ORIGINAL PAPER

The Subthreshold Autism Trait Questionnaire (SATQ): Development of a Brief Self-Report Measure of Subthreshold Autism Traits

Stephen M. Kanne · Jennifer Wang · Shawn E. Christ

Abstract The current study was motivated by a need for a self-report questionnaire that assesses a broad range of subthreshold autism traits, is brief and easily administered, and is relevant to the general population. An initial item pool was administered to 1,709 students. Structural validity analysis resulted in a 24-item questionnaire termed the Subthreshold Autism Trait Questionnaire (SATQ; Cronbach's alpha coefficient = .73, test–retest reliability = .79). An exploratory factor analysis suggested 5 factors. Confirmatory factor analysis indicated the 5 factor solution was an adequate fit and outperformed two other models. The SATQ successfully differentiated between an ASD and student group and demonstrated convergent validity with other ASD measures. Thus, the current study introduces and provides initial psychometric support for the SATQ.

Keywords Autism · Broader Autism Phenotype · AQ · BAPQ · Subthreshold

S. M. Kanne (✉)
Department of Pediatrics, Psychology Section, Baylor College of Medicine, Texas Children's Hospital, 1102 Bates, Ste. 730, Houston, TX 77030-2399, USA
e-mail: smkanne@texaschildrens.org

J. Wang
Department of Health Psychology and Thompson Center for Autism and Neurodevelopmental Disorders, University of Missouri, Columbia, MO 65212, USA

S. E. Christ
Department of Psychological Sciences, University of Missouri, Columbia, MO 65211, USA

Introduction

Autism Spectrum Disorders (ASD) are associated with significant impairments across several domains, including communication, behavior, and social interaction (American Psychiatric Association 1994). Recent research examining phenotypic expression of ASD traits has extended our knowledge beyond individuals diagnosed with an ASD to include relatives of an individual with ASD and individuals in the general population. For example, researchers have found that autistic traits are more common in relatives of individuals diagnosed with ASD (Hurley et al. 2007), and are continuously distributed in the general population (Constantino and Todd 2003). Presumably, ASD traits range from the greatest expression in individuals diagnosed with an ASD with decreasing phenotypic expression in non-autistic relatives of autistic individuals, termed the broader autism phenotype (Hurley et al. 2007), and the least expression in the subthreshold traits found in the general population (Constantino and Todd 2003, 2005).

Subthreshold ASD Traits

Individuals in the general population exhibit a wide range of communicative and social abilities, and may demonstrate a range of other idiosyncratic behaviors. Some individuals have a high level of social facility whereas others often struggle in their social interactions. Individuals vary in their ability to use the foundational skills necessary for successful social interactions and communication, such as appropriate facial expressions, pragmatic language skills, modulation of eye contact, appreciation of personal space, appropriate voice intonation, as well as level of empathy, insight, social referencing, joint attention, and use of gestures. Similarly, individuals vary in their level of

interest in topics, need for structure and routine, and other idiosyncratic behaviors. Variability in these areas may reflect normal variation in the population, or depending on their impact and specific manifestation, may be indicative of a disorder. If an individual demonstrates deficits in these areas that have a sufficient impact on their functioning, the diagnosis of an ASD may be warranted.

Researchers have recently argued that the core deficits in ASD, especially social-communication deficits, can be conceptualized as a continuously distributed trait in the population rather than a distinct or discrete disorder (Constantino and Todd 2003; Spiker et al. 2002). Clinicians realize that deficits in these areas do not necessarily mean an individual has an ASD and thus employ sophisticated diagnostic tools to help understand the extent, nature, and etiology of the deficit (e.g., Filipek et al. 2000; Volkmar et al. 1999).

Individuals may demonstrate ASD traits that are not of sufficient magnitude, quality, or number (i.e., subthreshold traits) to warrant a diagnosis of an ASD. Studying individuals with sub-threshold symptomatology of various disorders has been a useful approach allowing researchers to gain valuable understanding. For example, research involving individuals with schizophrenia-related symptoms suggests that they may experience cognitive and emotional difficulties analogous to those problems found in individuals with schizophrenia, but in a more diminished form (Delawalla et al. 2006; Kerns 2006). Other studies have looked at individuals with sub-threshold ASD traits; for example, Jobe and White (2007) found that increased endorsement of autism-related symptomatology in young adults was associated with increased loneliness and a decrease in the number and duration of friendships. Kanne et al. (2009) found that undiagnosed young adults reporting higher sub-threshold ASD traits also report increases across a wide range of psychiatric and psychosocial problem areas including depression/anxiety, interpersonal relationships, and personal adjustment. Individuals with sub-threshold traits of various disorders may have the advantage of being more tolerant of various testing environments, allowing for the use of a broader range of assessment tools and approaches.

The study of subthreshold traits requires some level of quantification of the targeted symptoms. With respect to ASD, there are a number of measures that assess autism symptoms and provide quantitative results. These vary from screening questionnaires filled out by parents, teachers, or the individual (i.e., self-report), to more complete assessment tools that involve extensive interviewing or direct observation. The motivation for the current study arose from a need for a brief, self-report measure that assesses a broad range of subthreshold ASD traits in the general population.

Assessing ASD Traits

As noted, there are a number of measures that assess autism symptoms and provide a quantitative result. Many of these are used to assess individuals as part of a clinical diagnostic referral and thus require extensive clinical knowledge and training and can also be very time consuming to complete. Such measures include the Autism Behavior Checklist (Krug et al. 1988), the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2002), the Autism Diagnostic Interview—Revised (ADI-R; Lord et al. 1994), and the Child Autism Rating Scale (Schopler et al. 1980). Note that many of these measures, though they produce a raw score and are thus "quantitative," vary with respect to their scale of measurement and are not necessarily continuously distributed in the population. With respect to the ADOS, Gotham and colleagues recognized this limitation and have developed a calibrated severity score based on ADOS total scores in an attempt to create a usable severity metric (Gotham et al. 2009). Alternatively, questionnaires are an oft-used method to ascertain the presence of phenotypic ASD characteristics. The ease, timeliness, and ability to gather information from a wide range of sources afforded by these questionnaires often offsets the limitations associated with them, such as response bias, informant reporting, and a psychometric emphasis on sensitivity.

