Title

Assessing gender differences in autism spectrum disorder using the Gendered Autism Behavioral Scale (GABS): An exploratory study

Authors

Clarke, Elaine
Hull, Laura
Loomes, Rachel
et al.

Publication Date

2021-10-01

DOI

10.1016/j.rasd.2021.101844

Peer reviewed

Assessing Gender Differences in Autism Spectrum Disorder Using the Gendered Autism Behavioral Scale (GABS): An Exploratory Study

Elaine Clarke,[1*] Laura Hull,[2*] Rachel Loomes,[3] Carolyn E.B. McCormick,[4] Stephen J. Sheinkopf,[5,6,7] William Mandy[2]

Affiliations:

[1] Department of Education and Information Studies, University of California, Los Angeles, Los Angeles, CA, 90095, USA

[2] Research Department of Clinical, Educational & Health Psychology, University College London, 1-19 Torrington Place, London, WC1E 7HB, UK

[3] South London & Maudsley NHS Foundation Trust, London, SE5 8AZ, UK

[4] Department of Human Development and Family Studies, Purdue University, 1200 West State Street, West Lafayette, IN, 47907, USA

[5] Brown Center for the Study of Children at Risk, Women & Infants Hospital, 50 Holden Street, Providence, RI, 02908, USA

[6] Department of Psychiatry & Human Behavior, Warren Alpert Medical School of Brown University, 700 Butler Drive, Providence, RI, 02906, USA

[7] Hassenfeld Child Health Innovation Institute, Brown University, 121-4121 South Main Street, Providence, RI, 02912, USA

*Joint first authors

Contact Information:

Elaine Clarke: ebclarke42118@g.ucla.edu

GENDERED AUTISM BEHAVIORAL SCALE

Laura Hull: laura.hull.14@ucl.ac.uk Corresponding Author

Rachel Loomes: Rachel.Loomes@slam.nhs.uk

Carolyn E.B. McCormick: mccorm37@purdue.edu

Stephen J. Sheinkopf: stephen_sheinkopf@brown.edu

William Mandy: w.mandy@ucl.ac.uk

Abstract

Introduction: Gendered differences in autism spectrum disorder (hereafter, 'autism') symptomatology, may contribute to delayed diagnoses for autistic females. The aim of this study was to develop a coding system, the Gendered Autism Behavioral Scale (GABS), to identify and measure hypothesized components of non-traditional autism phenotypes.

Methods: Two independent cohorts of autistic individuals completed modules 3 and 4 of the Autism Diagnostic Observation Schedule (ADOS, 2nd edition). Video-recorded ADOS administrations were coded using the GABS, and separate coding teams analyzed each cohort. Cohort 1 from the United Kingdom consisted of 22 males and 22 females, aged 9-15 years. Cohort 2 from the United States consisted of 40 males and 20 females, aged 4-59 years.

Results: The coders achieved acceptable inter-rater reliability both within and across coding teams. In exploratory analyses, gender differences between codes were assessed within cohorts. Within Cohort 1, there were significant gender differences, of a moderate size, on several individual items as well as the Managing Emotions subscale and the Total GABS score. Within Cohort 2, significant gender differences were found for two individual items.

Conclusions: This study demonstrated the feasibility of the GABS across different sites. Validity tests resulted in partial replication of gender differences on the GABS. Preliminary evidence from the GABS suggests that valuable data on hypothesized non-traditional autism phenotypes could be extracted from widely employed assessments such as the ADOS. Future work could capitalize on the GABS' utility for secondary data analysis to study gender differences in ASD in larger, adequately powered samples.

Keywords: Gender, Sex, Diagnosis, ADOS, Autism Spectrum Disorder

Introduction

Historically, autism spectrum disorder (hereafter 'autism') has been diagnosed more commonly in males than females (Fombonne, 2009). However, a recent meta-analysis suggests the prevalence of autism may be underestimated in females, particularly in clinical settings (Rachel Loomes et al., 2017). There is also evidence that on average, females receive ASD diagnoses later than their male peers (Begeer et al., 2013; McCormick et al., 2020) and may require additional needs to receive a diagnosis (Dworzynski et al., 2012). These findings have led some to assert autism in females is currently underdiagnosed (Jamison et al., 2017; Kopp & Gillberg, 1992). Missed or late diagnosis can result in exclusion from early interventions and other services across development, meaning females who may benefit from such services may be unable to access them in a timely matter, if at all (Rogers & Vismara, 2008).

