The measurement properties of the Spence Children’s Anxiety Scale- Parent version in a large international pooled sample of young people with Autism Spectrum Disorder

Final Author Names Listed (in order of authorship with and without qualifications)


*These two authors have contributed substantially and equally to the conceptualization, development, analyses and writing up of this paper and are both joint first authors.

Authors’ Professional Affiliations

1 Department of Psychology, National University of Singapore, 9 Arts Link, Singapore 117570, Singapore;
2 Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology & Neuroscience, King's College London, DeCrespigny Park, Denmark Hill, London, SE5 8AF, UK;
3 Department of Clinical Psychology, Faculty of Medicine and Health Sciences, Norwich Medical School, University of East Anglia, Norwich, UK;
4 Olga Tennison Autism Research Centre, School of Psychological Science, La Trobe University, Victoria 3086, Australia.
5 Cooperative Research Centre for Living with Autism (Autism CRC), Long Pocket, Brisbane, Queensland, Australia
6Clinical Psychology, Ridley Building, Institute of Neuroscience, Newcastle University, Newcastle, NE1 7RU, UK;
7Institute of Health and Society, Newcastle University, Sir James Spence Institute, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne; NE1 4LP, UK
8Evelina London Children’s Hospital, St Thomas’ Hospital, London, UK;
9Department of Psychology, College of Family, Home and Social Sciences, Brigham Young University, Provo, USA;
10Department of Psychology, College of Arts and Sciences, Marquette University, Cramer Hall, Milwaukee, WI 53201-1881, USA
11Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford University, Stanford, CA 94305, USA
12Wales Autism Research Centre, School of Psychology, Cardiff University, Cardiff, CF10 3AT, UK;

Corresponding Author’s Address

Dr. Iliana Magiati, Department of Psychology, National University of Singapore, 9 Arts Link, #02-24, Block AS4. Singapore 117570.

Tel. +65 65161143.

Email: psyim@nus.edu.sg

Running title: Measurement properties of the SCAS-P in ASD

Number of manuscript/ text pages: 24

Number of tables: 9

Number of figures: 1

Grant Information: see Acknowledgements for a breakdown of funding for all studies in this pooled international dataset.
The measurement properties of the Spence Children’s Anxiety Scale- Parent version in a large international pooled sample of young people with Autism Spectrum Disorder

Lay Abstract
Young people with ASD are often affected by anxiety. Assessing and measuring anxiety in ASD reliably and accurately is challenging, as research has yet to identify which existing anxiety measures are most useful and relevant when used with youth with ASD. The present study examined the measurement properties and factor structure of the Spence Children’s Anxiety Scale-Parent Version (SCAS-P) in a large international pooled sample of youth with ASD. Data from 870 participants from 12 studies in the UK, USA and Singapore were pooled. The accuracy of the existing SCAS-P full scale and its subscales was as good as that reported in typically developing children. The subscale measuring fear of animals, insects, environment and doctors, however, had poor accuracy. Thirty items out of the existing 38 SCAS-P items were identified to measure anxiety more accurately than the existing SCAS-P, but the way these items were organized together into subtypes was different to the anxiety subtypes commonly found in typically developing children. Moreover, these subscales were not consistent in one half compared to the other half of this large sample. The limitations of the present study, the use of SCAS-P to screen for anxiety problems in ASD and future research directions are discussed.

Scientific Abstract
Anxiety-related difficulties are common in ASD, but measuring anxiety reliably and validly is challenging. Despite an increasing number of studies, there is no clear agreement on which existing anxiety measure is more psychometrically sound and what is the factor structure of anxiety in ASD. The present study examined the internal consistency, convergent, divergent and discriminant validity, as well as the factor structure of the Spence Children’s Anxiety Scale-Parent Version (SCAS-P), in a large international pooled sample of 870 caregivers of youth with ASD from 12 studies in the UK, USA and Singapore who completed the SCAS-P. Most were community recruited, while the majority had at least one measure of ASD symptomatology and either cognitive or adaptive functioning measures completed. Existing SCAS-P total scale and subscales had excellent internal consistency and good convergent, divergent and discriminant validity similar to or better than SCAS-P properties reported in typically developing children, except for the poorer internal consistency of the physical injury subscale. Confirmatory Factor Analysis (CFA) of the existing SCAS-P six-correlated factor structure was a poor fit for this pooled database. Principal component analysis using half of the pooled sample identified a 30-item five correlated factor structure, but a CFA of this PCA-derived structure in the second half of this pooled sample revealed a poor fit, although the PCA-derived SCAS-P scale and subscales had stronger validity and better internal consistency than the original SCAS-P. The study’s limitations, the use of the SCAS-P to screen for DSM-derived anxiety problems in ASD and future research directions are discussed.

Key Words: autism spectrum disorder, anxiety, parent report, measurement, assessment, reliability, validity, factor structure.
The measurement properties of the Spence Children’s Anxiety Scale: Parent version in a large international pooled sample of young people with Autism Spectrum Disorder

Introduction

Individuals with ASD have significantly higher rates of clinically elevated anxiety symptoms (10-84%) or diagnosed anxiety disorders (about 40%) compared to individuals without ASD or with other conditions (van Steensel et al., 2011; White et al., 2009). Anxiety-related difficulties can significantly interfere with and negatively impact development, functioning and quality of life in ASD (Davis III et al., 2014; Pellecchia et al., 2016). However, the identification and measurement of anxiety in ASD is often complex and challenging.