When choosing a questionnaire, considerations include the purpose for which that questionnaire was originally designed (e.g., screening, diagnostic), ascertainment method (e.g., self-report, informant report), and the targeted age range. For example, the modified checklist in toddlers (M-CHAT; Robins et al. 2001) is a popular and widely used tool developed in the United Kingdom that screens for ASD symptoms in the general population. However, the M-CHAT targets toddlers in the general population rather than individuals of a wider age range referred for specific developmental concerns. The Childhood Autism Spectrum Test, or CAST, is a questionnaire completed by parents developed to detect ASD symptoms in primary school children that may be more mild (Scott et al. 2002), and has subsequently been used in several population-based prevalence studies (Baron-Cohen et al. 2009; Ronald et al. 2006). When looking for questionnaires that assess ASD phenotypic expression for an older age range, the field narrows. Four candidate measures that have been used widely are the Social Communication Questionnaire (SCQ; Berument et al. 1999), the Broader Autism Phenotype Questionnaire (BAPQ; Hurley et al. 2007), the Social Reciprocity Scale (SRS; Constantino et al. 2000), and the Autism-Spectrum Quotient (AQ; Baron-Cohen et al. 2001).

The SCQ is a well-validated screening tool for high risk individuals that covers the full breadth of ASD

characteristics and has a set of questions that apply to a specific age, allowing it to be used across a greater age range. The SCQ, derived from the ADI-R algorithm, is completed by the primary caregiver. Like the ADI-R, the SCQ has a version whose questions target the specific age range of 4–5 years of age, the "lifetime" version, and a version that pertains to existing symptom manifestation, the "current" version. The SCQ is somewhat unique in that it was validated on individuals ages 4–18 years that already carried an ASD diagnosis. The format of the SCQ allowed for the construction of "cut-off" scores into the three domains related to an ASD diagnosis. Several factors, however, limit its use in assessing subthreshold ASD traits. For example, the yes/no format of the SCQ precludes a dimensional response for each area probed, making it less suited to exploring subthreshold symptoms, and its method of validation supports its use in those strongly suspected of ASD rather than the general population.

The Social Responsiveness Scale (SRS), formally called the Social Reciprocity Scale (Constantino et al. 2000), is a questionnaire that assesses phenotypic ASD expression in the general population. The SRS is informant based, filled out by a parent or teacher, and assesses behaviors associated with the full range of autistic symptomatology. Using the SRS, researchers were able to demonstrate that reciprocal social behavior, a core feature of ASD, is continuously distributed in the general population (Constantino et al. 2000; Constantino and Todd 2003). In addition to overall social reciprocity deficits, the SRS also includes questions that include other ASD specific behaviors as these may represent the extremes of difficulty. Examples of questions include "has difficulty relating to peers," "is able to imitate others actions," "gets frustrated trying to get ideas across in conversations," "thinks or talks about the same thing over and over," "has a good sense of humor, understands jokes," "has overly serious facial expressions," and "is inflexible, has a hard time changing his or her mind." One restriction of the SRS currently is that its ascertainment method is informant based; that is, there is not a self-report version of the SRS available currently, though brief self-report SRS variants have been used (Kanne et al. 2009; Reiersen et al. 2008).

More recently, researchers have more specifically articulated the notion of a broader autism phenotype, which refers to the expression of an ASD phenotype in non-autistic relatives of an individual with ASD. The expression of ASD symptoms in this population are thought to be milder in expression, but similar in quality. Hurley and colleagues have introduced the BAPQ (Hurley et al. 2007) to assess for these features. The BAPQ consists of 36 questions specifically designed to assess for the broader autism phenotype, such as social personality, rigid personality characteristics, and higher level language difficulties (e.g., pragmatics). In addition to the total score, the BAPQ produces three subscales that ostensibly map onto the three DSM-IV (American Psychiatric Association 1994) domains of social deficits, restricted and repetitive behaviors, and communication difficulties. The BAPQ was found to have high sensitivity and specificity (>70%) for detecting BAP in this population, and the authors support its use as a screening and diagnostic tool. However, as the authors clearly state, this measure was designed and validated to assess autism traits in relatives of individuals diagnosed with ASD, and not in the general population. If the measure had instead been designed and validated in the general population, there may be significant differences with respect to the final items and the scale's psychometric properties.

Another self-report questionnaire has been developed by Simon Baron-Cohen and colleagues (Baron-Cohen et al. 2001). The Autism-Spectrum Quotient (AQ) is a self-report questionnaire that was designed to assess the degree to which an adult with normal intelligence has ASD traits. Baron-Cohen validated this measure on four groups; a group of adults with Asperger Syndrome (n = 58); a group of controls (n = 174); a group of students (n = 840); and a group of individuals who had won a math contest (the UK Mathematics Olympiad; n = 16). The measure differentiated the Asperger Syndrome group from the controls and the students, showed a normal distribution in the controls, a significant difference between genders (with males higher) and types of scientists (with mathematicians higher). Additional versions of the AQ have since been developed, including an adolescent version (Baron-Cohen et al. 2006), a child version (Auyeung et al. 2008), and a brief version (Hoekstra et al. 2011).

Whereas the AQ appears to satisfy the need for a self-report measure of subthreshold ASD traits, it was designed to be especially sensitive to an ASD presentation that is most characteristic of Asperger Syndrome, leading to the question of whether it includes items that cover a broader range of symptoms. This potential weakness became more evident when specific items from the AQ were directly compared to items from the SRS, BAPQ, and SCQ. For example, the AQ does not have questions pertaining specifically to eye contact, being perceived as odd or strange, perception of facial expressions, being physically awkward, using gestures, and sharing enjoyment as do many of the other measures. Thus, it may be "missing" a portion of subthreshold ASD traits.

