John Carroll University

Carroll Collected

2023

The Autism Symptom Dimensions Questionnaire: Development and psychometric evaluation of a new, open-source measure of autism symptomatology

Thomas Frazier

Anastasia Dimitropoulos

Leonard Abbeduto

Melissa Armstrong-Brine

Shanna Kralovic

See next page for additional authors

Authors

Thomas Frazier, Anastasia Dimitropoulos, Leonard Abbeduto, Melissa Armstrong-Brine, Shanna Kralovic, Andy Shih, Antonio Y. Hardan, Eric A. Youngstrom, and Mirko Uljarević

The Autism Symptom Dimensions Questionnaire: Development and psychometric evaluation of a new, open-source measure of autism symptomatology

Thomas W. Frazier[1,2] 🔆 | Anastasia Dimitropoulos[3] | Leonard Abbeduto[4] |
Melissa Armstrong-Brine[5] | Shanna Kralovic[6] | Andy Shih[2] | Antonio Y. Hardan[7] |
Eric A. Youngstrom[8,9] | Mirko Uljarević[10] | Quadrant Biosciences - As You Are Team[1]

[1]Department of Psychology, John Carroll University, University Heights, OH, USA

[2]Autism Speaks, Princeton, NJ, USA

[3]Department of Psychological Sciences, Case Western Reserve University, Cleveland, OH, USA

[4]Department of Psychiatry and Behavioral Sciences and MIND Institute, University of California, Davis, Davis, CA, USA

[5]MetroHealth Autism Assessment Clinic, MetroHealth Medical Center, Cleveland, OH, USA

[6]Division of Developmental Behavioral Pediatrics and Psychology, University Hospitals Case Medical Center, Cleveland, OH, USA

[7]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA

[8]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, NC, USA

[9]Helping Give Away Psychological Science, Chapel Hill, NC, USA

[10]Melbourne School of Psychological Sciences, Faculty of Medicine, Dentistry, and Health Sciences, The University of Melbourne, Victoria, Australia

Correspondence
Thomas W. Frazier, Department of Psychology, John Carroll University, 1 John Carroll Boulevard, University Heights, OH 44118, USA.
Email: tfrazier@jcu.edu

Funding information
Hartwell Foundation; Autism Speaks; the Hartwell Foundation; Research Foundation; Research Foundation; Autism Speaks, Grant/Award Number: N/A; F. Hoffmann-La Roche AG; Simons Foundation; Forest Laboratories; Bristol Myers Squibb; National Institutes of Health; Brain and Behavior Research Foundation; American Psychological Association

Abstract

Aim: To describe the development and initial psychometric evaluation of a new, freely available measure, the Autism Symptom Dimensions Questionnaire (ASDQ).

Method: After development and revision of an initial 33-item version, informants completed a revised 39-item version of the ASDQ on 1467 children and adolescents (aged 2–17 years), including 104 with autism spectrum disorder (ASD).

Results: The initial 33-item version of the ASDQ had good reliability and construct validity. However, only four specific symptom factors were identified, potentially due to an insufficient number of items. Factor analyses of the expanded instrument identified a general ASD factor and nine specific symptom factors with good measurement invariance across demographic groups. Scales showed good-to-excellent overall and conditional reliability. Exploratory analyses of predictive validity for ASD versus neurotypical and other developmental disability diagnoses indicated good accuracy for population and at-risk contexts.

Interpretation: The ASDQ is a free and psychometrically sound informant report instrument with good reliability of measurement across a continuous range of scores and preliminary evidence of predictive validity. The measure may be a useful alternative to existing autism symptom measures but further studies with comparison of clinical diagnoses using criterion-standard instruments are needed.

Autism spectrum disorder (ASD) is a heterogeneous set of conditions characterized by qualitative impairment in social interaction and communication and the presence of restricted and repetitive behaviors and interests.[1] Comprehensive and reliable measurement of autism symptoms is crucial for early identification, entry into clinical services, and ongoing longitudinal tracking, including evaluation of response to intervention. Informant report measures are also a key component of longitudinal tracking in many clinical and research settings.[2–4] However, the available informant report instruments measuring core autism symptoms are either commercial measures (e.g. Social Responsiveness Scale, Second Edition; Autism Impact Measure),[2,5] intended for screening within a narrow age band (e.g. Modified Checklist for Autism in Toddlers),[6] or focus only on one symptom domain (e.g. Repetitive Behavior Scale-Revised).[7] Commercial measures can add significant expense in clinical and research applications and are therefore ill-suited for frequent longitudinal tracking or the development of large sample cohorts. This problem specifically disadvantages research and clinical practice in low- and middle-income countries.[8] Measures such as the Autism Spectrum Quotient,[3] which is only applicable to cognitively able individuals, are difficult to implement in practice settings in which pre-appointment completion of informant measures is desired and estimates of the cognitive level of the child are not available. Finally, many existing informant-completed measures often have unclear construct validity,[9,10] require responses to a large numbers of items, which can be burdensome for parent caregivers (e.g. the Pervasive Developmental Disorder Behavior Inventory has 124 items in the standard set), use dichotomous responses (e.g. Social Communication Questionnaire [SCQ]), or were not expressly designed to assess Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) symptom criteria, which results in limited coverage of sensory symptoms.

This article describes the development and psychometric evaluation of a new open-source, freely available informant report measure, the Autism Symptom Dimensions Questionnaire (ASDQ). The ASDQ was created to address the limitations of existing measures, covering core autism symptoms in children and adolescents based on the DSM-5 symptom criteria. Initial work focused on the development and factor structure of a 33-item version. Using results from this initial work, the ASDQ was expanded to include 39 items, with the goal of identifying specific social communication/interaction and restricted/repetitive behavior factors corresponding to DSM-5 criteria and recent factor analyses of other measures.[10–12] Using the revised 39-item ASDQ version, the present psychometric evaluation aimed to (1) examine factor structure and (2) evaluate the psychometric properties of the revised scales, including measurement invariance, overall and conditional reliability, and convergent and discriminant validity. These analyses also explored the screening efficiency of the revised ASDQ compared with parent-reported ASD diagnoses, to

---

What this paper adds

- The Autism Symptom Dimensions Questionnaire (ASDQ) is a new, freely available measure of autism symptoms.
- The ASDQ showed reliable and accurate measurement of autism symptoms.
- The measure had good screening efficiency for autism spectrum disorder relative to other developmental conditions.

---

determine the potential value of the measure in screening and diagnostic contexts.