Emerging evidence suggests people with ASD experience both “typical” anxiety (i.e. worries about separation, achievement or social evaluation, common specific phobias) as well as more idiosyncratic anxiety presentations more specifically relating to ASD (i.e., worries about change, specific sensory related fears; social anxiety without fear of negative evaluation; see Ozsivadjian et al., 2012; Kerns et al., 2014; Trembath et al., 2012; Rodgers et al., 2016). Furthermore, there are often significant difficulties in disentangling anxiety from ASD symptoms (Lecavalier et al., 2014). For instance, worries about change may be part of generalized worry or a feature of restricted, repetitive behaviors (Kerns & Kendall, 2012). Nevertheless, a small number of factor analytic studies have been able to distinguish anxiety from ASD (i.e., White et al., 2012; Renno & Wood, 2013), suggesting that these can likely be disentangled both statistically and clinically (see also Kerns et al., 2016).

Children with ASD also vary considerably in their intellectual and verbal abilities,
insight, emotional understanding and expression, and physiological symptoms (Cook et al., 2013; Didehbani et al., 2012; Mazefsky et al., 2011; Oszivadjian et al., 2012), which can impact their ability to rate or describe their emotional states (Grondhuis & Aman, 2012; Lecavalier et al., 2014). This often necessitates obtaining multi-informant reports from caregivers or significant others. Although anxiety rating agreement between caregivers and children with ASD has ranged from poor (i.e. Kaat & Lecavalier, 2015; Kerns et al., 2015; Renno & Wood, 2013) to moderately good (i.e. Magiati et al., 2014; Blakeley-Smith et al., 2012; Oszivadjian et al., 2014; van Steensel et al., 2012), Storch et al. (2012) found that parental reports contributed significantly to the clinical diagnosis of anxiety. Thus, parent report remains an important way of assessing anxiety in youth with ASD.

Another current consideration is whether adapted or modified anxiety measures are required, or whether existing measures developed for typically developing children have adequate reliability and validity for youth with ASD (Kerns & Kendall, 2012; Ollendick & White, 2012). There is, to date, no anxiety-specific informant measure developed specifically for ASD\(^1\), with the exception of the Anxiety Scale for Children – Autism Spectrum Disorder (ASC-ASD; Rodgers et al., 2016) which showed promising psychometric properties. A small, but growing, number of studies have examined the psychometric properties of existing caregiver-reported anxiety measures when used with 7-18 year old individuals with ASD and

---

\(^1\) A small number of measures of broader psychopathology, including but not specifically focusing on anxiety, have been developed and tested in people with ASD. A modified structured caregiver interview piloted with children with ASD (Autism Comorbidity Interview-Present and Lifetime; ACI-PL; Leyfer et al., 2006) is currently not available for researchers or clinicians to use. The ASD-Comorbid for Children (Matson & Wilkins, 2008) and the Baby and Infant Scale for children with Autistic Traits (BISCUIT; Matson et al., 2009) are informant rating scales with evidence of satisfactory psychometric properties, but which assess only a small number of anxiety symptoms as part of mixed worry/ depressed or anxiety/ repetitive behaviour subscales or which have a very limited age range.
verbal or cognitive functioning standard scores above 60 (see Table 1 for summary of caregiver reported measures; see also reviews by Lecavalier et al., 2014 and Wigham & McConachie, 2014). The parent measures examined tended to have satisfactory to excellent internal consistencies, and superior convergent and divergent validity than child self-reports. Few studies have examined alternative clinical cut-offs or have identified a consistent anxiety factor structure in ASD (see Table 1; White et al., 2015).

**The present study: rationale and research aims/questions**

Despite the recent availability of a promising anxiety specific measure based on ASD-relevant anxiety factors (Rodgers et al., 2016), there is still a need to further examine the factor structure of “typical” anxiety in ASD (White et al., 2015) and to identify a psychometrically sound measure to assess traditional DSM-derived anxiety symptoms in ASD in order to elucidate the common anxiety features across different populations.

The present study therefore examined the psychometric properties and factor structure of the Spence Children’s Anxiety Scale-Parent version (SCAS-P). The SCAS-P provides a total, as well as six DSM-oriented subscale, scores and covers a wide range of anxiety symptoms, including a number of common specific phobia (the most commonly reported anxiety disorder in ASD; van Steensel et al., 2011) and social anxiety items assessing fear of negative evaluation (which may help in distinguishing caregiver-reported social avoidance due to fear of negative evaluation or due to ASD-related social communication challenges). It was selected in this study as it (a) has strong psychometric properties with typically developing children from diverse backgrounds (i.e. Spence, 1999; Nauta et al., 2004; Whiteside & Brown, 2008; Wang et al., 2015; Arendt et al., 2014); (b) is designed to parallel
DSM anxiety disorders’ criteria (albeit DSM-IV-TR criteria, as it was developed before the release of DSM-5 in 2013); (c) it is one of the most frequently used anxiety measures in ASD research (Grondhuis & Aman, 2012; Wigham & McConachie, 2014) and (d) it is a freely available informant measure officially translated in more than 20 languages, making it cost-effective and easily accessible internationally in resource-limited clinical and research settings.

In terms of its measurement properties, Zainal and colleagues (2014) reported that the SCAS-P full scale and subscales had acceptable to good internal consistency (α= .60 to .88), and satisfactory sensitivity (.75) and specificity (.71) against the K-SADS clinical interview using a cut-off score of one standard deviation above the SCAS-P normative mean (Nauta et al., 2004) in a preliminary study of 32 youth with ASD. It has shown good preliminary evidence of moderate to good parent-child agreement (ICC= .59 to .69; Magiati et al., 2014; Ozsivadjian et al., 2014), although May and colleagues (2015) found poor parent-child agreement in a younger sample of cognitively able children with ASD. To our knowledge, no study has yet examined the measurement properties of the SCAS-P for use in children with ASD with a sufficiently large sample, nor has its factor structure been explored in ASD. The present study therefore aimed to investigate the psychometric properties (internal consistency, convergent, divergent and discriminant validity) and factor structure of the SCAS-P in a large, international, multi-site pooled sample of children with ASD.
Methods

Participants

Data were pooled from 12 different studies from the UK (N=9 studies; 465 participants), Singapore (N=1 study; 241 participants) and the USA (N=2 studies; 164 participants; see Table 2 for participant characteristics and details of original studies).