Current Study

The current study was motivated by a need for a questionnaire that assesses a broad range of subthreshold autism traits, is brief and easily administered, is relevant to the

general population (i.e., individuals not directly related to those with autism), and uses a self-report format. Existing self-report measures assessing ASD traits each has limitations, with the BAPQ being designed and validated for use in relatives of those with ASD and the original AQ being longer and emphasizing Asperger traits rather than a broader spectrum. We desired a measure that was shorter in length, more comprehensive in the breadth of ASD related traits assessed, and applicable to the general population. Instead of attempting to both shorten the original AQ and add specific questions, or extend the BAPQ to the general population, we took a different approach and constructed a new self-report questionnaire, which we term the Sub-threshold Autism Questionnaire (SATQ), felt to represent a range of ASD traits.

Methods

Participants

This study consisted of three samples (see Table 1 for participant demographics):

Total Student Sample

One thousand seven hundred and nine undergraduate students (Age: M (SD) = 18.4 (0.99); Gender: 38.9% male) completed an initial 32-item version of the SATQ as part of a larger survey in return for course credit (see *SATQ Initial Development* below for description of scale development).

This project was approved by a University of Missouri IRB.

Student Subsample

Responses were totaled for the initial 32-item SATQ on the Total Student Sample described above. Separate totals for the three domains corresponding to the 3 primary diagnostic criterion ASD domains of communication, social skills, and restricted and repetitive behaviors were calculated based on an a priori assignment of items to one of the 3 domains. As part of a separate study that was examining the relationship between cognitive profiles and individuals who endorsed a great deal of symptoms in one of the three DSM-IV oriented domains, one hundred ninety-six participants were recruited whose SATQ total score was in the top 15% overall, in the top 10% on one of the three domain subscales, or in the bottom 10% overall.

ASD Sample

Seventeen individuals with a diagnosed ASD were recruited from the community to complete the BAPQ, AQ, and initial 32-item SATQ. These participants were young adults who received clinical services at an interdisciplinary academic medical center specializing in diagnosis and treatment of ASD. Diagnostic decisions of ASD for these individuals were made following a center-based diagnostic interview based on the ADI-R and observation focusing on DSM-IV criteria. Evaluations had been conducted by a pediatrician and/or neuropsychologist; if there was

Table 1 Demographic information for total student sample, student subsample and ASD sample

	Total student sample	Student subsample	ASD sample
Age			
n	1,709	196	17
Mean (SD)	18.4 (0.99)	18.8 (0.95)	20.9 (2.1)
Range	17–28	18–23	18–24
Gender (% male)	39	44	88
24-item SATQ total score			
Mean (SD)	23.1 (7.1)	22.7 (9.5)	40.8 (13.6)
Range	3–55	4–46	7–57
BAPQ total score			
Mean (SD)	–	99.2 (24.9)	147.5 (26.1)
Range	–	43–176	92–177
AQ total score			
Mean (SD)	–	16.2 (5.8)	31.7 (7.0)
Range	–	4–37	15–44
Estimated IQ[a]			
Mean (SD)	–	107.8 (9.9)	–
Range	–	76–131	–

[a] Based on the WASI (Wechsler 1999)

disagreement the results were discussed jointly to reach a consensus diagnosis. Three of these individuals had a diagnosis of Asperger Syndrome, five had diagnoses of Pervasive Developmental Disorder—Not Otherwise Specified, and the remaining nine had diagnoses of Autistic Disorder. These individuals were given the option of completing the questionnaire online or returning a paper version in return for $20 incentive.

Measures

Each participant from the Student Subsample and the ASD Sample completed the initial 32-item SATQ, the BAPQ, and the AQ. In addition, the participants in the Student Subsample also completed the 2-subtest version of the Wechsler Abbreviated Scale of Intelligence.

Broader Autism Phenotype Questionnaire

The Broader Autism Phenotype Questionnaire (BAPQ) (Hurley et al. 2007) is a self report measure assessing the expression of an ASD phenotype in non-autistic relatives of an individual with ASD. The report consists of 36 items assessing traits associated with ASD such as social personality, rigid personality characteristics, and higher level language difficulties. The BAPQ produces a total score and scores on three subscales that ostensibly map onto the three DSM-IV (American Psychiatric Association 1994) domains of social deficits, restricted and repetitive behaviors, and communication difficulties. The BAPQ was found to have high sensitivity and specificity (>70%). Higher scores reflect more severe symptomatology.

Autism Quotient

The Autism-Spectrum Quotient (AQ) (Baron-Cohen et al. 2001) is a self report measure that assesses the presence of autistic traits in adults with normal intelligence. The questionnaire consists of 50 items assessing traits associated with ASD including social skills, attention switching, attention to detail, communication, and imagination. Psychometric study of the AQ indicated good test–retest reliability and internal consistency (Baron-Cohen et al. 2001). Higher scores reflect more severe symptomatology.

Wechsler Abbreviated Scale of Intelligence

The Vocabulary and Matrix Reasoning subtests from the Wechsler Abbreviated Scale of Intelligence (WASI) (Psychological Corporation 1999) were administered to estimate general intellectual ability.

SATQ Initial Development

In an effort to construct a brief, easily administered, self-report questionnaire that assesses for the presence of a broad range of ASD traits, questions from the SRS, SCQ, BAPQ, and AQ were extracted and sorted into categories of perceived overlapping theme (e.g., social enjoyment, eye contact). Instruments were chosen for review based on their general use and acceptance, psychometric properties, and ascertainment method. The notion was to develop and refine a large item pool based on questions from measures with established validity, rather than starting with untested items. Two experienced clinicians (i.e., SC, SK), identified general rubrics for the questions without knowledge of the questionnaire from which they were derived. These categorical labels were then distilled into questions framed in a self-report format thought best to capture the theme of the underlying questions. A conscious attempt was made to include questions that covered the three general areas identified by the DSM-IV regarding social deficits, communication deficits, and restricted and repetitive behaviors, though the degree of overlap between the social and communication areas made this problematic. Several questions were retained conservatively even if they were not represented across questionnaires as they were thought to represent extremities of deficits rather than representing a larger category (e.g., "I am unusually sensitive to textures, sights, smells, tastes, or sounds").