METHOD

Initial development

Initial development of the ASDQ focused on a 33-item version (see Appendix S1 for an extensive description of this work). The development process found that the 33-item measure had good scale reliability and strong convergent validity with an established autism symptom measure, but that additional work was needed to update the factor structure and further examine key psychometric characteristics. Based on the results of this work, the 33-item ASDQ was revised by adding six items to identify separate factors corresponding to nine autism symptom dimensions identified in previous analyses of other screening measures,[10,11,13,14] including those shown in factor analyses to yield separate factors within the DSM-5 criteria. Two of the newly added items assessed relationships (items 15 and 17), one evaluated sensory sensitivity (item 27), two focused on sensory interests (items 29 and 31), and one examined restricted interests (item 34) (Appendix S2). Psychometric evaluation focused on this revised instrument. Based on our initial factor analyses and specific item content additions, a nine symptom factor plus a general bifactor model was hypothesized a priori to provide optimal fit.

Participants

Informants were recruited using the Prolific online data collection service (https://prolific.co/). Inclusion criteria for the Prolific panel included: residence in the USA; having a dependent child aged 2 to 17 years; and informant proficiency in English. Adults were not included in this study because the intention was to develop and validate a measure for children and adolescents and a narrower age range was desired for evaluating measurement invariance across age. Data were collected from 23rd June 2021 to 7th July 2021. Additional methodological details are included in Appendix S3.

Measures

Demographic and health Information

Parent informants provided their age, sex, relationship to the child, household income, US region of residence, and highest level of education. Parents also reported the child's age, sex, race (race was used following National Institutes of Health definitions to provide a more comprehensive evaluation of possible measurement differences across specific populations), ethnicity, level of speech/language, and all previous clinical diagnoses.

ASDQ

The 39-item ASDQ is available online in Microsoft Word and PDF file formats (https://osf.io/pyq8r/). The Flesch–Kincaid reading level for the final measure was US grade 7.9. All items were rated using a 5-point Likert frequency scale (1, never; 2, rarely; 3, sometimes; 4, often; 5, very often). See Appendix S2 for additional details.

The Stanford Social Dimensions Scale

The Stanford Social Dimensions Scale[14] was collected to evaluate the convergent validity of the ASDQ using a measure of social communication/interaction behavior. The Stanford Social Dimensions Scale is a 71-item informant report questionnaire that captures different aspects of social functioning. Items are rated on a 5-point Likert scale using the following choices: never; rarely; sometimes; often; and always. For the present study, a total score was used.

SCQ

The SCQ is a dichotomously keyed (yes/no) 40-item rating scale based on DSM, Fourth Edition, Text Revision symptoms. Lifetime ratings referenced the child's behavior throughout their developmental history, increasing diagnostic validity.[15] For the present study, a subsample of cases from our preliminary work using the Hartwell Foundation's KidsFirst cohort (n = 175) were administered the revised 39-item ASDQ and the SCQ. This was done to evaluate convergent validity.

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire[16] is a 25-item informant report measure appropriate for use in children aged 2 to 17 years. Items are rated using a 3-point Likert scale. The Strengths and Difficulties Questionnaire was expected to demonstrate moderate convergent validity based on the observation that many individuals with ASD have significant internalizing or externalizing behavior problems. Total raw scores were evaluated.

Executive and adaptive functioning measures

Two additional measures were collected, which evaluate daily living skills and executive functioning, that is, the Daily Living Skills and the Executive Functioning Scale[17] respectively. The Daily Living Skills is based on review of existing adaptive function measures. It includes 89 items that evaluate multiple daily living content areas and is strongly correlated with the Vineland Daily Living Skills Scale (r = 0.93). The Executive Functioning Scale was based on review of existing informant report executive functioning measures and the larger literature on executive functioning. The measure includes 70 items assessing five domains of executive functioning and the total score shows a strong correlation with the 24-item Behavior Rating Inventory of Executive Functioning (r = 0.85). Total item average scores were used for both measures.

Procedure

After completing informed consent electronically, caregiver informants completed demographic information and all questionnaires for the identified child. Compensation was provided after study completion based on the average expected completion time (US$10).

Statistical analyses

Factor structure

To identify the factor structure of the revised 39-item ASDQ, the sample was randomly divided into approximately equal exploratory (n = 734) and confirmatory (n = 733) subsamples. A series of exploratory structural equation models (ESEMs) were estimated.[18] The first model was a single-factor model. The second model included four specific factors and a general ASD bifactor, consistent with the initial analyses of the 33-item ASDQ. The next three models were 8-, 9-, and 10-factor ESEMs corresponding to variations in the a priori hypothesized model based on item content. The nine-factor model (model 9a) included nine specific factors, one for each item content area and a general ASD bifactor and was hypothesized a priori based on the content of items added to the revised instrument. Once the best-fitting ESEM was identified in the exploratory subsample and replicated in the confirmatory subsample, an equivalent confirmatory factor analysis (CFA) model for ordinal data[19] (model 9b) was estimated in both subsamples and the total sample. For this CFA model, item-factor correspondence was determined by assigning each item to the factor with its largest loading. Items 36 to 39 were only assigned to the general

ASD factor because these items did not have a clear loading pattern in our preliminary analyses of the 33-item version; these items emphasize symptoms that are most relevant for older children with average or above cognitive ability and they have the highest rates of nonresponse. Models were estimated and fitted using standard structural model fit statistics and consideration of model interpretability and parsimony (Appendix S4).

Measurement invariance

The CFA model derived from the best-fitting ESEM was used as the basis for the evaluation of measurement invariance across age groups (ages 2–4 years, 5–7 years, 8–12 years, and 13–17 years), sex (male, female), race (White, other), and ethnicity (Hispanic, non-Hispanic). Standard procedures were followed to evaluate fitting across successively restrictive models (see Appendix S4 for additional information).[20]

Reliability

Using the optimal factor solution, classical test theory reliability (internal consistency),[21] factor reliability,[22] and item response, theory analyses[23] were conducted for each scale (Appendix S4). Reliability estimates falling in the ranges 0.70 to 0.79, 0.80 to 0.89, and greater than 0.90 were considered fair, good, and excellent.[24] Test–retest reliability and sensitivity to change data were not available.

Convergent and discriminant validity

Convergent and discriminant validity was computed using bivariate correlations. Specific predictions for strong and weak convergent and discriminant validity are presented in Appendix S4.