Caregivers of 870 children (763 males, 87.7%) aged 5.58- 18.67 years old (M = 11.6, SD = 2.77) participated. Most of the informants were mothers (75%; Table 2). All 870 participants had a professional clinical diagnosis of Autism or ASD (727; 83.7%), Asperger’s (111; 12.8%) or Pervasive Developmental Disorder-Not Otherwise Specified (32; 3.7%). In all three countries, community professional diagnoses were based on national guidelines for ASD diagnosis involving the use of DSM-IV-TR or ICD-10 diagnostic criteria and a number of validated diagnostic assessments, including semi-structured caregiver interviews and child observations by qualified professionals (see Table 2). In addition, a number of studies also included a measure of autism symptom severity (n=479 participants; 55%; see Measures), of which 393 / 82.3% scored above clinical cut off on the respective caregiver-reported ASD measure or the clinician-rated Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2001²).

² 130 of 130 (100%) met SRS cut-off (Constantino & Gruber, 2005) in studies 6, 11 and 12; 102 of 122 (83.6%) met SCQ cut-off (Rutter et al., 2003) in studies 1, 9 and 10; 162 of 238 (68.1%) met the DBC-ASA cut-off of 14 in Study 7 (Steinhausen & Metzke, 2004); and 118 of 124 (95.2%) met ADOS autism-spectrum cut-off (Lord et al., 2001) in studies 3, 4, 11 and 12.
Twenty-one participants (2.4%) from Study 3, who had a caregiver-reported co-
occurring clinical diagnosis of anxiety disorder and who participated in Study 3 because they
were seeking treatment for anxiety related concerns, constituted the clinical subsample in the
present pooled database. All other participants were the “unselected” community sample,
recruited from non-help seeking for anxiety settings (i.e. national or local autism research
databases, special or mainstream schools, parent support groups, clinics where referral was
for ASD but not for anxiety, etc.).

Measures

**Anxiety.** The *Spence Children's Anxiety Scale-Parent Version* (SCAS-P; Spence,
1999; Nauta et al., 2004) is a caregiver-completed DSM-IV-TR derived anxiety measure
comprising 38 items rated on a 4-point scale (from 0 to 3; higher scores=more anxiety)
assessing symptoms in six subscales: separation (6 items; score range 0-18), social (6 items;
0-18), generalized (6 items; 0-18), panic/agoraphobia (9 items; 0-27), physical injury/
specific phobias (5 items; 0-15) and obsessive compulsive disorder (6 items; 0-18). A SCAS-
P total score (range 0-114) one SD or more above the normative mean (mean 14.2; SD=9.7 in
Nauta et al., 2004) is considered to be clinically elevated (Spence, personal communication,
October 2012). In typically developing children, the SCAS-P has a factor invariance of six
factors across age, gender and different countries, excellent convergent and divergent
validity, acceptable to excellent internal consistency, and good discriminant validity between
anxiety disordered and non-clinically anxious groups (Nauta et al., 2004; Whiteside &
Brown, 2008; Li et al., 2011; Zhao et al., 2012). It also appears to have promising
psychometric properties in youth with ASD in studies in the UK, Singapore and Australia
Autism Symptom Severity. The caregiver or teacher-reported Social Responsiveness Scale (SRS; Constantino & Gruber, 2005; n=130), the 29-item Developmental Behavior Checklist (Einfeld & Tonge, 2002) Autism Screening Algorithm score (DBC-ASA; n=238)\(^3\), the Social Communication Questionnaire (SCQ; Rutter et al., 2003; n=122) or the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2001; n=124) were used to establish the number of participants scoring above cut-off scores as a way to confirm caregiver reported clinical diagnosis of ASD and to measure autism symptom severity in some of the studies (see Table 2).

IQ/ Adaptive Functioning. The Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 1991; n=20), the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999; n=161), the Stanford Binet Intelligence Scales–Fifth Edition (SB5; Roid, 2003; n=54), the Kaufman Brief Intelligence Test–Second Edition (KBIT-2; Kaufman & Kaufman, 2005; n=43) or the Scales of Independent Behavior–Revised Short Form (SIB-R Short; Bruininks et al., 1996; n=239) were used as a measure of overall IQ or adaptive functioning in in some of the studies (see Table 2). All these measures have a normative mean of 100 and a standard deviation of 15.

Procedure

---

\(^3\) The Developmental Behavior Checklist (Einfeld & Tonge, 2002) Autism Screening Algorithm score (DBC-ASA) has been found to discriminate well between children with disabilities with and without ASD (Brereton et al., 2002; Steinhausen & Metzke, 2004).
All studies were approved by their respective institutional ethics committees. The caregivers and youth were recruited according to the studies’ approved research protocols and completed the measures summarized above and in Table 2.

**Statistical Analytical Plan**

**Missing data.** No participant was excluded, and there was no more than 10% missing data for any one particular measure (Bennett, 2001). All missing data were managed and replaced according to manual/measure guidelines.