Female (Non-Traditional) Autism Phenotypes

One explanation for the underdiagnosis of autistic females is the Female Autism Phenotype hypothesis (Bargiela et al., 2016; Lai et al., 2015) which states that the behavioral presentation of autism may vary across genders in subtle but important ways. Females may express the same underlying characteristics associated with autism through different behavioral presentations than males, which are not always identifiable given the current conceptualization of autism. Importantly, these behaviors (although generally associated with females) can also be observed in males and people of non-binary gender – and there are autistic females who will not display any or all of these proposed behavioral characteristics. Thus, in this study we will refer to 'Non-Traditional Autism Phenotypes', to capture how differences in symptomatology across the spectrum of autism may converge with differences across the spectrum of gender. Core autistic traits include social/communication difficulties, and restricted and repetitive interests

including sensory over- or under-sensitivity (American Psychiatric Association, 2013). Variation in males and females' development and socialization may lead to differences in the precise ways these underlying traits are expressed, resulting in some individuals expressing their autism in qualitatively different ways which are not identified by clinicians using current diagnostic tools (Lai et al., 2015). An alternative scenario is that delayed or missed diagnoses in females is the result of milder presentations that impact the age at which autism related symptoms manifest or rise to clinically significant levels. Measurement designed to capture non-traditional phenotypes could help resolve these issues.

The current understanding of autism—and the development of measures used in its assessment—is based on predominantly male samples, reflecting the gender ratios observed at the time. These measures may not encapsulate the full range of "behavioral exemplars" of autism, and may therefore be inadvertently biased against females (Hiller et al., 2014; Lai et al., 2015). The following areas have the greatest evidence for sex/gender differences in how autistic traits are expressed and were thus included in our development of the GABS.

Camouflaging. The concept of camouflaging–minimizing the appearance of autistic characteristics, whether consciously or not–was first proposed by Lorna Wing (Wing, 1981). Recent research has determined camouflaging is used by autistic females, as well as males and non-binary people, as a way to fit in with other people and form connections (Hull, Petrides, et al., 2017; Hull et al., 2020). Camouflaging as a coping strategy may in some cases be an effortful and not always successful process that, through chronic stress, could increase the risk for mental health problems (Bargiela et al., 2016; Cassidy et al., 2019; Lai et al., 2017). Autistic females have been shown to camouflage to a greater extent than males (Hull et al., 2020; Lai et al., 2017; Schuck et

al., 2019) although there is some research suggesting no gender differences in camouflaging (Cage & Troxell-Whitman, 2019; Livingston et al., 2020). While camouflaging may be harder to observe during a diagnostic interview, autistic individuals may self-report using camouflaging strategies, or demonstrate camouflaging attempts which can be identified by others (Hull, Petrides, et al., 2017).

Social relationships. In verbally and cognitively able samples, autistic girls and women show greater motivation to form friendships than autistic boys and men (Dean et al., 2014; Head et al., 2014). On average, autistic girls may have similar levels of friendship motivation to typically developing girls (Sedgewick et al., 2015), in contrast to common clinical perceptions of reduced sociality in autism. This raises the possibility that social motivation in autistic girls may be a factor influencing the likelihood of diagnosis in females compared to males (Little et al., 2017). Despite evidence of relatively high social motivation, autistic women report difficulties maintaining friendships, suggesting difficulties with social relations are far from absent (Kanfiszer et al., 2017). It is therefore important that behavioral assessments of autism measure the desire for, quality, and maintenance of close relationships, such as friendships, in addition to the presence or absence of these relationships.

Internalizing. Internalizing difficulties such as anxiety and depression are common in autism, especially amongst females (Hartley & Sikora, 2009; Hiller et al., 2014; Lai et al., 2019). In contrast, autistic males have on average higher levels of externalizing problems (Mandy et al., 2012; May et al., 2012). Internalizing problems are less visible to caregivers and teachers (who typically serve as informants for studies of behavior problems in children and adolescents) and therefore may contribute to underreporting of clinically significant behavioral differences in autistic females.

