA, Blakeley-Smith A, Nicholls S, Dasari M, Flanagan E, et al (2009) Cognitive-behavioural group treatment for anxiety symptoms in children with high functioning autism spectrum disorders: a pilot study. Focus Autism Other Dev Disabil 24(1): 27–37.
28. Silverman WK, Albano AM (1996) The Anxiety Disorders Interview Schedule for Children: DSM-IV. Child and Parent versions. San Antonio TX: Psychological Corporation.
29. March JS (1998) Multimodal Anxiety Scale for Children. North Tonawanda, NY: Multi-Health Systems.
30. Reynolds CR, Richmond BO (1997) What I think and feel: a revised measure of Children’s Manifest Anxiety. J Am Acad Child Adolesc Psychiatry 36(4): 545–55.
31. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, et al (1997) The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. J Am Acad Child Adolesc Psychiatry 36(4): 545–55.
32. Spence SH (1995) The Social Worries Questionnaire. Social skills training: Enhancing social competence with children and adolescents. Windsor: NFER-Nelson.
33. Sprue SB (1998) A measure of anxiety symptoms among children. Behav Res Ther 36: 545–566.
34. Guy W (1976) Clinical global impressions. In W. Guy (Ed.), ECDEU Assessment Manual for Psychopharmacology, Revised. (218–222.). Rockville, MD.: National Institute of Mental Health.
35. Chopita BF, Yim L, Mofitt CE, Unemoto LA, Francis SE (2000) Assessment of symptoms of DSMIV anxiety and depression in children: A Revised Child Anxiety and Depression Scale. Behav Res Ther 38: 835–855.
36. Terwee CB, Janzen EP, Repjaben H, de Vet HWC (2009) Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 18: 1113–1123.
37. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, et al (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 19: 539–549.
38. De Vet HWC, Terwee CB, Mokkink LB, Knol DL (2011) Measurement in medicine. A practical guide. Cambridge: Cambridge University Press.
39. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, et al (2012) COSMIN checklist manual. VU University Medical Center, Department of Epidemiology and Biostatistics. (www.cosmin.nl).

Supporting Information

Checklist S1 PRISMA checklist. (DOC)

Acknowledgments

We are grateful to Dr Caroline Terwee for training in the COSMIN methodology, and to Denise Brett for formatting references.