**Data Harmonization.** Although pulling data together has many advantages in terms of increasing power, allowing the creation of larger and likely more representative of the population samples and maximizing the use of smaller-scale data (Griffith et al., 2015; Hussong et al., 2014), integration of data from different studies often employing different measures is also very challenging. Careful consideration therefore needs to be given to how the datasets can be harmonized as much as possible. Integrating data from intellectual, adaptive and autism symptom severity measures was complex, as different measures and informants were employed. To address this, and following methods discussed by Griffith and colleagues (2015), Hussong et al., (2013) and Schaap et al. (2011), we harmonized cognitive and adaptive functioning data, all of which had the same normative mean of 100 and SD of 15 by creating an ordinal “approximate level of functioning” classification variable with scores ranging from 1 (=standard score <40 corresponding to severe/profound ID range as per DSM) to 8 (SS ≥120, corresponding to superior functioning as per the Wechsler scales’

---

4 Average caregivers’ ratings were used to tabulate the subscale and total scores for eight participants in Study 7 for whom SCAS-P data were given by both caregivers.
classification). To ensure higher levels of consistency, we harmonized only caregiver-reported autism symptom severity measures by converting the raw scores from the different autism screening measures to a 0 to 1 metric (dividing the total raw score of each participant by each scale’s maximum score).

**Statistical Analyses.** To examine their fit in our ASD sample, a series of SCAS-P factor models identified in the literature (see Results) were constructed in CFA models within a general structural equation model framework with a robust maximum likelihood (MLR) estimator in MPLUS v5 (Muthén & Muthén, 2012) to generate model fit indices. The MLR estimator was used due to non-normally distributed data. The Satorra-Bentler Scaled Chi-Square test was used to compare models to account for the scaling due to the use of the robust estimator. As the Chi-Square statistic tends to underestimate goodness of fit in large sample sizes (Bollen, 1989), Hu and Bentler (1999) recommend examining multiple indices RMSEA ≤.06; CFI ≥.95; and TLI≥.95 to determine the model fit.

Participants (n=849) were then randomly assigned into two groups (n₁= 425; n₂= 424) with similar rates of participants from the three countries. Each group was deemed

---

5 8= Standard Score ≥120 (superior); 7= 110-119 (high average); 6= 90-109 (average); 5= 80-89 (low average); 4= 70-79 (borderline); 3= 55-69 (mild ID range); 2= 40-54 (moderate ID range); and 1= < 40 (severe or profound ID range); categories based on the Wechsler scales classification and DSM-5 ID ranges. All cognitive and adaptive measures employed in the different studies correlate very highly with each other in data presented in their respective manuals. We acknowledge that different measures assess different components of intellectual/adaptive functioning and similar, but also different, underlying constructs, which is why we did not use the actual scores, but rather the broader classification categories. This approach maximizes use of all available data, while still acknowledging that the scores obtained are highly correlated, but not directly comparable – this is the reason we did not use the continuous standard scores.

6 In Study 12, 43 participants had ADOS, SRS, and SCQ data; the SRS scores were used for these participants, because its scoring method provides greater score range.

7 The clinical subsample of 21 participants from Study 3 were not included in these analyses, as they had a known identified anxiety disorder and were recruited into the research study due to specifically seeking intervention for anxiety related difficulties.
sufficiently large to explore the underlying factor structure using Principal Component Analysis (PCA) and validating the identified factor structure using CFA (Norris & Lecavalier 2010; Tabachnick & Fidell, 2013; Yong & Pearce, 2013). The two subgroups were not significantly different in terms of chronological age, gender, ASD symptomatology, functioning, and SCAS-P anxiety total score (p>.05; all were small effect size differences, except harmonized functioning which was small-to-medium). The PCA (with direct oblimin rotation, as the factors were theoretically and empirically correlated; Spence, 1998) was run with the first group. Parallel analysis and scree plots determined the number of factors. The PCA derived model was then tested for model fit in the second subsample.

The reliability and validity of the existing SCAS-P and PCA-derived SCAS-P was examined in the full sample. Pearson’s, biserial, Spearman’s Rho and Cramer’s V correlations examined the relationship between SCAS-P total and subscale scores with demographic variables, harmonized autism symptomatology, and functioning classification. Welch ANOVAs with Bonferroni corrections, Chi square tests, and Games–Howell Post-Hoc tests were used to examine differences (i) among the three countries; (ii) between those who met ASD cut-off in the screening measures (n=394) and those who had professional diagnoses of ASD but did not meet ASD cut-off or did not have measures of autism symptomatology completed (n=476); and (iii) between community (n=849) and clinically referred for anxiety (n=21) participants. Effect sizes were interpreted as per Cohen (1988; e.g. r<.30 small; .30-.49 medium; >.50 large; d <.49 small; .50-.79 medium; large >.80).
Results

Preliminary analyses: differences between samples/countries and between subsamples with and without ASD screening data

Excluding the clinically referred for anxiety subsample from Study 3 (n=21), the three country subsamples differed statistically significantly in chronological age, SCAS-P total score, ASD symptomatology, functioning classification, and gender rates with small to large effect size differences (Table 3). Post-hoc Games–Howell tests showed that the UK subsample had higher SCAS-P total score and ASD symptomatology compared to the US and Singaporean participants with medium-to-large effect size differences. The UK and US participants were also significantly older and higher functioning than those from Singapore with a small and large effect size difference respectively (Table 3). However, the country differences in SCAS-P total scores were no longer significant after controlling for age, gender, functioning, and ASD symptomatology (F(2, 355)=1.91, p=.15, d=.049).

Also excluding the clinically referred subsample, there were statistically significant differences in chronological age, SCAS-P (total, social and generalized anxiety scores only), mean functioning classification, and gender between the participants with professional diagnoses of ASD who scored above clinical cut-off in one of the different screening measures employed in the pooled studies (n=376) and the subsample with professional diagnoses who scored below recommended clinical cut-offs in the screening measures administered or who had no ASD symptom ratings available (n=473), but effect sizes of these differences were mostly small (p<.05, .007 ≤ effect sizes ≤ .19; Table 4), with the exception of functioning classification, where the participants not meeting cut-off or not having ASD
screening data available had somewhat higher functioning rankings (Table 4). The difference in SCAS-P anxiety total score between the two subsamples remained significant when age, gender, and functioning classification were controlled for, but the effect size of the difference was very small (Table 4). Because of the small effect size differences between those with and without ASD screening measure data, and as all studies had recruited participants with valid professional diagnoses following established national procedures, we proceeded with full sample analyses.