Relational interests. While some evidence suggests autistic females have lower levels of restricted interests than males (Frazier et al., 2014; Frazier & Hardan, 2017), it has been argued this reflects differences in type rather than intensity (Duvekot et al., 2017; Mandy et al., 2012). This hypothesis suggests females are more likely to have interests focused on relationships with or between others, including animals, celebrities, or fictional characters, whereas males' interests may be focused on systems and objects such as vehicles or scientific phenomena (Grove et al., 2018; Hiller et al., 2014). The intensity of these relationship-based interests may not be identified during clinical assessments if the topics are seen as (gender) typical, even though interests may be as intense as those of males.

Autism Diagnostic Observation Schedule, Second Edition (ADOS-2)

The ADOS-2 is a gold-standard assessment tool for autism (Lord et al., 2012). This semi-structured standardized assessment is one of the main observational tools used when diagnosing the condition (Rogers et al., 2016). However, like many autism assessments, the ADOS-2 was developed using a majority male sample, reflecting gender ratios observed when the original ADOS was developed (Lord et al., 2000). It is therefore possible the ADOS-2 and similar assessments, such as the Autism Diagnostic Interview, Revised (ADI-R; Lord et al., 1994) have reduced sensitivity for non-traditional presentations of autism, including females (Beggiato et al., 2017; Lai et al., 2011).

If this is the case, using the ADOS-2 alongside a measure assessing non-traditional phenotypes could increase the sensitivity for identifying autism in women and girls, and improve researchers' ability to study gender differences in autism in a standardized manner. The incorporation of a gender-specific assessment into autism diagnostic batteries would have substantial utility in efforts to empirically test for the presence and/or prevalence of the non-traditional autism phenotype across the entire

autistic population. Further, the ability to use such a measure for secondary data analysis using recordings of ADOS assessments or similar clinical interactions could enable researchers to reexamine existing data for evidence of gender differences, rather than devoting time and energy to recruiting new participants.

Aims

The aim of the current study was to develop a method of measuring variation in autistic presentation alongside a standardized tool, the ADOS-2. This study is the first known pilot of a coding frame specifically designed to pick up non-traditional behaviors with an existing and widely used assessment tool. The objectives were as follows:

1) To develop a coding frame measuring hypothesized behaviors from a non-traditional autism phenotype
2) To investigate whether said framework can measure these behaviors via module 3 and 4 ADOS recordings
3) To test the inter-rater reliability of this coding frame both in and across two independent research sites.
4) To explore gender differences in autism symptom presentation using the GABS coding frame in two independent cohorts.

Methods

Overview of Design

The Gendered Autism Behavioral Scale (GABS) coding frame was developed by clinicians with extensive experience diagnosing autism in girls and women using both the ADOS and clinical judgement. The GABS was then refined using a small sample of ADOS videos, before being piloted in two larger cohorts. Module 3 and 4 ADOS assessments completed at two independent sites were video recorded and coded. This study was conducted with the approval of both the South West – Frenchay Research Ethics Committee in the United Kingdom and the Lifespan Hospitals Institutional Review Board in the United States. Written consent was sought from participants and their families to use ADOS assessment recordings for this research. Eight participants and their families gave additional written consent for their ages and ADOS recordings to be shared across sites for reliability training.

Participants

Cohort 1. Twenty-two autistic males and 22 autistic females were recruited through a specialist National Health Service (NHS) clinic in the UK. All participants received a clinician consensus autism diagnosis based on ICD-10 criteria (World Health Organization, 2018), were aged between 9 and 15 years old at their initial assessment, and had an IQ above 70, i.e. no intellectual impairment, as assessed by their performance on the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV) (Wechsler, 2003). There were no significant gender differences in respect to age or IQ. See Table 1 for participant characteristics.