Author Contributions

Conceived and designed the experiments: SW HM C. Performed the experiments: SW HM C. Analyzed the data: SW HM C. Contributed reagents/materials/analysis tools: SW HM C. Wrote the paper: SW HM C.
64. Ebesutani C, Chorpita BF, Higa-McMillan CK, Nakamura BJ, Regan J, et al (2010) A
58. Ross CJM, Davis TMA, Hogg DY (2007) Screening and assessing adolescent
44. Higa-McMillan CK, Smith RL, Chorpita BF, Hayashi K (2008) Common and
29. Perrin S, Last C (1992) Do childhood anxiety measures measure anxiety?
40. Brown-Jacobson AM, Wallace DP, Whiteside SPH (2011) Multi-method, multi-
20. Paget KD, Reynolds CR (1984) Dimensions, levels and reliabilities on the Revised
30. Reynolds CR, Paget KD (1981) Factor analysis of the revised Children’s
10. Muris P, Steerneman P (2001) The revised version of the Screen for Child Anxiety
67. Cole DA, Hoffman K, Tram JM, Maxwell SE (2000) Structural differences in
7. Ollendick TH (2011) Psychometric properties of the Spence Child Anxiety
17. Muris P, Steerneman P (2001) The revised version of the Screen for Child Anxiety
26. Muris P, Steerneman P (2001) The revised version of the Screen for Child Anxiety
22. Muris P, Steerneman P (2001) The revised version of the Screen for Child Anxiety
60. Muris P, Dreessen L, Bo¨gels S, Weckx M, van Melick M (2004) A questionnaire
25. Depress Anxiety 22(4): 292–299.
79. Wisneski JJ, Mulick Glenshaft JL, Coury DL (1987) Test retest reliability of the
56. Astoria CA, Carelli DA, Klein ME, Klein MB (2006) Brief measures to screen for
84. Muris P, Steerneman P (2001) The revised version of the Screen for Child Anxiety
78. Wisneski JJ, Mulick Glenshaft JL, Coury DL (1987) Test retest reliability of the
83. Perrin S, Last C, Brookman RR, Whyatt HM, Decker BW, et al (2008) Agreement in
55. Baldwin JS, Dadds MR (2007) Reliability and validity of parent and child versions of
45. Baldwin JS, Dadds MR (2007) Reliability and validity of parent and child versions of
59. Baldwin JS, Dadds MR (2007) Reliability and validity of parent and child versions of
72. Paget KD, Reynolds CR (1984) Factor analysis of the revised Children’s Manifest Anxiety Scale for blacks, whites, males, and females with a national normative sample. J Consult Clin Psychol 49(3): 352–9.
16. White SW, Schry AR, Maddox BB (2012) Brief Report: The assessment of anxiety in high-functioning adolescents with autism spectrum disorder. J Autism Dev Disord 42: 1138–1145.
2. Chorpita BF, Moffin CE, Gray J (2005) Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. Behav Res Ther 43: 809–322.
33. Perrin S, Last C (1992) Do childhood anxiety measures measure anxiety? J Abnorm Child Psychol 20(4): 367–378.
21. Reynolds CR, Paget KD (1981) The revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. J Am Acad Child Adolesc Psychiatry 40(3): 1203–6.
23. Perrin S, Last C. (1992) Do childhood anxiety measures measure anxiety? J Abnorm Child Psychol 20(4): 367–378.
4. Perrin S, Last C (1992) Do childhood anxiety measures measure anxiety? J Abnorm Child Psychol 20(4): 367–378.
5. Perrin S, Last C. (1992) Do childhood anxiety measures measure anxiety? J Abnorm Child Psychol 20(4): 367–378.
82. Canavera KE, Wilkins KC, Pincus DB, Ehrenreich-May JT (2009) Parent-report measures of anxiety and depression in a community sample of adolescents. J Anxiety Disord 23(8): 1098–1105.
57. Baldwin JS, Dadds MR (2007) Reliability and validity of parent and child versions of the Multidimensional Anxiety Scale for Children in community samples. J Am Acad Child Adolesc Psychiatry 46(2): 252–260.
71. MASC factor structure and discriminant validity in a mixed clinical outpatient sample. Depress Anxiety 25: 929–936.
85. Muris P, Merckelbach H, Van Brakel A, Mayer AB (1999) The revised version of the Screen for Child Anxiety-Related Emotional Disorder-71 as an assessment tool for school phobia in primary care paediatrics. J Pediatr Psychol 31(5): 512–521.
93. Stearns L, Last C, Higa-McMillan CK, Chorpita BF, Klett RG, et al (2008) The psychometric properties of the MASC in five European countries. J Anxiety Disord 25(1): 109–119.
94. Stearns L, Last C, Higa-McMillan CK, Chorpita BF, Klett RG, et al (2008) The psychometric properties of the MASC in five European countries. J Anxiety Disord 25(1): 109–119.
93. Russell E, Sofronoff K (2005) Anxiety and social worries in children with Asperger syndrome. Aust N Z J Psychiatry 39: 633–638.
94. Spence SH, Barrett PM, Turner CM (2003) Psychometric properties of the Spence Children's Anxiety Scale with young adolescents. J Anxiety Disord 17: 605–625.
95. Whiteside SP, Brown AM (2008) Exploring the utility of the Spence Children's Anxiety Scales parent- and child-report forms in a North American sample. J Anxiety Disord 22: 1440-1446.
96. Whiteside SPH, Gryczkowski MR, Biggs BK, Fagen R, O'Connor D (2012) Validation of the Spence Children's Anxiety Scale's obsessive compulsive subscale in a clinical and community sample. J Anxiety Disord 26: 111–116.
97. Lewin AB, Peris TS, de Nadai AS, McCracken JT, Piacentini J (2012) Agreement between therapists, parents, patients, and independent evaluators on clinical improvement in pediatric obsessive compulsive disorder. J Consult Clin Psychol 80(6): 1103–1107.
98. Browne MW, Cudek R (1993) Alternative ways of assessing model fit. In: Bollen KA & Long JS, editors. Testing Structural Equation Models. Newburg Park, CA: Sage. 136–162.
99. Grounds SJ, Aman MG (2012) Assessment of anxiety in children and adolescents with autism spectrum disorders. Res Autism Spectr Disord 6: 1345–1365.
100. White SW, Oswald D, Ollendick T, Scahill L (2009) Anxiety in children and adolescents with autism spectrum disorders. Clin Psychol Rev 29(3): 216–229.
101. MacNeil BM, Lopes VA, Minnes PM (2009) Anxiety in children and adolescents with autism spectrum disorders. Res Autism Spectr Disord 3: 1–21.
102. Leyer OT, Folstein SE, Ricalman S, Davis NO, Dinh E, et al (2006) Comorbid psychiatric disorders in children with autism: interview development and rates of disorders. J Autism Dev Disord 36: 849–861.
103. Hedges DW, Brown BL, Sisalb DA (2009) A direct comparison of effect sizes from the clinical global impression-improvement scale to effect sizes from other rating scales in controlled trials of adult social anxiety disorder. Hum Psychopharmacol Clin Exp 24: 35-40.
104. Hale WH, Crocetti E, Raajmakers Q, van Meurs WHJ (2011) A meta-analysis of the cross-cultural psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED). J Child Psychol Psychiatry 52(1): 80–90.
105. Reutenskiold L, Öst LG, Ollendick T (2008) Exploring child and parent factors in the diagnostic agreement on the Anxiety Disorders Interview Schedule. J Psychopathol Behav Assess 30: 279–290.
106. Blakeley-Smith A, Reaven J, Ridge K, Hepburn (2012) Parent-child agreement of anxiety symptoms in youth with autism spectrum disorders. Res Autism Spectr Disord 6: 707–716.
107. World Health Organization (2007) International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY). Geneva: World Health Organization.
108. Gotham K, Bishop SL, Huynh V, Huerta M, Land S, et al (2013) Exploring the relationship between anxiety and insistence on sameness in autism spectrum disorders. Autism Res 6: 1: 33–41.
109. Halliet V, Roncaud A, Golvert E, Ames C, Woodhouse E, et al (2013) Exploring anxiety symptoms in a large-scale twin study of children with autism spectrum disorders, their co-twins and controls. J Child Psychol Psychiatry 54: 11: 1176–1185.