Preliminary screening efficiency

To explore possible screening efficiency, the sample was split into training (approximately 60% of the sample, $n = 880$) and testing (approximately 40% of the sample, $n = 587$) subsamples;[25] two criteria were evaluated: ASD versus neurotypical plus other developmental disability (total sample) and ASD versus other developmental disability (at-risk-only sample). The first criterion estimates potential screening accuracy in primary care or population settings and the second criterion estimates potential screening accuracy in tertiary care settings where only at-risk cases are evaluated. These criteria were derived from informant-reported diagnoses, which showed very good correspondence (80.0% overall accuracy) with classifications based on an SCQ cut score of 12, which has previously been shown to have good sensitivity and specificity.[26]

To estimate potential screening value, nonparametric estimates of the area under the curve from receiver operating characteristic analyses quantified screening efficiency using the ASDQ total item average score as the predictor. The predictive validity of the ASDQ total item average scores was also evaluated by identifying the optimal cut score in the exploratory subsample using Youden's J. The cut scores were applied to the testing subsample to determine whether validity was maintained in a holdout sample. The total sample cut score was also examined in the relevant subgroups (ages 2–4 years, female, other race, and Hispanic ethnicity) to examine whether predictive accuracy was maintained.

A rough guideline for evaluating area under the curve values is: less than 0.60, poor; 0.60 to 0.69, fair; 0.70 to 0.79, good; 0.80 to 0.89, excellent if the comparison group is clinically meaningful; and 0.90 to 1.00, exceptional only if the design and comparison are appropriate.[27]

Statistical power

The exploratory and confirmatory sample sizes ($n = 734$ and 733) were expected to be adequate for factor analyses with up to 39 indicators (>18 cases per indicator)[28] and for item response theory analyses of the total score and subscales.[29] For convergent and discriminant validity analyses, the total sample was expected to have excellent power ($1–\beta > 0.96$) for detecting at least small Pearson bivariate correlations ($r \geq 0.10$, $\alpha = 0.05$, two-tailed) and statistical power was expected to be excellent ($1–\beta > 0.94$) for detecting differences in dependent correlations as small as $\Delta r = 0.09$. Analyses estimating screening efficiency had excellent power ($1–\beta \geq 0.95$) for detecting areas under the curve ≥ 0.660.

Statistical significance was set at $\alpha = 0.05$, two-tailed; effect size magnitude was emphasized. Data preparation, descriptive analyses, internal consistency reliability used Cronbach's α, model reliability used MacDonald's omega (ω), and bivariate correlations used SPSS v28 (IBM Corp., Armonk, NY, USA).[30] Factor, measurement invariance, and item response analyses were computed in MPlus v8.7.[31] Receiver operating characteristic analyses were computed using the R package pROC. All R packages (R Foundation for Statistical Computing, Vienna, Austria) and procedures were implemented in v4.1.2,[32] using RStudio v2021.09.1 (Posit, Boston, MA, USA).

RESULTS

Participant characteristics

The sample included 1467 children (aged 2–17 years; Table 1). Biological mothers predominated as informants. Informants were older and slightly less educated in the groups with developmental disability and ASD. Household income was lower in the group with developmental disability and lowest in the group with ASD relative to neurotypical children. Children were older in the groups with developmental disability and

THE AUTISM SYMPTOM DIMENSIONS QUESTIONNAIRE: DEVELOPMENT AND PSYCHOMETRIC
EVALUATION OF A NEW, OPEN-SOURCE MEASURE OF AUTISM SYMPTOMATOLOGY

5

TABLE 1 Demographic and clinical characteristics across ASD, developmental disability, and neurotypical controls in the revised 39-item ASDQ sample

	Neurotypical	Non-ASD developmental disability	ASD	χ^2/F (p)
	$n = 1012$	$n = 351$	$n = 104$	
Informant				58.5 (<0.001)
Biological mother	548 (54.2)	237 (67.5)	80 (76.9)	
Biological father	434 (42.9)	97 (27.6)	19 (18.3)	
Other/not reported	30 (2.9)	17 (4.9)	5 (4.8)	
Informant age in years:months, mean (SD)	37:6 (7:11)	39:8 (8:6)	40:6 (7:2)	13.7 (<0.001)
Highest parental education				26.8 (0.003)
Less than high school	6 (0.6)	2 (0.6)	1 (1.0)	
High school or general equivalency diploma	90 (8.9)	38 (10.8)	11 (10.6)	
Some college	178 (17.6)	94 (26.8)	32 (30.8)	
College graduate	427 (42.2)	132 (37.6)	40 (38.5)	
Graduate degree or higher	295 (29.2)	80 (22.8)	18 (17.3)	
Unknown	16 (1.6)	5 (1.4)	2 (1.9)	
US region				10.9 (0.205)
Northeast	188 (18.6)	51 (14.5)	16 (15.4)	
Midwest	215 (21.3)	69 (19.7)	23 (22.1)	
South	402 (39.8)	168 (47.9)	50 (48.1)	
West	203 (20.1)	62 (17.7)	15 (14.4)	
Other/chose not to provide	4 (0.4)	1 (0.3)	0 (0)	
Household income				43.1 (<0.001)
<US$25 000	70 (6.9)	36 (10.3)	17 (16.3)	
US$25 000–34 999	65 (6.4)	31 (8.8)	15 (14.4)	
US$35 000–49 999	105 (10.4)	50 (14.4)	14 (13.5)	
US$50 000–74 999	224 (22.1)	80 (22.8)	23 (22.1)	
US$75 000–99 999	178 (17.6)	55 (15.7)	14 (13.5)	
US$100 000–149 999	238 (23.5)	58 (16.5)	12 (11.5)	
US$150 000–199 999	80 (7.9)	29 (8.3)	7 (6.7)	
≥US$200 000	48 (4.7)	9 (2.6)	2 (1.9)	
Unknown	4 (0.4)	3 (0.9)	0 (0)	
Child age in years:months, mean (SD)	8:1 (4:7)	10:11 (4:7)	9:10 (4:8)	49.4 (<0.001)
2–4 years	315 (31.1)	47 (13.4)	21 (20.2)	84.0 (<0.001)
5–8 years	322 (32.0)	89 (25.4)	30 (28.8)	
9–11 years	134 (13.3)	53 (15.1)	14 (13.5)	
12–17 years	236 (23.4)	161 (46.0)	39 (37.5)	
Child biological sex (male)	546 (54.0)	194 (55.3)	79 (76.0)	21.6 (<0.001)
Ethnicity				
White	849 (83.9)	311 (88.6)	87 (83.7)	4.7 (0.096)
Black/African American	124 (12.3)	36 (10.3)	20 (19.2)	6.0 (0.050)
Middle Eastern	5 (0.5)	3 (0.9)	0 (0)	1.2 (0.538)
East Asian	43 (4.2)	15 (4.3)	2 (1.9)	1.3 (0.512)
South Asian	28 (2.8)	4 (1.1)	0 (0)	5.7 (0.057)
Pacific Islander	7 (0.7)	4 (1.1)	0 (0)	1.5 (0.461)

(Continues)

T A B L E 1 (Continued)