**Comparison with SCAS-P norms**

The pooled unselected subsample had significantly higher mean SCAS-P total and subscale scores than published norms ($p<.001$; Nauta et al., 2004; Table 5). Using the suggested cut-off score of >24, 76.2% (16/21) participants in the clinical subsample as compared to 46.8% (397/849) from the unselected subsample scored above cut-off.

**Item analyses**

All the items had at least .33 corrected item-total correlations, except item 16 (“my child needs to keep checking that s/he has done things right”; $r=.13$). The average corrected item-total correlation for the 38 SCAS-P items was .52 ($SD=.11$). The average corrected item-subscale correlation was .53 ($SD=.11$) and the average corrected subscale-total correlation was .69 ($SD=.10$).

**Internal Consistency**

Internal consistency for the total score items was excellent at $\alpha=.93$. For the subscales, all Cronbach's alphas were >.75, except physical injury which was suboptimal (Table 5).

**Validity**
**Convergent validity.** The SCAS-P total and subscale scores had moderate to strong positive correlations with the DBC-anxiety subscale ($r_{(236)} = .64$ for total; $r_{(236)} = .32$-$ .56$ for subscales; all $p < .001$; $n = 238$ from Study 7 only in which used the DBC).

**Divergent validity.** SCAS-P had non-significant or small correlations (<.20) with age, gender, and overall functioning classification, with the exception of a positive small-to-medium effect size relationship between age, functioning classification and social phobia, and between overall functioning and social and generalized anxiety (Table 6). ASD symptomatology had significant medium-to-large positive correlations with SCAS-P total and subscale scales, except for the physical injury subscale which had a small correlation (Table 6). Using data from Study 7 only, the SCAS-P had small-to-moderate correlations with the DBC-disruptive/antisocial subscale ($r_{(236)} = .47$ for total; $r_{(236)} = .23$-$ .44$ for subscales; all $p < .001$).

**Discriminant validity.** The clinical subsample from Study 3 ($n=21$) had significantly higher SCAS-P total scores, generalized and social anxiety scores than the unselected subsample (Table 7).

**Confirmatory Factor analysis**

Four different factor models examined by Nauta et al. (2004) in the normative SCAS-P sample were explored (Table 8). The one factor model provided a better fit for the data than the null model, but the six uncorrelated factor model did not improve the model fit. The six correlated factor model provided a better fit than six uncorrelated factors. Lastly, the 5-factor model with Generalized Anxiety as a second-order factor provided a significantly better fit than the previous models. Overall, however, the indices suggested that none of these models
provided adequate fit for our ASD sample. A PCA was thus conducted to explore the SCAS-P factor structure in our sample\(^8\).

**Revised Factor Structure Analyses**

**Principal Component Analysis.** PCA was first run on the 38 SCAS-P items with the randomly selected first half of the sample’s participants \((n_1=425)\), with direct oblimin (delta=0) rotation. The Kaiser-Meyer-Olkin test of sampling adequacy was excellent at .93 (Field, 2009). Bartlett’s test of sphericity was also significant \((\chi^2(703, n=425) = 7891.76, p < .001)\) indicating that the inter-item correlations were adequate for PCA. The points of inflexion on the Scree plot suggested one, three, or five components (Figure 1), while eight factors had eigenvalues >1. Parallel analysis (O'Connor, 2000) suggested five components.

Extracting for five components and suppressing correlation coefficients <.38 to ensure that the selected items had minimal relation with the component, the initial five components accounted for 51.1% of the variance. As several items had communalities significantly below .40 (Costello & Osborne, 2005) or factor loadings below .38, they were removed sequentially and the PCA was re-run each time until a simple structure was achieved (Thurstone, 1947). Items 16, 23, 29, and 34 were removed sequentially because of low communality, followed by items 3 and 36 because of low factor loadings <.38, and lastly items 17 and 38 due to cross-loading. A simple factor structure of five components was obtained with the remaining 30 items (see Table 9) accounting for a total of 57.9% of the variance\(^9\).

---

\(^8\) We also ran the CFA with the participants scoring above cut-off in the autism screening measures only and the results were consistent with the CFA for the full sample, therefore we report the full sample analyses.

\(^9\) We first ran the PCA following the same steps and criteria only with the participants who scored above cut-off in one of the three caregiver-reported screening measures employed in the different studies and found the same 5-factor structure explaining very similar variance (54.8%), the only difference being that two more SCAS-P
The five components had significant positive moderate-to-high inter-correlations ($r \geq .38$). The first component, Social/Generalized Anxiety symptoms, accounted for 35.2% of the variance and its nine items were mainly on social anxiety, but with some generalized and separation anxiety symptoms. The second component accounted for 7.3% of the variance and contained 5 items, measuring mostly separation anxiety symptoms but also fear of darkness and feeling scared. The third component contained eight items measuring somatic or panic symptoms accounting for 6.5% of the variance. The fourth factor contained four items all assessing obsessive-compulsive symptoms and accounted for an additional 4.7% of variance. The last component was a Specific Phobia factor containing four items on medical/dental phobia and fear of public places or using public transport and accounting for 4.2% of the variance.