Cohort 2. Forty autistic males and 20 autistic females were recruited through a US patient registry, the Rhode Island Consortium for Autism Research and Treatment (RI-CART) Participants were recruited from Rhode Island hospitals, outpatient clinics, group

homes, community/parent groups, schools, and autism-related community events. All participants had a clinical diagnosis of autism (Autistic Disorder, ASD, Asperger's syndrome, or pervasive developmental disorder-not otherwise specified [PDD-NOS]), an ADOS-2 comparison score above the autism spectrum cut-off, were aged between 4 to 59 years old at the time of their enrollment and had no apparent intellectual impairment. IQ data from the Kaufman Brief Intelligence Test, Second Edition (KBIT-2) (Kaufman & Kaufman, 2004) or Stanford-Binet Intelligence Scales, Fifth Edition (SB-5) (Roid, 2003) were available for the majority (n = 42) of this sample. For these individuals, a standard score of 85 or higher was used as the inclusion criteria. IQ data were not available for the remaining 18 participants. For these individuals, an adaptive behavior composite (ABC) standard score of 50 or higher, on the Vineland Adaptive Behavior Scales, Second Edition (Vineland-II) (Sparrow, Balla, & Cichetti, 2005) was used in lieu of an IQ measure. Prior research has shown that Vineland-II ABC standard scores are strongly positively correlated with IQ (Bishop et al., 2015; Sparrow & Cicchetti, 1985). Two male participants were matched to each female participant on the basis of age and either IQ or Vineland-II ABC scores.

Table 1. *Characteristics of male and female participants across Cohorts 1 and 2.*

	Cohort 1			Cohort 2		
	Females (*n* = 22)	Males (*n* = 22)	Gender differences	Females (*n* = 20)	Males (*n* = 40)	Gender differences
	M (SD)	M (SD)		M (SD)	M (SD)	
Age	11.59 (2.44)	11.05 (1.94)	p = .421	14.00 (9.64)	14.39 (12.60)	p = .596
ADOS-2 Comparison Score	6.95 (2.32)	5.41 (2.27)	p = .032	7.09 (2.04)	7.52 (1.80)	p = .621

WISC-IV Overall Score	101.72 (18.81)	100.91 14.12)	p = .971	—	—	—
KBIT-2 Overall Score	—	—	—	110.15 (18.23)	102.00 (16.64)	p = .274
SB-5 Overall Score	—	—	—	98 (—[a])	98.20 (10.45)	p = .958
Vineland-II ABC Score	—	—	—	71.82 (9.77)	71.04 (13.67)	p = .650

Note: Gender comparisons of descriptive characteristics across cohorts were not conducted.
[a]The SB-5 was only used to measure IQ in one female participant, thus a standard deviation for female participants' scores on the SB-5 could not be calculated.

Procedures

Systematic development of the GABS. Initial ideas for codes were developed at the UK site through consideration of: (1) the empirical literature on autistic sex/gender differences, including comprehensive systematic reviews (Hull, Mandy, et al., 2017; Lai et al., 2015; Loomes et al., 2017; Van Wijngaarden-Cremers et al., 2014); (2) published

accounts of autistic sex/gender differences by autistic individuals (Holliday Willey, 2015; Simone, 2010); (3) consultation with experienced autism clinicians. Codes were further developed through iterative discussion between three research reliable ADOS administrators over a period of several weeks, and initial piloting was conducted with three male and three female cases not included in either cohort. Item phrasings and coding instructions were modified following this pilot before being finalized for use in the present study.

Raters. Both cohorts were coded by teams of two, for a total of four raters. Cohort 1 was coded by two psychology doctoral students specializing in the study of autism. These coders were also closely involved in the initial development of the GABS manual and coding frame, and thus were not blind to the hypotheses of the current study. Cohort 2 was coded by a post-baccalaureate research assistant in an autism lab and an undergraduate cognitive science student. The research assistant was trained to score the GABS by the Cohort 1 coders, and was not blind to the study hypothesis. However, the undergraduate student, who coded the majority of the Cohort 2 data, was blind to the study hypotheses. The research assistant only coded Cohort 2 data for the calculation of inter-rater reliability.

Coding. After the initial pilot and subsequent amendments, the GABS was used to code Cohort 1 by the UK researchers. Within Cohort 1, 40% of the cases (9 males and 9 females) were coded for inter-rater reliability. Following the completion of Cohort 1, the UK site trained a US site researcher in the coding system and achieved inter-rater reliability (k = 0.69) across eight ADOS assessment recordings. This researcher then trained a second, independent researcher at the US site to code the GABS. Within Cohort 2, 22% of the cases (9 males, 4 females) were coded for interrater reliability.