	Neurotypical	Non-ASD developmental disability	ASD	χ^2/F (p)
	$n = 1012$	$n = 351$	$n = 104$	
Native American	18 (1.8)	8 (2.3)	1 (1.0)	0.8 (0.657)
Unknown ethnicity	2 (0.2)	1 (0.3)	0 (0)	0.3 (0.849)
Chose not to provide	12 (1.2)	3 (0.9)	1 (1.0)	0.3 (0.868)
Hispanic or Latino	117 (11.6)	40 (11.4)	11 (10.6)	1.5 (0.824)
Level of speech				116.5 (<0.001)
Nonspeaking	3 (0.3)	5 (1.4)	10 (9.7)	
Minimal speech	94 (9.3)	31 (8.8)	31 (30.1)	
Fluent speech	912 (90.4)	315 (89.7)	62 (60.2)	
Non-ASD diagnoses				
Intellectual disability/GDD	–	10 (2.8)	6 (5.8)	2.1 (0.150)
Speech/language disorder	–	75 (21.4)	16 (15.5)	1.7 (0.193)
ADHD	–	146 (41.6)	29 (27.9)	6.1 (0.014)
ODD/conduct disorder	–	25 (7.1)	5 (4.9)	0.7 (0.415)
Anxiety disorder	–	111 (31.6)	19 (18.4)	6.8 (0.009)
Specific learning disorder	–	33 (9.4)	3 (2.9)	4.6 (0.032)
Motor/coordination disorder	–	16 (4.6)	2 (1.9)	1.4 (0.231)
Depressive disorder	–	50 (14.2)	8 (1.8)	3.0 (0.083)
Bipolar disorder/mania	–	7 (2.0)	1 (1.0)	0.5 (0.488)
Obsessive compulsive disorder	–	11 (3.1)	5 (4.9)	0.7 (0.405)
Tic disorder	–	6 (1.7)	1 (1.0)	0.3 (0.593)
Feeding/eating disorder	–	16 (4.6)	0 (0)	4.9 (0.029)

Data are n (%) unless otherwise stated. Non-ASD diagnoses do not sum to 100% because children could be diagnosed with more than one condition. Cognitive level information was only completed for $n = 886$. Continuous variables (parent informant and child age) were tested using univariate analysis of variance using the F statistic and associated statistical significance (p). Abbreviations: ASD, autism spectrum disorder; ADHD, attention-deficit/hyperactivity disorder; GDD, global developmental delay; ODD, oppositional defiant disorder.

ASD and, consistent with ASD prevalence, the sex ratio was approximately 3:1 in the group with ASD but more balanced in the neurotypical and developmental disability groups. The sample was predominantly White (85%), but Black (12.0%) and Hispanic (11.5%) groups were well represented. The most frequently reported non-ASD diagnoses in the groups with ASD and developmental disability were speech/language disorder, attention-deficit/hyperactivity disorder (ADHD), and anxiety. Depression and specific learning disability were also common in the group with developmental disability. Details on the age and language level distribution and missing data handling are provided in Appendix S5.

Factor structure

In the exploratory and confirmatory samples, optimal fit was identified for the a priori hypothesized nine specific symptom factor plus a general bifactor ESEM solution (Table S1). Specific factors were readily interpreted as basic social communication (seven items), affiliation (three items), perspective taking (four items), peer relationships (three items), repetitive motor behavior (four items), sensory interests (three items), insistence on sameness (four items), sensory sensitivity (three items), and restricted interests (four items) (Table 2 and Figure S1). For this model, the nine specific symptom factors had primary loadings that were consistent with the model hypothesized a priori (Table S2). All items had significant loadings on the general ASD factor. Factor correlations indicated small-to-medium residual relationships, with the strongest relationships being between repetitive motor behavior and sensory interests ($r = 0.59$) and between insistence on sameness and sensory sensitivity ($r = 0.48$). Correlations based on scale scores followed the same pattern but were larger in magnitude because these reflect relationships without partitioning variance due to the general ASD factor (Table S3).

The CFA model (model 9b) fitted adequately in spite of not estimating minor item cross-loadings (Table S1). This model was used to evaluate measurement invariance, model reliability, and variance accounted for by the specific and general ASD factors. Using this model, the general ASD factor explained 42% of the common variance, whereas the specific factors accounted for smaller amounts of common variance (basic social communication = 6%, affiliation = 6%, perspective talking = 2%, peer relationships = 5%,

THE AUTISM SYMPTOM DIMENSIONS QUESTIONNAIRE: DEVELOPMENT AND PSYCHOMETRIC
EVALUATION OF A NEW, OPEN-SOURCE MEASURE OF AUTISM SYMPTOMATOLOGY

7

TABLE 2 ASDQ factor-item mapping and relationship to the DSM-5 criteria

Factor	Equivalent DSM-5 symptom criteria	Items and descriptions
Social communication and interaction domain		
Basic social communication	A2: nonverbal communication[a]	1. Start interactions 4. Eye contact 5. Use gestures 6. Communicate clearly 8. Share enjoyment 10. Respond appropriately 11. Converse about others
Affiliation[b]	Mixed between A1 and A3 A1 – initiate or respond to social interactions A3 – absence of interest in peers	2. Be with family/friends 3. Stay connected 15. Relationships are important
Perspective taking	A1: Reciprocity[c]	7. Offer comfort 12. Read social cues 13. Expected behavior by situation 14. Understand others
Peer relationships	A3: Relationships[d]	9. Peer play 16. Seek out interactions 17. Close friends
Restricted and repetitive behaviors domain[e]		
Repetitive motor behavior	B1: Repetitive motor, object use, and speech	18. Flap hands 19. Repetitive body movements 20. Repeat sounds or words 21. Play repetitively
Sensory interests	B4: Sensory sensitivities and interests	29. Engage with parts of objects 30. Fascinated by sensory 31. Preoccupied with patterns
Insistence on sameness	B2: Insistence on sameness	22. Consistent schedule 23. Difficulty transitioning 24. Follow strict rules or rituals 25. Trouble being flexible
Sensory sensitivities	B4: Sensory sensitivities and interests	26. Sensitive to noise 27. Upset in crowds 28. Dislike certain stimuli
Restricted interests	B3: Restricted, fixated interests	32. Think about same topic 33. Fixated on one activity 34. Too much time on interest 35. Narrow interest

[a]The ASDQ factor basic social communication includes both verbal and nonverbal elements of communication but is consistent with previous factor analyses suggesting that this DSM-5 symptom criterion (A2) may be too narrow.

[b]Affiliation item content is consistent with previous factor analyses identifying a social motivation factor.

[c]The perspective taking factor includes some overlap in content with the DSM-5 A1 reciprocity criterion, such as the back and forth of interaction/offering comfort, but also focuses more on the underlying cognitive problem that likely drives problems with reciprocity.