**Item analyses and reliability of revised PCA-derived SCAS-P.** The item means and SDs, corrected item-total, corrected item-subscale and corrected subscale-total correlations of the revised 30-item SCAS-P are presented in Table 5. All items had at least .34 corrected item-total correlations. The average corrected item-total correlation, item-subscale correlation and subscale-total correlations were .54 ($SD=0.09$), .59 ($SD=0.08$) and .62 ($SD=0.06$) respectively. Internal consistency for the 30-SCAS-P items was .93 and $\alpha \geq .69$ for the subscales (Table 5).

**Validity of the PCA-derived SCAS-P.** The PCA-derived SCAS-P total and its subscales had good convergent validity with the DBC-anxiety subscale in Study 7 ($r (236)$ items were included in the final solution. Given the consistent findings, we report on the analyses for the full sample, as the strength of this pooled database lies primarily in the large, diverse sample.
=.65 for total; .38-.55 for subscales; all \( p < .001 \). In the same study subsample, the revised SCAS-P total and subscale scores were positively correlated with the DBC-disruptive/antisocial subscale \( r (236) = .26-.46, p < .001 \) with small to medium effect size correlations. In the full sample, the PCA-derived SCAS-P total and subscales scores had non-significant or significant small correlations with age, gender and functioning classification, indicating divergent validity (with the exception of small-to-medium effect size correlations between mixed generalized/social anxiety subscale and age, functioning classification, and ASD symptomatology; Table 6). The SCAS-P total and the mixed generalized/social anxiety subscale scores were significantly higher in the clinical subsample than the unselected participants, providing some preliminary evidence of discriminant validity, although effect sizes were small (Table 7). The original SCAS-P full-scale correlated \( r (868) = .99 \) with the PCA-revised SCAS-P full scale. The original SCAS-P subscales were also positively and highly correlated with their corresponding revised subscales (i.e., \( r = .88-.95 \)), except for the PCA-derived Physical Injury/Specific Phobia subscale which was moderately correlated with the original SCAS-P Physical Injury subscale \( r (868) = .57 \).

**Confirmatory Factor Analysis.** A CFA was then run in the other half of the pooled sample \( n = 424 \) to test the goodness of fit of the PCA derived model summarized above (and in Table 8). The model fit indices (CFI, TFI, RMSEA) suggested that the PCA derived model did not fit the other half of our sample adequately (Table 8).
Discussion

Measurement properties of the original 38-item SCAS-P

Overall, the original SCAS-P full scale and all but one subscales had excellent internal consistencies and convergent, divergent and discriminant validity similar to or better than SCAS-P data reported in typically developing (TD) children (Whiteside & Brown, 2008; Li et al., 2011; Nauta et al., 2004). The Physical Injury subscale had a low Cronbach’s alpha of .55, consistent with previous findings (Li et al., 2011; Magiati et al., 2016; Nauta et al., 2004) suggesting that this particular subscale, although commonly endorsed in our sample, may be inconsistent psychometrically.

However, the SCAS-P total correlated .47 with the DBC antisocial/aggressive subscale, suggesting some overlap between anxiety and externalizing behaviors in our sample. This is consistent with other qualitative and quantitative studies reporting that often anxiety in ASD may come across or be expressed as irritability or aggression (Ozsivadjian et al., 2012; Mazefsky et al., 2012; Lydon et al., 2015). The positive association of age and functioning with the SCAS-P social anxiety subscale is developmentally expected - increasing social anxiety symptoms are often observed with age in typically developing children, while studies in ASD have also shown that with increases in age and ability comes increased awareness of social isolation and difference, a risk factor for developing anxieties about social situations (i.e. Kerns et al., 2016; Sukhodolsky et al., 2008). At the same time, the original SCAS-P performed less well in the present study in terms of structural and discriminant validity as compared to other published studies with typically developing children (Nauta et al., 2004; Whiteside & Brown, 2008).
Factor structure of the original SCAS-P in ASD

The SCAS-P factor models in the existing literature (see Nauta et al., 2004) did not provide an adequate fit for our data, suggesting that the underlying conceptual and structural basis of the SCAS-P anxiety symptoms may to some extent be different in ASD. Other studies exploring factor structures of existing anxiety measures in ASD have also reported similarly poor fits - Renno & Wood (2013) and White et al. (2015) using the Multidimensional Anxiety Scale for Children-Parent version (MASC-P; March et al., 1997); and Stern et al. (2014) using the Screen for Child Anxiety Related Disorders (SCARED-P; Birmaher et al., 1997). White et al. (2015) found a mixed factor of separation anxiety and panic items and two separate social anxiety factors (i.e., evaluation and performance focused) using the MASC-P, while Stern et al. (2014) observed a mixed panic and generalized anxiety factor comprising a mixture of items from the other subscales, such as school phobia, using the SCARED-41-P. It is possible that the lack of fit of the factor structures of existing measures in different samples of participants with ASD may be explained by the fact that anxiety symptoms in ASD may be manifestations of broader and more complex underlying pathways and processes, such as emotional dysregulation, alexithymia, hyper-arousal, intolerance of uncertainty and/or sensory sensitivities (e.g. Kerns et al., 2016; White et al., 2014; Mazefsky et al., 2012; South & Rodgers, 2017). These broader underlying ASD-related vulnerabilities may also explain the findings of “mixed” factors in the earlier and the present studies in ASD. Others, however, have reported preliminary evidence of a good fit (i.e., Hallett et al., 2013a, using the CASI-Anxiety). Of the “traditional” anxiety factors, separation and social anxiety subscales, and panic subscale (physiological symptoms), to a lesser extent, were most consistently replicated in the studies mentioned earlier.
Measurement properties of the revised PCA-derived 30-item SCAS-P

The PCA suggested an alternative five-factor SCAS-P structure, with a number of items loading on factors not commonly suggested in the existing literature (Nauta et al., 2004). For example, social and generalized anxiety items loaded together on one “mixed” factor. As discussed earlier, it is possible that, compared to normative samples, some individual SCAS-P items load onto different anxiety factors because of different, and ASD-distinct, underlying pathways explaining these fears and anxieties (White et al., 2015; Kerns et al., 2016; South & Rodgers, 2017). For example, two items relating to fears of using public toilets or being in crowds load onto the social anxiety and panic/agoraphobia factors respectively in the normative SCAS-P. However, these items load onto the specific phobias factor in the PCA-derived SCAS-P in this study, which could be because these fears may be related to specific overwhelming sensory experiences in toilets and crowds for those with ASD, rather than to “traditional” social anxiety concerns (see Kerns, 2016 for more on the clinical implications of this in the differential diagnosis of anxiety in ASD).