Measures

Gendered Autism Behavioral Scale (GABS). The GABS consists of four categories: A. Social Adaptation, B. Social Relationships, C. Managing Emotions and D. Interests. More information about the structure of the GABS is detailed in the Results section.

Autism Diagnostic Observation Schedule (ADOS). The ADOS-G (Lord et al., 2000) and ADOS-2 (Lord et al., 2012) are successive forms of a semi-structured, standardized assessment of communication, social interaction, and play or imaginative use of materials for individuals who have been referred for suspected autism. Modules 3 and 4 of the ADOS-2, aimed at verbally fluent children/young adolescents, and verbally fluent older adolescents/adults respectively, were used in this study. The ADOS is widely used and its psychometric reliability and validity have been demonstrated in a number of samples (Lord et al., 2012; Lord et al., 2000). All ADOS assessments coded for the current study were administered by research reliable assessors.

Data Analysis

Inter-rater reliability. In Cohort 1, two raters coded a random selection of 18 of the 44 (40%) ADOS recordings using the GABS. In Cohort 2, two raters coded a random selection of 13 of the 60 (22%) ADOS recordings using the GABS. Inter-rater reliability was judged by calculating the Cohen's κ coefficient for each item (Cohen, 1968). For the majority of codes, a standard Cohen's κ was calculated. For items with ordinal rather than categorical codes, weighted Cohen's κ was calculated.

Correlations between ADOS and GABS. To explore similarities and differences between ADOS and GABS subscales, correlations between domain scores on each measure were run separately in the two cohorts.

Gender comparison. To establish whether there were gender differences on individual GABS items, Fisher's exact tests were performed in 2x2 and 2x3 contingency tables. Cramer's V effect size was calculated to indicate if the effect was small (.10), moderate (.30) or large (.50) (Gravetter & Wallnau, 2017). Gender differences were calculated for each of the GABS subscales and a total GABS score, using t-tests and Mann-Whitney U tests as appropriate.

Results

Development of the GABS

Following the initial generation of ideas for items, behavioral codes were separated into four categories to represent the four key proposed components of the non-traditional autism phenotype: Social Adaptation, Social Relationships, Managing Emotions, and Interests & Hobbies. There are multiple items in each category (see Table 2 for details). Each GABS item is accompanied by a brief descriptor, in addition to more detailed explanation in the training manual. Codes are either dichotomous (i.e. 0= no evidence of behavior, 1 = behavior present) or are a scale with up to four options. For example, when coding item C1, 'Internalizing Difficulties', there are three options (0 = no evidence of symptoms, 1 = possible symptoms, 2 = definite symptoms).

Social Adaptation. This category covers behaviors proposed to contribute to the camouflaging of autism, namely masking of autistic characteristics (such as stimming, or inappropriate facial expressions), and compensation for autistic difficulties/differences (e.g. forcing oneself to make eye contact). This category also includes items on the extent of self-reflection the individual shows during the ADOS assessment.

Social Relationships. This category focuses on friendships, social behaviors, and skills that facilitate initiation and maintenance of friendships, including those

displayed during the ADOS assessment. Key codes focus on the individual's understanding and reporting of friendships, as well as their response to conversational cues from the interviewer.

Managing Emotions. This category encompasses the individual's reporting and demonstration of emotions, and what influences those emotions. In particular, items emphasize how well the individual identifies and expresses internalizing and externalizing emotions, and whether there are any specific social behaviors (such as acceptance or rejection by peers) which the individual endorses as impacting their emotions.

Interests & Hobbies. The final category covers the nature and intensity of a participant's interests, which can include responses to direct questions about interests, and any interests brought up spontaneously. Interests are sorted into two types: physical-mechanical (where the primary feature involves object-oriented such as taxonomies and mechanisms), and relational (where the primary feature involves people or animals).