The result was an initial item pool of 32 questions. Sixteen questions were worded such that higher scores represented higher ASD traits, and 16 were worded such that higher scores represented lower ASD traits to avoid a yes/no response bias. Respondents are instructed: "For each item, please use the scale below to rate the extent to which it describes you *on most days*. There are no right or wrong answers. Please answer all of the items the best that you can." They respond on a 4-point Likert Scale with 0 = False, not at all true, 1 = Slightly true, 2 = Mainly true, and 3 = Very true. Table 2 lists the initial 32 SATQ questions.

Statistical Analyses

To improve the structural validity of the scale, item response distributions generated from the Total Student Sample (n = 1,709) were examined first to identify those items that were highly skewed, reducing the scale's validity (Clark and Watson 1995). After removing items with excessive skewness, item-total correlations were examined to identify items not accurately assessing the same underlying construct (Clark and Watson 1995).

The size of the Total Student Sample allowed a split-half analysis. The dataset was divided randomly in half and

Table 2 Initial 32 SATQ items with factor loadings and initial communalities for the final 24 items from first half of the split-half total student sample (n = 855)

	Original assignment on ASD triad	Initial communalities	Factor				
			1	2	3	4	5
% Variance accounted for by factor		Total = 46.4	17.6	10.2	7.3	6.2	5.1
Social interaction & enjoyment							
I like being around other people	Social	0.47	**0.75**	0.02	0.01	0.04	0.06
I enjoy social situations where I can meet new people and chat (i.e. parties, dances, sports, games)	Social	0.47	**0.72**	−0.02	0.05	0.05	0.13
I seek out and approach others for social interactions	Social	0.36	**0.62**	−0.14	0.02	0.08	0.00
I like to share my enjoyment with others	Social	0.31	**0.52**	−0.01	0.13	0.07	0.08
Others consider me warm, caring, and/or friendly	Social	0.30	**0.48**	0.16	0.27	−0.06	−0.07
I respond appropriately to other people's emotions (for example, comforting someone who is upset)	Social	0.25	**0.40**	0.10	0.28	0.10	−0.10
I can have a back and forth conversation (listen well and change topics appropriately)	Communication	0.32	**0.37**	−0.04	0.32	0.37	−0.07
I use many gestures when speaking with others such as shrugging, "talking with my hands," nodding my head, etc.	Communication	0.16	**0.32**	−0.10	0.12	0.07	−0.13
Oddness							
Others think that I am strange or bizarre	Social	0.39	0.06	**0.70**	0.02	0.10	−0.03
I have some behaviors that others consider strange or odd	RRB	0.34	0.01	**0.67**	−0.04	−0.03	−0.01
I sometimes say things that others tell me are rude or inappropriate	Communication	0.24	0.02	**0.52**	−0.02	−0.02	0.02
I use odd phrases or tend to repeat certain words or phrases over and over again	Communication	0.20	−0.09	**0.45**	−0.05	0.11	0.011
I am very interested in things related to numbers (i.e. dates, phone numbers, etc.)	RRB	0.08	−0.08	**0.18**	−0.03	0.17	0.11
Reading facial expressions							
I am good at knowing what others are feeling by watching their facial expressions or listening to the tone of their voice	Communication	0.37	0.18	−0.06	**0.69**	0.12	0.01
I can sense that someone is not interested in what I'm saying by reading their facial expressions	Communication	0.31	0.08	−0.10	**0.61**	0.09	0.05
I make eye contact when talking with others	Communication	0.22	0.30	0.03	**0.31**	0.20	0.06
Expressive language							
I am good at using words to express my thoughts and ideas	Communication	0.29	0.20	−0.05	0.20	**0.60**	0.01
I have difficulty getting my ideas across to others in a conversation	Communication	0.30	0.12	0.27	0.05	**0.57**	0.13
I have a good imagination	RRB	0.17	0.15	−0.18	0.17	**0.22**	0.08
Rigidity							
I am comfortable with spontaneity, such as going to new places and trying new things	RRB	0.31	0.40	−0.14	0.20	0.03	**0.46**
I tend to stick to routines in my day to day life, preferring to do things the same way	RRB	0.13	−0.01	−0.01	−0.03	0.07	**0.45**
I am considered "laid back" and am able to "go with the flow"	RRB	0.20	0.19	−0.01	0.24	−0.07	**0.40**
I sometimes take things too literally, such as missing the point of a joke or having trouble understanding sarcasm	Communication	0.21	−0.11	0.18	0.05	0.36	**0.38**
I tend to focus on individual parts and details more than the big picture	RRB	0.13	−0.08	0.14	−0.12	0.11	**0.26**
Items from original 32 item pool that were removed							
I have interests that occupy much of my time and thoughts (more so than most of my peers)	RRB	–	–	–	–	–	–
Other people can typically tell what I am thinking or feeling based on my facial expressions	Communication	–	–	–	–	–	–

Table 2 continued

	Original assignment on ASD triad	Initial Communalities	Factor				
			1	2	3	4	5
% Variance accounted for by factor		Total = 46.4	17.6	10.2	7.3	6.2	5.1
I notice subtle patterns where most others do not	RRB	–	–	–	–	–	–
I am unusually sensitive to textures, sights, smells, tastes or sounds	RRB	–	–	–	–	–	–
I am awkward or less coordinated compared to my peers	RRB	–	–	–	–	–	–
I am not very good at chatting with others	Communication	–	–	–	–	–	–
I have trouble connecting with my peers	Social	–	–	–	–	–	–
Others say that I speak in a strange way or with an odd tone (robotic, too soft, too loud, etc.)	Communication	–	–	–	–	–	–

Extraction method: principal axis factoring, 5 factors forced (from PFA), varimax rotation

Bold value represents the highest factor loading

an exploratory factor analysis (EFA) was conducted on the first half (n = 855) using a principal axis factor (PAF) analysis with Varimax rotation to examine latent variables. A Parallel Analysis was conducted using the PAF parameters (i.e., 855 subjects, 24 variables, with 100 replications). Parallel Analysis is a Monte-Carlo based simulation method used to determine how many factors to retain from a factor analysis, and is often preferred over other methods such as examination of a Scree plot (Ledesma and Valero-Mora 2007; Zwick and Velicer 1986).