[d]Peer relationships are more narrow than the DSM-5 A3 relationships criterion by focusing mostly on peer relationships and interactions, although item 16 also includes content regarding seeking any playful interactions.

[e]ASDQ restricted/repetitive behavior factors tend to show strong correspondence with the DSM-5 restricted/repetitive behavior symptom criteria with the notable exception that the DSM-5 B4 is split into the two types of sensory symptoms.

Abbreviations: ASDQ, Autism Symptom Dimensions Questionnaire; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

repetitive motor behavior = 9%, sensory interests = 7%, insistence on sameness = 7%, sensory sensitivity = 6%, restricted interests = 10%).

Measurement invariance

The nine specific symptom factors hypothesized a priori with a general bifactor CFA model[33] showed evidence of measurement invariance of factor loadings, thresholds, and

residual variance (strict invariance) across sex, age, race, and ethnicity (Table S4).

Reliability

Model reliability was high for the general ASD factor ($\omega = 0.97$) and specific factors ($\omega \geq 0.85$). Internal consistency reliability was excellent for the total scale ($\alpha = 0.95$) and at least adequate for all subscale scores ($\alpha \geq 0.75$; Table S5). Conditional reliability

estimates indicated excellent reliability (≥0.90) for the total ASD scale from very low (θ [trait level] = −2.1) to extremely high (θ = +4.0) scores. Adequate or better reliability (≥0.70) was present for subscale scores in the range from low average (θ = −1.0) to very high scores (θ = +2.6) (Figure 1). Theta scores showed strong correlations with ASDQ total scores (r = 0.97) and subscales (r = 0.92–0.98).

The revised 39-item ASDQ showed a relatively normal score distribution (Figure S2). Coupled with strong reliability through very low scores, this suggests that the ASDQ total score adequately reflects continuous variation in autism symptom levels through much of the general population.

Convergent and discriminant validity

ASDQ total scores showed strong correlations with SCQ and Strengths and Difficulties Questionnaire total scores and with reported ASD diagnosis (Table 3; see Table S6 for the SCQ sample characteristics). Moderate correlations (r = 0.39–0.50) were observed between the ASDQ and Strengths and Difficulties Questionnaire internalizing and externalizing, executive functioning, and adaptive functioning measures. The expected patterns for convergent and discriminant validity were observed.

Evaluation of potential convergent and discriminant validity of the ASDQ and SCQ subscales was largely consistent with expectation (Table S7). Specifically, correlations between similar SCQ-ASDQ scale pairs (in bold) were significantly higher (r = −0.58) than correlations between dissimilar scale pairs (r = −0.44; $t_{(6)}$ = 3.93, p = 0.008).

Potential screening accuracy

Potential screening accuracy was very good (89.6%–94.9%; Table S8) with very-good-to-excellent area under the curve = 0.927 in all cases and 0.876 in at-risk cases (Figure 2). The optimal cutoff point (Youden's J) was 2.7 for screening in a primary care or population context (sensitivity = 0.88, specificity = 0.91) and 2.9 for screening in a tertiary care setting (sensitivity = 0.83, specificity = 0.89) accuracy. Scores greater than 3.2 had high specificity, occurring in nearly half of cases with ASD (sensitivity = 55.8%) but in only 5.4% of participants with developmental disability and 0.3% of neurotypical participants (specificity = 98.4%; likelihood ratio = 34.9). Accuracy levels remained high across ages 2 to 4 years, female, other race, and Hispanic subgroups, with some variation across subgroups (87.3%–92.9%). Potential screening accuracy in the at-risk sample was also good (80.2%).

DISCUSSION

The ASDQ is a brief (39-item), informant report measure of core autism symptoms informed by the DSM-5 symptom criteria and recent factor analyses of autism symptom measures. Initial work developed a 33-item version with good reliability and construct validity. However, coverage of specific autism symptom constructs was limited. Based on this, six items were added to improve construct coverage and a comprehensive psychometric evaluation of the revised instrument was conducted. This evaluation focused on factor structure, measurement invariance, classical test theory
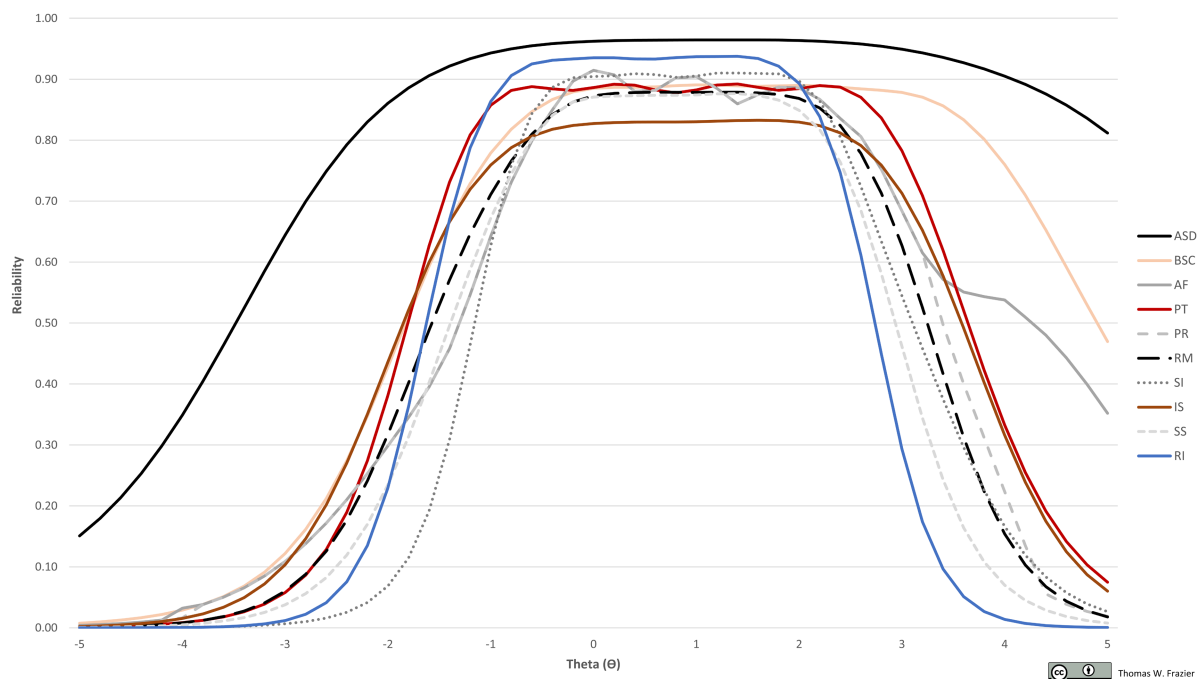
FIGURE 1 Item response theory-derived conditional reliability across the latent trait for the total ASD scale and subscales. Abbreviation: ASD, autism spectrum disorder; BSC, basic social communication; AF, affiliation; PT, perspective taking; PR, peer relationships; RM, repetitive motor behavior; SI, sensory interests; SS, sensory sensitivities; RI, restricted interests.