[INSERT TABLE 2 HERE]

Inter-rater reliability

κ coefficients were calculated separately for each cohort for all 18 GABS items (Table 3). Due to low counts for camouflaging observed in the ADOS, A1 (masking) and A2 (compensating) were combined to create a general camouflaging code. For item B1 (Reported friendships), κs could not be calculated as in both cohorts one rater gave the same rating for all participants. This suggests little variation on this item, calling its usefulness into question. No items had a κ coefficient of 0-0.20 (none to slight agreement).

In Cohort 1, two items (B3: Quality of Reported Friendships and C5: Violation of Sameness or Rigidness on Emotions) had only fair agreement (κ = 0.21-0.40). Due to the low agreement on these items, they were removed from the main analysis in both cohorts. Seven items had moderate agreement (κ = 0.41-0.60), four items had substantial agreement (κ = 0.61-0.80) and four items had almost perfect agreement (0.81-1.00). Of the 15 items included for further analysis, the mean κ was 0.65, indicating overall the items included in the main analysis showed substantial agreement.

In Cohort 2, two items had moderate agreement (κ = 0.41-0.60), seven items had substantial agreement (κ = 0.61-0.80) and six items had near-perfect agreement (0.81-1.00). Of the 15 items included for further analysis, the mean κ was 0.79, indicating overall the items included in the main analysis showed substantial agreement.

Table 3. *Cohort 1 and 2 κ coefficients for all items*

GABS Item	Description	Cohen's Kappa	
		Cohort 1	Cohort 2
A1 + A2	Camouflaging (masking + compensation)	1.0	0.95

A3	Self-reflection on social behavior	0.53	0.80
B1	Reported friendships	NA[a]	NA[a]
B2	Understanding of friendships	0.44	0.70
B3	Quality of reported friendships	0.22	0.90
B4	Age of friends	0.89	0.80
B5	Nature of friendships	0.67	0.70
B6	Responding to conversational cues	0.80 [b]	0.60 [b]
B7	Social interest	0.44 [b]	0.90 [b]

GENDERED AUTISM BEHAVIORAL SCALE

C1	Internalizing difficulties	0.51	0.70
C2	Externalizing difficulties	0.89	0.60
C3a	Communicating about emotions (verbal)	0.51 [b]	1.0 [b]
C3b	Communicating about emotions (non-verbal)	0.42 [b]	0.70 [b]
C3c	Communicating about emotions (different to communicating about other topics)	0.44	0.90
C4	Influence of social acceptance or rejection on emotions	1.0	0.90
C5	Violations of sameness or rigidness on emotions	0.40	0.70
D1	Reported interests	0.68 [b]	.90 [b]

D2	Type of interests	0.73	0.80

[a] unable to calculate κ as one rater is a constant. [b] weighted Cohen's κ.

Correlations between ADOS and GABS

Correlations between ADOS domain scores (Social Affect and Restricted, Repetitive Behaviours) and GABS subscale scores (Social Adaptability, Managing Emotions, Social Relationships, and Interests) are presented in Table 4. Scores for Cohort 1 are below the diagonal and scores for Cohort 2 are above the diagonal.

In Cohort 1, GABS Managing Emotions, Social Relationships, and Interests subscales were all significantly negatively correlated with ADOS Social Affect scores. GABS Social Relationships was also negatively correlated with ADOS RRB scores. This suggests that the GABS codes, despite being conceptually similar, pick up behaviours not otherwise captured in the ADOS. **add detail on RICART findings/consistency across sites**

Table 4. Correlations between GABS subscale scores and ADOS-2 domain scores. Cohort 1 (UK) is presented below the diagonal, Cohort 2 (US) above the diagonal. Significant correlations are in bold.

	GABS Social Adaptability	GABS Managing Emotions	GABS Social Relationships	GABS Interests	ADOS Social Affect	ADOS Restricted Repetitive Behaviours
GABS Social Adaptability						
GABS Managing Emotions	0.26					
GABS Social Relationships	**0.47*** **	0.21				
GABS Interests	0.29	0.14	**0.35***			
ADOS Social Affect	-0.29	**-0.32***	**-0.42****	**-0.33***		
ADOS Restricted Repetitive Behaviours	-0.23	-0.13	**-0.35***	-0.10	0.21	

*p≤.05; **p≤.01; ***p≤.001

Gender comparison

Results of Fisher's exact tests for individual GABS items are displayed in Table 5 for Cohort 1, and Table 6 for Cohort 2. Scores for each subscale were compared between genders, using t-tests when data were normally distributed (Managing Emotions) and Mann-Whitney U tests when data were not normally distributed (Social Adaptation, Social Relationships, Interests). To ensure items had equal weighting in their subscale, they were all given the same minimum and maximum possible score: items with options 0 and 1 were given scores 0 and 2 for calculating total score, and items with options 0, 1 and 2 were given scores 0, 1 and 2 for calculating total score.