Next, a confirmatory factor analysis (CFA) was conducted with the 2nd half of the Total Student Sample (n = 854) using the AMOS statistical package. Three models were evaluated. The first model was the most parsimonious, one factor solution, also propounded by some researchers (e.g., Constantino et al. 2004). The second model, which forced 3 factors, was based on the notion of the existence of the 3 DSM-IV criterion domains in ASD: social deficits, communication problems, and restricted and repetitive behaviors. The third was the model predicted by the factor structure found in the first half EFA.

Several statistics common to CFA were used to evaluate each model's fit. First, goodness of fit was evaluated using the χ^2 test statistic. In CFA, lower χ^2 values relative to the number of degrees of freedom indicate a better fit. The Standardized Root Mean Square Residual (SRMR), based on the fitted residuals, was also used. An SRMR of less than .05 indicates a "good" fit whereas values smaller than .10 are deemed an "acceptable" fit (Hoekstra et al. 2008; Schermelleh-Engel et al. 2003).

Four other statistics were used for cross model comparisons: the Goodness of Fit Index (GFI), Parsimony Goodness of Fit Index (PGFI), Expected Cross Validation Index (ECVI), and Root Mean Square Error of Approximation (RMSEA). The GFI, ranging between zero and one, measures the relative amount of the variances and co-variances

predicted by the model. The PGFI modifies the GFI taking into account model complexity with a higher PGFI indicating a better model. The ECVI evaluates how well the model would generalize to other samples, with a smaller ECVI indicating a better fit. The RMSEA (Steiger 1990) is a measure of approximate fit in the population, substituting the null hypothesis of exact fit with one of close fit. The RMSEA is bounded by zero with values ≤.05 deemed a "good" fit, between .05 and .08 an "adequate" fit, and between .08 and .10 as a "mediocre" fit (Steiger 1990).

Results

SATQ Structural Validity

Average skewness of all 32 items was 0.64. Four items were identified that were extremely skewed, skewness >1.2 (#14: *I am awkward or less coordinated compared to my peers*; #16: *I am not very good at chatting with others*; #18: *I have trouble connecting with my peers*; and #28: *Others say that I speak in a strange way or with an odd tone (robotic, too soft, too loud, etc.)*. After removing these four items, item-total correlations were examined to identify items not accurately assessing the same construct. Item-total correlations ranged from r = 0.04–0.46. Four items with item-total correlations lower than r = 0.10 were identified and subsequently removed (#2: *I have interests that occupy much of my time and thoughts (more so than most of my peers)*; #9: *Other people can typically tell what I am thinking or feeling based on my facial expressions*; #15: *I notice subtle patterns where most others do not*; and #25: *I am unusually sensitive to textures, sights, smells, tastes or sounds*). Removing these 8 items resulted in the 24 item SATQ.

Total scores from the 24-item SATQ ranged from 3 to 55 in the Total Student Sample (n = 1,709). As shown in
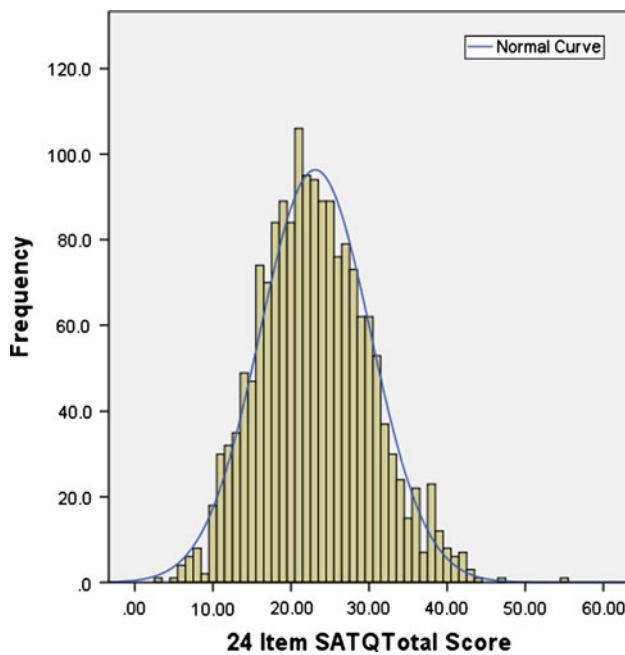
Fig. 1 Distribution of SATQ total scores (24 items) from total student sample (n = 1,709)

Fig. 1, the distribution of scores was generally continuous and slightly positively skewed (mean = 23.1, skewness = .30). Internal consistency for the 24-item SATQ was good, with a Cronbach's Alpha of .73 and an average item-total correlation of $r = 0.28$. Note that a broad underlying construct would be expected to have lower item-total correlations compared to a narrower construct with items that are more focused in content. Given the broad construct of subthreshold autism traits that includes possible deficits across several domains, including items thought to represent extremities of deficits, this item-total correlation average was deemed sufficient.

SATQ Exploratory Factor Analysis

Using the first half of the Total Student Sample, a Parallel Analysis was conducted using the PAF parameters (i.e., 855 subjects, 24 variables, with 100 replications) and indicated five factors should be retained. The five factors accounted for 46.4% of the total variance. Examining the items loading on each factor resulted in the following labels: Social Interaction and Enjoyment, Oddness, Reading Facial Expressions, Expressive Language, and Rigidity. Table 2 reports the initial communalities and factor loadings for each of the 24 items in the initial EFA. Three items (i.e., interested in numbers, good imagination, and focusing on details) with loadings less than 0.30 an any factor were retained and associated with the subscale upon which they had the highest loading due to their theoretical significance and because they may have meaningful discriminate value.