TABLE 3 Convergent and discriminant validity of the revised 39-item ASDQ total item average score

	Revised 39-item ASDQ	
	r	p
SCQ total	0.85	<0.001
Stanford Social Dimensions Scale	−0.53	<0.001
ASD diagnosis	0.62	<0.001
Executive functioning	0.47	<0.001
Strengths and Difficulties Questionnaire-internalizing	0.46	<0.001
Strengths and Difficulties Questionnaire-externalizing	0.50	<0.001
Daily living skills	−0.39	<0.001
Sex	0.12	<0.001
Intellectual disability/global delay	0.13	<0.001
Speech/language disorder	0.16	<0.001
Attention-deficit/hyperactivity disorder	0.22	<0.001
Anxiety disorder	0.16	<0.001
Obsessive compulsive disorder	0.12	<0.001
Age	−0.10	<0.001
Sex (ASD subsample only)	0.04	0.655
Race: White vs other	0.08	0.002
Ethnicity: Hispanic vs other	0.04	0.161
Oppositional defiant/conduct disorder	0.14	<0.001
Specific learning disorder	0.05	0.040
Motor/coordination disorder	0.07	0.005
Depressive disorder	0.07	0.004
Bipolar disorder/mania	0.08	0.003
Tic disorder	0.04	0.127
Feeding/eating disorder	0.06	0.026

Abbreviations: ASD, autism spectrum disorder; ASDQ, Autism Symptom Dimensions Questionnaire; SCQ, Social Communication Questionnaire.

scale reliability and item response theory conditional reliability, convergent and discriminant validity, and exploration of potential screening validity.

Results indicated that the 39-item ASDQ is a psychometrically sound instrument, with evidence of several positive psychometric properties, suggesting it is a highly promising measure for use in both research and clinical practice. Specifically, the ASDQ had a clear and replicable factor structure, which is consistent with previous multi-instrument analyses;[10,11] good measurement equivalence across age, sex, race, and ethnicity; convergent validity with established measures of autism symptoms and traits; and discriminant validity evidence among ASDQ subscales and with measures of related but distinct constructs and other developmental diagnoses. The ASDQ also showed preliminary evidence of utility in informing the screening process. If replicated in future work, the ASDQ would represent a major advance for the measurement of autism symptoms because currently open-source informant report measures of comparable coverage in terms of ages and cognitive levels

are not freely available. However, it is crucial that future replication studies overrecruit at younger ages and include a sufficient sample of individuals with cognitive and speech/language impairment to ensure that factor structure and screening efficiency are maintained in these subgroups.

The revised 39-item ASDQ had a highly differentiated factor structure that replicated well across exploratory and confirmatory subsamples. The final model included four social communication/interaction symptom factors, five restricted and repetitive behavior symptom factors, and a general ASD factor. These factors represent specific types of social communication/interaction and restricted and repetitive behavior symptoms and these distinctions have been found in previous research.[10,11,34,35]

To gain greater understanding of the well-recognized variety and heterogeneity of autism presentations, it is crucial to have a measure that can capture these subdomains. For example, using the ASDQ, future etiological studies will be better equipped to identify which etiological factors relate to specific core ASD symptoms. Having nine specific symptom factors that correspond to the DSM-5 criteria (with affiliation and peer relationships assessing A3 and sensory interests and sensory sensitivities evaluating B4) is also useful for helping clinicians understand whether a child might be exhibiting symptoms that meet the diagnostic threshold. Fine-grained measurement of specific symptom areas may also facilitate identification of symptom subgroups, enable more detailed longitudinal tracking, and provide useful secondary outcomes for clinical trials (see Appendix S6 for additional discussion of the ASDQ findings), although it will be important to examine factor structure in more discrete ages and functional levels to avoid inflation of inter-item correlations that may occur in very broad samples.

Convergent and discriminant validity evidence indicates that the ASDQ is measuring core autism symptoms, showing lower correlations with measures of other psychopathologies. This finding is important because it suggests that ASDQ scores are only showing the anticipated influences from other forms of psychopathology, an expected finding given the level of comorbidity seen in ASD.[36] Future research is needed with other autism measures and with measures of other types of psychopathology and clinical factors (e.g. IQ, language) to better understand ASDQ measurement specificity, determine whether adjustment for other factors will produce a more distinct measure of autism symptoms, and further explore convergent and discriminant validity.

Although an exploratory aspect of this study, the present results suggest that the ASDQ has good potential to inform screening. Accuracy was maintained across relevant subsamples, including young children (aged 2–4 years), females, other (non-White) races, and Hispanic ethnicity. Intriguingly, the ASDQ item average also performed comparably to statistical learning methods, producing roughly equivalent accuracy in most analyses. Although additional work is needed in larger and more diverse samples of young children, including children who are referred but not yet diagnosed, these findings suggest that the ASDQ total item
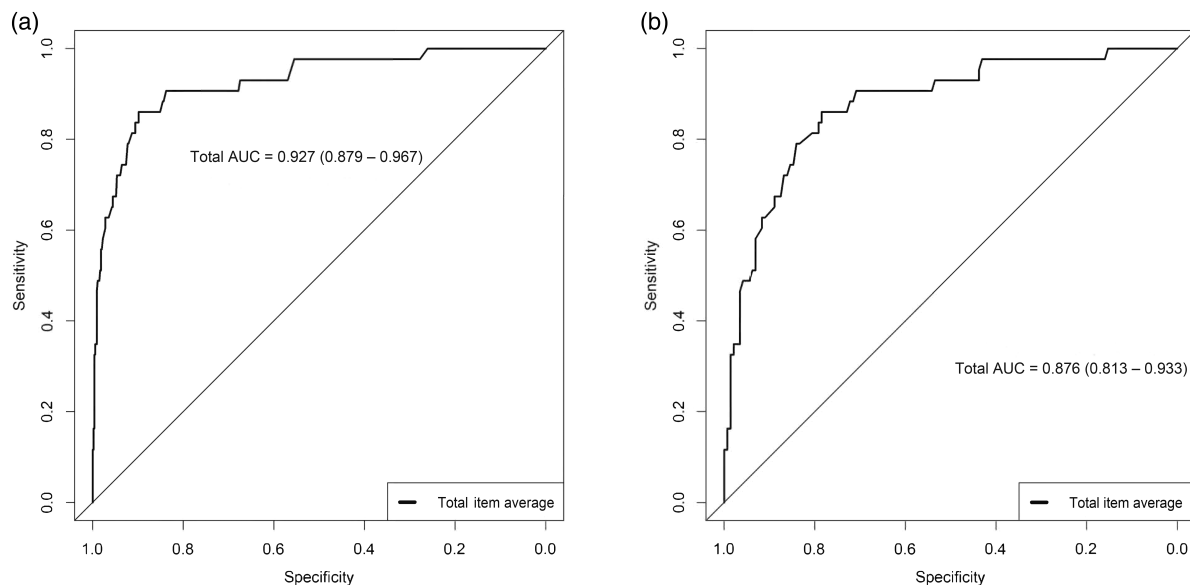
FIGURE 2 Receiver operating characteristic curves depicting the prediction of ASD diagnosis in the test subsample using the revised 39-item ASDQ total item average score, in all cases (a) and in at-risk (b) cases. Abbreviations: ASD, autism spectrum disorder; ASDQ, Autism Symptom Dimensions Questionnaire; AUC, area under the curve.

average may be a sufficient approach to inform screening and diagnostic differentiation across demographic groups.