In Cohort 1, significant gender differences were found for the Internalizing Problems, Impact of Acceptance or Rejection on Emotions, and Type of Interest items. Females were significantly more likely to report internalizing symptoms ($\chi 2(1, n = 44) = 5.87, p = .03$) and emotional impact due to social acceptance and rejection ($\chi 2(1, n = 44) = 5.35, p = .05$) than males. Females also volunteered more relational interests ($\chi 2(1, n = 44) = 12.03, p < .01$) than males. Females ($m = 3.38$, SD = 1.10) had higher scores on the Managing Emotions subscale, which measures display and reporting of internalizing and externalizing symptoms ($t(42) = -2.01, p = .05$) than males ($m = 2.47$, SD = 1.81), and females ($m = 12.35$, SD = 4.57) had higher Total GABS scores ($t(42) = -2.55, p = .02$) than males ($m = 9.00$, SD = 4.13). No other significant gender differences were found. See Table 5 for subscale and total GABS scores by gender, and effect sizes for differences.

Table 5. *Mean scores for GABS subscales and total scale for males (n = 22) and females (n = 22) in Cohort 1.*

Subscale (Range)	**Female mean (SD)**	**Male mean (SD)**	**Difference**	**Effect size (Cohen's d)**
Social Adaptation (0-4)	0.84 (0.77)	0.50 (0.74)	U = 284.00, p = 0.27	d = 0.30
Managing Emotions (0-12)	3.38 (1.10)	2.47 (1.81)	t(42) = -2.01, p = 0.05	d = 0.61
Social Relationships (0-7)	3.42 (1.68)	2.45 (1.77)	U = 309.50, p = 0.11	d = 0.20

Interests (0-4)	2.68 (1.16)	2.45 (0.60)	U = 250.00, p = 0.25	d = 0.05
Total GABS (0-27)	10.52 (3.87)	7.61 (3.80)	t(42) = -2.51, p = .02	d = 0.78

In Cohort 2, significant gender differences were found for Externalizing Problems and Frequency and Intensity of Interest. Males reported significantly more externalizing symptoms ($\chi 2(1, n = 60) = 6.56$, $p = .01$) and more intense interests ($\chi 2(3, n = 60) = 8.81$, $p = .032$) than females. There were no significant gender differences in subscale scores. Females ($m = 16.38$, SD = 5.05) and males ($m = 14.52$, SD = 5.42) also did not significantly differ in Total GABS scores ($t(58) = 0.784$, $p = .076$). See Table 6 for subscale and total GABS scores by gender, and effect sizes for differences.

Table 6. *Mean scores for GABS subscales and total scale for males (n = 40) and females (n = 20) in Cohort 2.*

Subscale (Range)	**Female mean (SD)**	**Male mean (SD)**	**Difference**	**Effect size (Cohen's d)**

Social Adaptation (0-4)	1.23 (1.2)	0.8 (1.16)	U = 317.00, p = 0.14	d = 0.36
Managing Emotions (0-12)	5.75 (1.83)	5.93 (2.63)	t(58) = -.27, p = 0.79	d = 0.08
Social Relationships (0-7)	2.90 (1.68)	2.55 (1.99)	U = 348.50, p = 0.41	d = 0.19
Interests (0-4)	3.15 (0.93)	2.85 (0.80)	U = 298.50, p = 0.07	d = 0.35
Total GABS (0-27)	16.38 (5.42)	14.52 (5.05)	t(58) = 1.32, p = .19	d = 0.36

Discussion

This study developed the GABS coding frame as a measure of hypothesized behaviors of behavioral presentation of autism in girls and women. Two coding teams were trained to reliability on the GABS and piloted the codes in two independent cohorts. The first, relatively homogenous cohort tested the application of the GABS, and the second, more heterogenous cohort demonstrated the GABS could be used reliably across research sites. In preliminary exploratory analyses, GABS scores from both cohorts were used to examine whether the behaviors assessed by GABS items occurred more often in females than males. In Cohort 1, females had higher scores on the Managing Emotions subscale and Total GABS score. Significant gender differences were also found for the Internalizing Problems, Impact of Acceptance or Rejection on Emotions, and Type of Interest items. In Cohort 2, there were no significant gender differences in the subscale or total GABS scores. Significant gender differences were found for Externalizing Problems and Frequency and Intensity of Interest items.