Table 3 Goodness-of-Fit indicators of models for Confirmatory Factor Analysis of the Total Student Sample (n = 854) comparing 1 factor, 3 factor, and 5 factor solutions

	1 Factor model	3 Factor model	5 Factor model
χ^2	2,134.9 ($p <$ 0.001)	1,861.0 ($p <$ 0.001)	981.3 ($p <$ 0.001)
df	252	250	237
χ^2/df	8.47	7.44	4.14
SRMR	0.091	0.090	0.064
RMSEA	0.094	0.087	0.061
GFI	0.796	0.820	0.906
PGFI	0.669	0.683	0.715
ECVI	2.615	2.299	1.298

SATQ Confirmatory Factor Analysis

As shown in Table 3, three models were evaluated in the CFA. The first was the most parsimonious one factor solution. The second model extracted 3 factors following the notion of capturing the three criterion areas in the DSM-IV. The third was the model predicted by the 5 factor structure found in the EFA on the first half of the split-half sample (shown in Fig. 2). Note that in this model, 4 items had loadings on a second subscale above 0.30, and were thus allowed a path to a second subscale.

The RMSEA was between 0.08 and 0.10 in the 3 factor and 1 factor models, indicating a mediocre fit; however, the RMSEA was between 0.05 and 0.08 for the 5 factor model, indicating a better, "adequate" fit. The SRMR was below 0.10 in all three of the models indicating and were thus an acceptable fit with the 5 factor model SRMR again the lowest and indicating the strongest fit. The χ^2 of the 1 and 3 factor models were much larger than χ^2 of the 5 factor model, and the GFI, PGFI, and ECVI also all indicated that the 5 factor model had the best fit.

Thus, the results of the CFA indicated that the 5 factor model predicted by the EFA had an adequate fit or better to the data and was a better fit than the 2 other models evaluated.

Student Subsample

For the subsample of 196 individuals who were recruited from the larger student sample, even given the sampling procedure with targeted specific scale results, the 24-item SATQ mean total score for the Student Subsample was only slightly lower than the Total Student Sample mean total score (M = 23.1 vs. 22.7). The larger variability in the Student Subsample reflects the selection process for the Student Subsample wherein individuals were selected based on subdomain scores and suggests that the subsample may not be fully representative of the Total Student Sample.

Fig. 2 SATQ 5 factor model
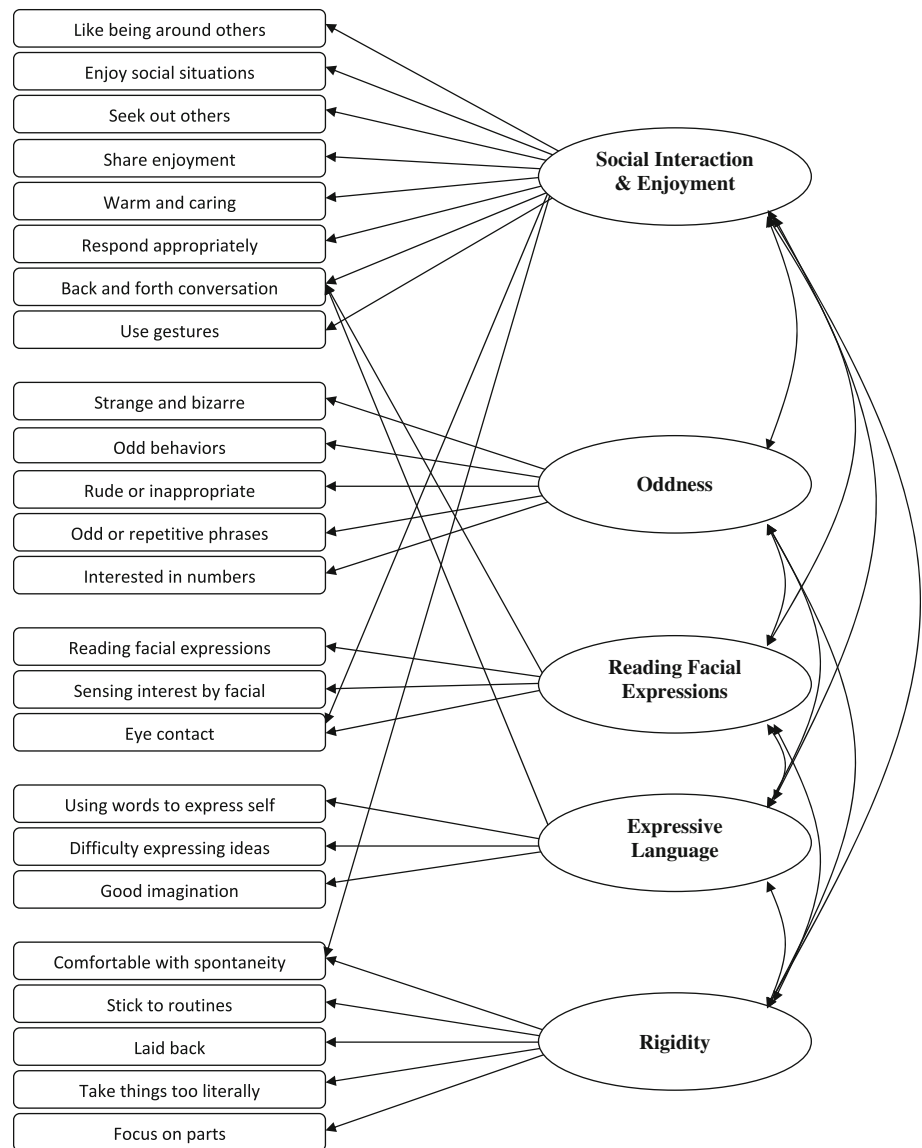
These participants concurrently completed the BAPQ and AQ questionnaires in addition to other cognitive measures as part of the separate study. Range, mean, and standard deviations for both all the measures can be found in Table 1. Mean BAPQ total score was 99.6 and mean AQ total score was 16.2 for this sample. Note that lower scores are indicative of less severe symptom presentation. One hundred and fifty nine participants who had fully completed both measures had their responses compared to the first assessment. Test–retest reliability was good, at $r = 0.79$, $p < .001$.

ASD Sample

As with the Student Subsample, ASD participants concurrently completed the BAPQ and AQ questionnaires. The mean total score for the 24-item SATQ was 40.8, mean

BAPQ total score was 147.5, and mean AQ total score was 31.7.