Limitations and future directions

The primary limitations of this study were having only one validated measure of psychopathology for evaluating discriminant validity, a limited number of items per ASD subscale, reliance on informant-reported clinical diagnoses of ASD and other developmental conditions from previously diagnosed individuals, and small subsamples across age and other demographic and clinical factors with a very limited sample size below age 5 years, thus precluding more detailed exploration of the potential screening utility of the instrument. Future work should include a larger sample of undiagnosed children aged 2 to 4 years, including stratification according to speech and language level and intellectual disability, to better assess the potential of the ASDQ as a screening tool. This will provide a more stringent test by focusing on undiagnosed individuals and will assist in determining whether modifications are needed to accurately differentiate cases with and without ASD in younger and more impaired samples. However, it is important to note that decreased reliability of informant-reported ASD diagnoses is more likely to attenuate validity estimates rather than inflate them. If sufficiently validated in future investigations, the ASDQ could be used in an evidence-based assessment capacity by a developmental pediatrician or child neurologist to screen for possible risk for ASD. For example, a clinician seeing a 4-year-old male child with no family history of ASD, but with concerns of social deficits, could use the ASDQ to inform the risk of ASD. A total item average score greater than 3.2 would warrant priority referral for tertiary care

evaluation with the probability of ASD moving from 3.7% (population estimate) to approximately 57% (elevated concern). Additionally, assuming good test–retest reliability and sensitivity to change are demonstrated in future research, the ASDQ could also be used after diagnosis to monitor autism symptom improvement during intervention.

The ASDQ factor structure requires replication in large ASD and developmental disability samples; several specific factors and scales had only a few items with high loadings. Although the resulting scales showed good overall and conditional reliability, content coverage of these domains is limited. Future instrument revisions may consider adding a small number of items to enhance content coverage for the weakest scales, particularly affiliation and sensory sensitivity. In the meantime, researchers investigating these domains with the ASDQ may simply choose additional measures, such as the Strengths and Difficulties Questionnaire[14] and the Dimensional Assessment of Repetitive Behavior,[37] which enhance coverage of key symptom domains.

Additional studies of convergent and discriminant validity with established criterion standard autism diagnostic instruments and informant report questionnaires are also needed, as well as studies of test–retest reliability and sensitivity to change. Investigations are also needed to examine different time intervals for recall to assess whether specific, shorter intervals maintain good validity for screening efficiency while also maximizing sensitivity to change.

A few ASDQ items are applicable only to individuals who use speech or older children. For example, items on restricted interests and high-level perspective taking are often not relevant for ages lower than 3 years. The approach of allowing informants to decide when a rating could not be provided is useful because most items were responded to; when items were rated not applicable, the instrument was still able to be

scored for nearly all participants. Future large-sample studies are needed to confirm this observation and identify whether loss of validity occurs for the total score or specific subscales due to nonresponse. Given the reading grade level of 7.9 and the potential screening use of the ASDQ, future instrument revisions may also consider attempting to decrease the reading level to ensure that a broader swathe of parent informants can interpret the items and provide accurate ratings.

Modestly sized developmental disability and neurotypical control samples are a further limitation of this study. Future studies with well-characterized neurotypical and developmental disability controls are needed to better understand ASDQ score distributions in these groups and provide more accurate estimates of screening and diagnostic validity. Another important next step will be to collect normative data on the ASDQ for computation of standard scores and inform clinical interpretation. The present data suggest very few typically developing and developmental disability controls with item average scores above 3.2 and very few cases with ASD with item average scores below 2.7 (5-point Likert scale), further reinforcing the screening potential of the ASDQ. However, large-sample normative data coupled with a large ASD sample are needed to develop multilevel likelihood ratios and related algorithms to use the ASDQ in an evidence-based medicine fashion to inform screening and clinical diagnosis.

In summary, the present data provide preliminary evidence that the ASDQ is a free, informant report measure of core autism symptoms that may be useful for screening and treatment monitoring, particularly if validated in larger at-risk and ASD samples, including those where ASD has not yet been diagnosed or where concern for ASD is not a predominant presenting problem. Despite relative brevity, the ASDQ offers comprehensive and reliable assessment of key domains of the core ASD phenotype, which is consistent with the DSM-5 ASD criteria. The instrument has potential to not only inform identification, but also has sufficient measurement precision to track individual differences from very low to very high symptom levels. If the present results are replicated, the ASDQ has strong potential to be widely adopted in future research and clinical practice.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Thomas W. Frazier https://orcid.org/0000-0002-6951-2667