Compared to the original SCAS-P, the revised PCA-derived SCAS-P had similar corrected internal consistencies (with the exception of the specific phobia subscale, which was better in the PCA-derived SCAS-P factor structure), item-total, item–subscale, and subscale-total correlations, divergent validity and discriminant validity, but somewhat stronger convergent validity, as shown by higher correlations with the DBC anxiety subscale. ASD symptomatology was less strongly associated with the PCA-derived SCAS-P subscales, with the exception of a medium association with mixed GAD/social subscale. This could potentially mean that PCA helped remove overlapping items that tap on both anxiety and ASD severity.
The moderate similarities in content and the high positive correlations between the original SCAS-P and the corresponding revised subscales suggest that the underlying anxiety factors may generally be similar. Nevertheless, the lack of a consistent, replicable, and adequate factor structure in the two randomly selected large subsamples in the present pooled dataset is difficult to interpret. Clearly, differences in recruitment and sample characteristics may partially explain the lack of consistency, as we were able to compare the two randomly selected subsamples on only some characteristics and measures. The large heterogeneity of ASD may contribute to less consistent structures, or possibly to different underlying structures for different subgroups. Furthermore, the items to which the caregivers responded were all enquiring about typical anxiety presentations and there was no further clarification or elaborations requested (i.e. whether what parents were describing related specifically to anxiety anticipation or to ASD symptoms and associated distress more broadly; see Kerns, 2016). It is possible that measures which will include both “traditional” as well as more ASD-distinct anxiety presentations and symptoms may derive more comprehensive, and thus more consistent, structures across different samples. Currently, however, due to the concerns with lack of consistency, it is likely premature to use the PCA-derived SCAS-P scale and factor structure for clinical or research purposes.

Comparing the measurement properties of the original/ revised SCAS-P to other existing anxiety measures in ASD

Both the original and revised SCAS-P versions and the SCARED-71-P have demonstrated promising evidence of discriminant validity in terms of significantly higher means in the clinical help-seeking for anxiety sample compared to the unselected community-recruited sample, although this needs to be confirmed with studies employing
clinical diagnostic interviews to establish sensitivity and specificity of the measure. Good sensitivity and specificity for the SCARED-41-P and SCARED-71-P has been reported in a clinically anxious help-seeking sample (Van Steensel et al., 2013), but much poorer accuracy was reported for a number of other parent and self-report anxiety measures developed for typically developing children when compared against the augmented for ASD Anxiety Disorder Interview Schedule Child/Parent (ADIS C/P; Kerns et al., 2015; Stern et al., 2014; van Steensel et al., 2012).

Both original and revised PCA-derived SCAS-P versions and the other anxiety measures examined in the literature so far had poor structural validity, except for the CASI-Anxiety which showed promising evidence of adequate model fit (Hallett et al., 2013a; Renno & Wood, 2013; Stern et al., 2014; White et al., 2015). Overall, with regards to existing anxiety measures, both the original SCAS-P and the SCARED-71-P and CASI-Anxiety appear to be promising in their use with children and adolescents with ASD, but future research needs to examine how to improve their structural validity, sensitivity and specificity for use as a screening tool with this population. It is likely that such improvements in the structural validity will come from including and piloting in existing measures addendum subscales measuring more ASD-specific anxiety presentations (i.e., see Kerns et al., 2017; Bearss et al., 2015; Rodgers et al., 2016).

**Strengths and Limitations of the present study**

This large pooled ASD sample which includes participants from 12 studies from three countries made it possible to examine structural validity and factor structure of, as well as to test the adequacy of the new structure in, a large diverse group of young people with ASD. As most of the participants were recruited through community settings, this allowed the
investigation of the use of the SCAS-P in non-clinically referred young people for anxiety with ASD.

At the same time, however, pooling participants from multiple studies and countries also presented several challenges and harmonization of data from different sources and measures is inherently difficult as there was no common measure for ASD severity, cognitive or adaptive functioning. Some loss in the richness and range of data available for harmonization is inevitable (in the present study, this was the case for scores from intellectual/adaptive functioning and autism symptomatology which were pooled together in ordinal categories, rather than used continuously). It is also possible that pooling data from 12 studies in three countries might have masked other potential differences, such as sociocultural differences, beyond the identified examined differences in age, gender, overall functioning, or ASD symptom severity. Furthermore, only some of the pooled studies had additional child report anxiety measures or clinician diagnostic interviews available, thus we were not able to investigate informant agreement nor to establish which participants met clinician-rated diagnostic criteria for an anxiety disorder. Lastly, although participants across all levels of intellectual/verbal/adaptive functioning were included, our pooled sample comprised more children and adolescents with ASD who were functioning within the non-impaired range in terms of cognitive/adaptive skills.