ASD vs. Non-ASD

Given the ordinal scale and lack of normal distributions of the BAPQ, SATQ, and AQ total scores, Spearman's rho was calculated between the three questionnaires for both groups. All three measures were highly correlated in both groups, ranging from $r = 0.73$ to $r = 0.79$ in the Student Subsample, and from $r = 0.68$ to $r = 0.83$ in the ASD sample.

To examine group differences between those with ASD and those without, a MANCOVA was conducted comparing the total scores from both the Student Subsample and ASD Sample on each measure, with Gender as a covariate given the significant difference in gender across the groups.

Dependent variables were the SATQ total score, AQ total score, and BAPQ total score with Group (Student Subsample and ASD Sample) as the fixed factor. Results indicated that the groups were significantly different from each other on each measure, with the Student Subsample consistently scoring lowest ($p < .001$ in all cases).

Given prior work showing that students with science degrees often score higher with respect to ASD symptoms (Baron-Cohen et al. 2001; Hoekstra et al. 2008; Wakabayashi et al. 2006), an ANOVA was conducted comparing SATQ total scores by general degree type. Degree types were grouped into two general categories of those associated with science (e.g., Engineering; Physical Sciences; Agriculture, Food, and Natural Resources) and those less associated with science (Business; Health Related Professions; Journalism and English; Fine Arts; Social Sciences; and Education). The groups were significantly different ($F(1, 1,535) = 19.0$, $p < .001$) with those with science associated degrees having a higher mean SATQ total score (24.5 vs. 22.7). Engineering students had the highest total (meant SATQ total score = 26.1) and those in Social Sciences had the lowest (meant SATQ total score = 21.5).

Hierarchical regression analyses were conducted for both the Student Subsample and the ASD Sample with SATQ total score serving as the dependent variable to examine the shared variability between the AQ and SATQ. AQ total score was entered in the first step of the model, followed by the BAPQ total score. By utilizing this approach, we were able to partial out variability in the SATQ related to the AQ, allowing us to examine the portion of remaining variance attributable solely to the BAPQ (i.e., partial correlation; pr^2). The notion is that, after accounting for the shared variance amongst the AQ and SATQ, is the SATQ capturing additional information found in the BAPQ that is not found in the AQ? The results of the regression analyses are summarized in Table 4. In the Student Subsample, the BAPQ uniquely accounted for 47% of the remaining variance in the SATQ total score that was not accounted for by AQ score (i.e., pr^2). Similarly, in the ASD Sample, the BAPQ uniquely accounted for 69% of the remaining variance in the SATQ total score that was not accounted for by AQ score. The results support the

contention that the SATQ is capturing information beyond that found in the AQ.

Discussion

Studying individuals who have various degrees of ASD phenotypic expression affords the opportunity to gain insight into autism as a disorder. By studying individuals with ASD traits, researchers have potential access to a larger base of participants and individuals who may be able to participate more readily in a wide range of experimental methodology. This research relies on the assumption that those who manifest ASD traits that are less severe in quality will demonstrate similar difficulties as those with more severe traits. Moreover, this research enriches our understanding of individuals with ASD traits who do not have an ASD diagnosis, which is important in its own right. This knowledge can be used to help enhance the functioning of those individuals as well as elucidate our knowledge of autism.

The current study was motivated by a need for a questionnaire that was brief, relevant to a broader population, and used a self-report format. We desired a measure that was shorter but also more comprehensive in the breadth of ASD related traits assessed. Instead of attempting to both shorten or add specific questions to an existing measure, we took a different approach: taking four oft-used and well-validated questionnaires, the SRS, the SCQ, the AQ, and the BAPQ, extracted all of the questions, listed them independently, and sorted them by perceived category. We then constructed self-report questions felt to represent the rubric reflected in the sorting. This procedure resulted in 32 questions in an initial item pool, which we term the Subthreshold Autism Traits Questionnaire, or SATQ.

This 32-question SATQ was then administered to 1,709 students. Item distribution analysis and item-total correlations resulted in the elimination of 8 questions to improve the structural validity of the measure, resulting in 24 questions. Internal consistency was good, with a Cronbach's Alpha of .73 and a test–retest of $r = .79$ (n = 187). Using a split-half analysis, an exploratory factor analysis

Table 4 Hierarchical regression analysis predicting SATQ total scores from AQ total score and BAPQ total score for student subsample and ASD sample

| | Step 1 | | | Step 2 | | | | |
| | AQ Total Score Entered | | | BAPQ Total Score Entered | | | | |
	R^2	df	ΔF	R^2	ΔR^2	pr^2	df	ΔF
Student Subsample	0.58	1,181	249.31**	0.67	0.09	0.47	1,180	51.00**
ASD Sample	0.70	1,15	37.64**	0.83	0.14	0.69	1,14	12.56*

* $p < .005$; ** $p < .001$

was performed on one half of the sample (n = 855) and 5 subscales were identified: Social Interaction and Enjoyment, Oddness, Reading Facial Expressions, Expressive Language, and Rigidity. Confirmatory factor analysis on the other half of the sample (n = 854) indicated that the 5 factor solution was an adequate fit and outperformed two other models evaluated.

The SATQ successfully differentiated between the ASD group and the student group, with the ASD group having a higher mean score. The SATQ demonstrated convergent validity with other measures that have been used to assess ASD traits, correlating strongly with both the AQ and the BAPQ in both the ASD and student group.

A direct comparison of the questions in the AQ and the SATQ is instructive as both are self-report and purport to assess subthreshold ASD traits in a general population. This informal comparison supported the contention that the AQ may focus on a specific phenotype related to Asperger Syndrome and as such may not include some questions that are important in assessing a broader range of subthreshold ASD traits. These questions are covered by other measures such as the SRS and the SCQ, but those two measures use informant based ascertainment methods. In looking at the SATQ, approximately 10 of its questions do not appear to be reflected on the AQ. These include question 1, 5, 7, 10, 19, 21–23, 31 and 32, and query aspects of expressive language, eye contact, facial expressions, odd behaviors, making gestures, empathy, sharing enjoyment, and others' perception of the individual being caring.