REFERENCES
1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Fifth ed. Arlington, VA: American Psychiatric Association; 2013.
2. Constantino JN, Gruber CP. The social responsiveness scale manual, second edition (SRS-2). Los Angeles, CA: Western Psychological Services; 2012.
3. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. J Autism Dev Disord. 2001;31(1):5-17.
4. Rutter M, Bailey A, Lord C. The Social Communication Questionnaire Manual. Los Angeles, CA: Western Psychological Services; 2003.
5. Kanne SM, Mazurek MO, Sikora D, Bellando J, Branum-Martin L, Handen B, et al. The Autism Impact Measure (AIM): initial development of a new tool for treatment outcome measurement. J Autism Dev Disord. 2014;44(1):168-79.
6. Robins DL, Casagrande K, Barton M, Chen CM, Dumont-Mathieu T, Fein D. Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHAT-R/F). Pediatrics. 2014;133(1):37-45.
7. Lam KS, Aman MG. The Repetitive Behavior Scale-Revised: independent validation in individuals with autism spectrum disorders. J Autism Dev Disord. 2007;37(5):855-66.
8. Durkin MS, Elsabbagh M, Barbaro J, Gladstone M, Happe F, Hoekstra RA, et al. Autism screening and diagnosis in low resource settings: Challenges and opportunities to enhance research and services worldwide. Autism Res. 2015;8(5):473-6.
9. Lau WY, Kelly AB, Peterson CC. Further evidence on the factorial structure of the Autism Spectrum Quotient (AQ) for adults with and without a clinical diagnosis of autism. J Autism Dev Disord. 2013;43(12):2807-15.
10. Uljarević M, Jo B, Frazier TW, Scahill L, Youngstrom EA, Hardan AY. Using the big data approach to clarify the structure of restricted repetitive behaviors across the most commonly used autism spectrum disorder measures. Molecular Autism. 2021;12(1):39.
11. Frazier TW, Hardan AY. Equivalence of symptom dimensions in females and males with autism. Autism. 2017;21(6):749-59.
12. Uljarevic M, Phillips JM, Schuck RK, Schapp S, Solomon EM, Salzman E, et al. Exploring Social Subtypes in Autism Spectrum Disorder: A Preliminary Study. Autism Res. 2020;13(8):1335-42.
13. Uljarevic M, Frazier TW, Phillips JM, Jo B, Littlefield S, Hardan AY. Mapping the Research Domain Criteria Social Processes Constructs to the Social Responsiveness Scale. J Am Acad Child Adolesc Psychiatry. 2019;58(10S):S311.

14. Phillips JM, Uljarevic M, Schuck RK, Schapp S, Solomon EM, Salzman E, et al. Development of the Stanford Social Dimensions Scale: initial validation in autism spectrum disorder and in neurotypicals. Mol Autism. 2019;10:48.

15. Lord C, Pickles A, McLennan J, Rutter M, Bregman J, Folstein S, et al. Diagnosing autism: analyses of data from the Autism Diagnostic Interview. J Autism Dev Disord. 1997;27(5):501-17.

16. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. Eur Child Adolesc Psychiatry. 1998;7(3):125-30.

17. Frazier TW, Crowley E, Shih A, Vasudevan V, Karpur A, Uljarevic M, et al. Associations between executive functioning, challenging behavior, and quality of life in children and adolscents with and without neurodevelopmental conditions. Frontiers in Psychology. 2022;13:1022700

18. Marsh HW, Morin AJ, Parker PD, Kaur G. Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. Annu Rev Clin Psychol. 2014;10:85-110.

19. Li CH. Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. Behav Res Methods. 2016;48(3):936-49.

20. Putnick DL, Bornstein MH. Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. Dev Rev. 2016;41:71-90.

21. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide To Their Development and Use. 2nd ed. New York: Oxford University Press; 1995. 231 p.

22. Revelle W, Condon DM. Reliability from alpha to omega: A tutorial. Psychol Assess. 2019;31(12):1395-411.

23. Embretson SE, Reise SP. Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. 372 p.

24. Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill, Inc.; 1994.

25. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with applications in R. New York: Springer; 2013.

26. Marvin AR, Marvin DJ, Lipkin PH, Law JK. Analysis of Social Communication Questionnaire (SCQ) Screening for Children Less Than Age 4. Curr Dev Disord Rep. 2017;4(4):137-44.

27. Youngstrom EA, Salcedo S, Frazier TW, Perez Algorta G. Is the Finding Too Good to Be True? Moving from "More Is Better" to Thinking in Terms of Simple Predictions and Credibility. J Clin Child Adolesc Psychol. 2019;48(6):811-24.

28. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. Educ Psychol Meas. 2013;76(6):913-34.

29. Forero CG, Maydeu-Olivares A. Estimation of IRT graded response models: limited versus full information methods. Psychol Methods. 2009;14(3):275-99.

30. IBM Corp. IBM SPSS Statistics for Windows. 28.0 ed. Armonk, NY: IBM Corp; 2021.

31. Muthén LK, Muthén BO. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén; 1998-2017.

32. R Core Team. R: A language and environment for statistical computing. Vienna, Austria; 2021. Available from: URL https://www.R-project.org/.

33. Morin AJS, Arens AK, Marsh HW. A Bifactor Exploratory Structural Equation Modeling Framework for the Identification of Distinct Sources of Construct-Relevant Psychometric Multidimensionality. Structural Equation Modeling-a Multidisciplinary Journal. 2016;23(1):116-39.

34. Bishop SL, Hus V, Duncan A, Huerta M, Gotham K, Pickles A, et al. Subcategories of restricted and repetitive behaviors in children with autism spectrum disorders. J Autism Dev Disord. 2013;43(6):1287-97.

35. Uljarevic M, Frazier TW, Phillips JM, Jo B, Littlefield S, Hardan AY. Quantifying Research Domain Criteria Social Communication Subconstructs Using the Social Communication Questionnaire in Youth. J Clin Child Adolesc Psychol. 2021;50(5):609-618.

36. Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G. Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. J Am Acad Child Adolesc Psychiatry. 2008;47(8):921-9.

37. Uljarević M, Frazier TW, Jo B, Scahill L, Youngstrom EA, Spakcman E, et al. Dimensional Assessment of Restricted and Repetitive Behaviors: Development and Validation of a New Measure. J Am Acad Child Adolesc Psychiatry. 2022. Online ahead of print.

SUPPORTING INFORMATION

The following additional material may be found online:

Table S1: Model fit indices from factor analyses of the revised 39-item ASDQ in the exploratory and confirmatory subsamples

Table S2: Standardized factor loadings for the nine-factor exploratory structural equation model with a general ASD bifactor computed using all 39 items from the revised ASDQ

Table S3: Factor and scale correlations from the ASDQ in the total revised 39-item ASDQ sample

Table S4: Measurement invariance model comparisons for the revised 39-item ASDQ across sex, age, race, and ethnicity

Table S5: Reliability statistics for the revised 39-item ASDQ scales

Table S6: Demographic and clinical characteristics of individuals from the KidsFirst subsample who completed the revised 39-item ASDQ and SCQ

Table S7: Bivariate correlations between ASDQ and SCQ subscales

Table S8: Prediction of ASD in the total and the at-risk samples using all 39 of the revised ASDQ items

Figure S1: Exploratory bifactor model with nine specific factors and a general ASD factor in the revised 39-item ASDQ sample

Figure S2: Revised 39-item ASDQ total item average distribution by diagnostic groups, neurotypical, developmental disability, and ASD

Appendix S1: Background, methodological details, results, and discussion for the development and validation of the initial 33-item ASDQ

Appendix S2: The ASDQ

Appendix S3: Additional methodological detail for validation of the revised 39-item ASDQ

Appendix S4: Additional statistical details for validation of the revised 39-item ASDQ

Appendix S5: Additional results for the analyses of the revised 39-item ASDQ

Appendix S6: Additional discussion of the ASDQ psychometric findings, associated implications, limitations, and future directions