Possible implications

For Clinical Practice. Following further replication, the SCAS-P appears to be a reasonably good choice as a freely available first line screening tool for typical DSM-derived anxiety in ASD, provided that it is employed alongside other multi-informant and multi-
method information gathering approaches (i.e. using the ASC-ASD by Rodgers et al., 2016 to also examine more ASD-related anxiety difficulties; or the ADIS C/P ASD addendum clinical interview by Kerns and colleagues, 2014; see also Vasa et al., 2016). If the original SCAS-P is used, we recommend the use of the total score as a general screen for anxiety, and not of the separate subscales, as the underlying subscale/ factor structure remains inconsistent and unclear and requires further exploration. We specifically caution against the use of the original SCAS-P Physical Injury subscale, as some of its items may not be tapping on a latent specific phobia factor. Furthermore, clinicians should not simply invite caregivers to “tick” symptoms off the checklist; instead, we encourage them to explore the informants’ responses further, in order to establish the precise content of anxiety concerns and to elucidate whether they may be best explained or understood as anxiety involving anticipation or as a distressed reaction as a consequence of their ASD symptoms (e.g., Kerns & Kendall, 2012; see Kerns et al., 2014; 2016 for evidence and recommendations differentiating ASD and anxiety symptomatology using systematic interviewing procedures).

For Future Research. Several unanswered questions remain. In our summary in Table 1, it appears that measures developed for typically developing youth generally do not have strong evidence for discriminant validity, sensitivity or specificity, or good factor structure fits when employed in youth with ASD. CFAs often reveal that the measures’ original structures have inadequate fits in ASD samples, although studies using EFAs have identified relatively similar factor structures consistently using different measures (i.e., Hallett et al., 2013a; Renno & Wood, 2013; Stern et al., 2014; White et al., 2015; this study).

Future research should thus look into (i) modifying, adapting or enhancing existing measures for individuals with ASD, while gathering normative ASD data; and/ or (ii)
developing and improving ASD-specific anxiety measures (Kerns et al., 2015; Sterling et al., 2015; White et al., 2015; Rodgers et al., 2016). Children’s cognitive profile and verbal skills also appear to affect anxiety symptom endorsement to some extent (Hallett et al., 2013b; Witwer & Lecavalier, 2007) and hence, it will be important to validate the measures, establish norms and cut-offs separately for ASD youths of different levels of ability.

Furthermore, youth with ASD also present with more idiosyncratic ASD-related anxiety (see Bearss et al., 2015; Kerns et al., 2014; Ozsivadjian et al., 2012; for a review, see Magiati et al., 2017). Future research should examine how ASD-related anxiety symptoms relate to traditional DSM-derived anxiety symptoms in young people with ASD (Kerns et al., 2014) and the underlying structural validity of measures including both traditional and ASD-related anxiety symptoms. Researchers can use existing work with focus group methodologies which have generated candidate anxiety items to develop an ASD-specific anxiety measure (Bearss et al., 2015; Rodgers et al., 2016) or augment modified measures of traditional DSM-defined anxiety symptoms with an ASD-addendum subscale (Ozsivadjian et al., 2014; Kerns et al., 2014).

Research efforts to better assess, understand and treat anxiety in children and adolescents with ASD have intensified and have considerably improved our understanding of anxiety in ASD. The present study contributes to this growing body of literature by examining a widely used and easily accessible caregiver-report anxiety measure. Future research efforts should focus on improving our understanding of the structure of, and relationship between, traditional and more ASD-related presentations of anxiety and our ability to reliably and validly identify these in individuals with ASD.
Acknowledgements

We would like to thank all the caregivers and families, schools and clinics, participants and staff involved in all studies and countries. Not in any particular order, Study 1 includes data from independent research commissioned by the UK National Institute for Health Research (NIHR) under the Research for Patient Benefit programme (PB-PG-0408-16069) and data collected by Dr. Janes and Dr. Connolly as part of doctoral theses for the qualification of Doctor of Clinical Psychology. Studies 2 and 5 are grateful to the Database of Children with ASD Living in the North East (www.daslne.org) for assistance with recruitment. Daslne is funded by the UK autism research charity Autistica and by The Children's Foundation. Studies 4 & 5 were supported by funding provided by Northumbria, Tyne and Wear NHS Foundation Trust awarded to Dr. Jacqui Rodgers. Study 3 was commissioned by the National Institute for Health Research (NIHR) under the Research for Patient Benefit programme (PB-PG-0408-16069). Study 6 was supported through a PhD studentship awarded to Dr. N. Darus by the Ministry of Health in Malaysia. Study 7 was supported by a Start-up grant awarded to Dr. I. Magiati from the National University of Singapore, Faculty of Arts and Social Sciences. Study 8 was funded by the NIHR Biomedical Research Centre (BRC) in Mental Health (grant code: PCCKASA) and the South London and Maudsley Charitable Funds (awarded to Prof. E. Simonoff). Study 9 was supported by the Charles Hawkins Fund. Study 10 was supported by a PhD studentship from the Wales Office of Research and Development for Health and Social Care, National Institute for Social Care and Health Research, awarded to Dr. M. Ujlarevic and a project grant from the Baily Thomas Charitable Fund awarded to Prof. S. Leekam. The data collection was carried out at Cardiff University by Dr. M. Ujlarevic and by Dr. J. Lidstone. Study 11 was supported by grants from Simons Foundation Autism Research Initiative, the Mosbacher Family Fund for Autism Research, The Escher Fund at the Silicon Valley Community Foundation, and Stanford University’s Child Health Research Institute. Study 12 was supported by grants from the Autism Society of Southeastern Wisconsin and the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), through grant number 8KL2TR000056 to Dr. A. Van Hecke. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH, NHS, NIHR or Department of Health. The authors report no conflicts of interest.
References


Disorders, 43(4), 851-859.


White, S. W., Bray, B. C., & Ollendick, T. H. (2012). Examining shared and unique aspects