In fact, the data provides some empirical support that the SATQ may be capturing information that is not provided by the AQ. Using hierarchical regression analyses, after accounting for the shared variance amongst the AQ and SATQ, the BAPQ uniquely accounted for 47% of the variance in the SATQ total score in the Student Subsample after accounting for the variance related to the AQ. In the ASD Sample, the BAPQ uniquely accounted for 69% of the variance in the SATQ total score after accounting for the variance related to the AQ. The results support the contention that the SATQ is capturing information beyond that found in the AQ.

Interestingly, the distribution of SATQ total scores in the large sample was positively skewed, very similar to the distribution reported in SRS total scores on a sample of 788 twin pairs (Constantino and Todd 2003). Though the specific value of the skewness was not reported in the Constantino and Todd (2003) study, examination of the distribution clearly indicates a positive skew. One possibility of the non-normal distribution is psychometric. That is, the questions on these measures are able to more sensitively assess higher ASD traits compared to lower, contributing to the positive skew. However, the positive skew could also reflect the nature of the underlying construct of

ASD traits in that the departure from average is limited in terms of social/communicative facility, with more variability in how ASD-related difficulties can emerge (e.g., an analogy with reaction time: there is only so fast a person can respond, with no limit on how long it takes them to respond). This is an area in need of further study.

One limitation of the current study is the focus on a student sample. It may be that individuals in the general non-student population may have different responses patterns. Also, the subsample used for the comparisons with the ASD group was a sample selected based on specific SATQ results and not a random selection, again suggesting caution in the interpretation of the results. Further psychometric studies are thus needed in non-student populations, with comparison across groups with varying presenting problems and diagnoses. Another limitation was the use of non-standardized measures in the diagnosis of ASD through the community-based clinic, though the diagnoses used assessments based on such measures and were confirmed by experienced clinicians. Future studies should include rigorously phenotyped individuals so that more in-depth comparisons between the SATQ and individual differences based on those measures (e.g., the Autism Diagnostic Observation Schedule) can be completed.

The current study introduces and provides initial psychometric support for the SATQ as a brief, 24-item questionnaire that assesses broad phenotypic ASD variation and subthreshold ASD traits. Though the purpose of the measure is to assess subthreshold ASD traits, it may be found that it has some utility as a screening or diagnostic tool. If so, then further psychometric analysis can help in determining useful cutoffs. Future studies should be conducted to assess further the difference between measures such as the SATQ, the AQ, and the BAPQ. Each was constructed based upon different theoretical underpinnings and with empirical support derived from different populations. Though each clearly assesses aspects of a common construct, how they differ may be meaningful and instructive.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Publishing, Inc.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism spectrum quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders, 38*(7), 1230–1240.

Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The autism-spectrum quotient (AQ)—Adolescent version. *Journal of Autism and Developmental Disorders, 36*(3), 343–350.

Baron-Cohen, S., Scott, F. J., & Allison, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry, 194*, 500–509.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry, 175*, 444–451.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*(4), 719–726.

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics, 21*(1), 2–11.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60*(5), 524–530.

Constantino, J. N., & Todd, R. D. (2005). Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry, 57*(6), 655–660.

Delawalla, Z., Barch, D. M., Fisher Eastep, J. L., Thomason, E. S., Hanewinkel, M. J., Thompson, P. A., et al. (2006). Factors mediating cognitive deficits and psychopathology among siblings of individuals with schizophrenia. *Schizophrenia Bulletin, 32*(3), 525–537.

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Jr., Dawson, G., et al. (2000). Practice parameter: Screening and diagnosis of autism: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology, 55*(4), 468–479.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): A study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders, 38*(8), 1555–1566.

Hoekstra, R., Vinkhuyzen, A., Wheelwright, S., Bartels, M., Boomsma, D., Baron-Cohen, S., et al. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of Autism and Developmental Disorders, 41*(5), 589–596.

Hurley, R. S. E., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 37*(9), 1679–1690.

Jobe, L., & White, S. (2007). Loneliness, social relationships, and a broader autism phenotype in college students. *Personality and Individual Differences, 42*, 1479–1489.

Kanne, S. M., Christ, S. E., & Reiersen, A. M. (2009). Psychiatric symptoms and psychosocial difficulties in young adults with autistic traits. *Journal of Autism and Developmental Disorders, 39*(6), 827–833.

Kerns, J. G. (2006). Schizotypy facets, cognitive control, and emotion. *Journal of Abnormal Psychology, 115*(3), 418–427.

Krug, D. A., Arick, J. R., & Almond, P. J. (1988). *Autism behavior checklist.* Austin, TX: PRO-ED.

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1–11.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). *Autism diagnostic observation schedule.* Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview—Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685.

Psychological Corporation. (1999). *Wechsler abbreviated scale of intelligence.* San Antonio, TX: Psychological Corporation.

Reiersen, A. M., Constantino, J. N., Grimmer, M., Martin, N. G., & Todd, R. D. (2008). Evidence for shared genetic influences on self-reported ADHD and autistic symptoms in young adult Australian twins. *Twin Research and Human Genetics, 11*(6), 579–585.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131–144.

Ronald, A., Happe, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., et al. (2006). Genetic heterogeneity between the three components of the autism spectrum: A twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(6), 691–699.

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood autism rating scale (CARS). *Journal of Autism and Developmental Disorders, 10*(1), 91–103.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): Preliminary development of a UK screen for mainstream primary-school-age children. *Autism, 6*, 9–31.

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: Evidence for a continuous severity gradient. *American Journal of Medical Genetics, 114*(2), 129–136.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. Apr 1990. *Multivariate Behavioral Research, 25*(2), 173–180.

Volkmar, F., Cook, E. H., Jr., Pomeroy, J., Realmuto, G., Tanguay, P., American Academy of, C., et al. (1999). Practice parameters for the assessment and treatment of children, adolescents, and adults with autism and other pervasive developmental disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(12 Suppl), 32S–54S.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The autism-spectrum quotient (AQ) in Japan: A cross-cultural comparison. *Journal of Autism and Developmental Disorders, 36*(2), 263–270.

Wechsler, D. (1999). *The Wechsler abbreviated scale of intelligence.* San Antonio, TX: The Psychological Corporation.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432–442.