A primary aim of this study was to help define a method for future investigations of gender differences in the expression of autism. More piloting is needed to determine 1) if the successful identification of gender differences in childhood and adolescence seen in Cohort 1 can be replicated and 2) if these gender differences can also be identified in older individuals such as those included in Cohort 2, perhaps through the inclusion of more age-appropriate items. With further piloting, the GABS' utility in secondary data analysis and ability to be used in complement with widely employed assessments such as the ADOS could make it a promising measure for studying autism gender differences at a large scale.

GABS Reliability

Reliability between coders within and across sites was acceptable to good. This suggests, in addition to being used reliably across multiple participants, that the GABS can be used across multinational sites and a variety of participant ages. This high level of transferability suggests further validation and application of the GABS in a range of different clinical contexts would be meaningful.

Interrater reliability analyses from both cohorts suggest that GABS items can be coded accurately after 6-8 hours of training. Raters in this study ranged broadly in their professional training and familiarity with autism, from doctoral students specializing in the condition to undergraduate volunteers with little prior familiarity. This suggests that with training, the behaviors scored via the GABS are readily apparent even to non-experts. Furthermore, initial reliability training between Cohorts 1 & 2 was conducted remotely; this did not appear to impede the efficiency nor ease of training to reliability. Given the global surge in remotely conducted research and telehealth as a result of the COVID-19 pandemic, the GABS' ease of use in virtual settings is notable.

Similar levels of inter-rater reliability were found for both cohorts, indicating that the GABS is transferable, with codes that can be understood by individuals of varying skill levels. However, the reliability of GABS administrations across cohorts is contingent upon the standardization of ADOS administrations across cohorts. The ADOS is a highly standardized instrument, but inevitably administrations will vary across individual assessors, research and clinical contexts, geographic locations, etc. Individuals' GABS scores are directly based upon the contents of their ADOS administrations. It is therefore important to note even minor differences in ADOS administrations across the two study cohorts could have influenced GABS results.

Total GABS scores were noticeably higher in Cohort 2 than Cohort 1 (although differences between cohorts were not compared statistically). Upon examination of the subscales, it seems that this is mostly driven by Cohort 2 participants (male and female) scoring higher on the 'Managing Emotions' subscale. One explanation for this is that adults and older children are better able to recognize and express their own emotions than younger children (Zeidner et al., 2003). Participants in Cohort 2 had a higher mean age compared to Cohort 1, and adults were included in Cohort 2 but not Cohort 1. We suggest that future studies should evaluate the GABS in a range of ages to determine mean scores across development. It may also be beneficial to adapt GABS questions or scoring to be more age-appropriate, following further examination across a broader range of samples.

Gender Differences in GABS Scores

No evidence for gender differences in camouflaging or social-relational constructs was found in either Cohort 1 or Cohort 2. Camouflaging behaviors may be relatively rare in both genders. Alternatively, this finding may be a product of attempting to measure camouflaging behaviors in the context of an observational measure, the ADOS-2, that does not include direct questions or observational presses specifically designed to detect these types of behaviors (Wood-Downie et al., 2020). Prior research indicates camouflaging is a clinically important phenomenon (Lai et al., 2017), and includes evidence for an association between camouflaging and mental health problems (Cassidy et al., 2019; Hull et al., 2018; Lai et al., 2017). Testing for the presence of camouflaging in autism may require modified observational presses and/or additional interview questions, beyond the scope of the ADOS-2 in its current form. Table 7 includes

suggested probes that could be incorporated into an ADOS module 3 or 4 administration to elicit descriptions of camouflaging.

[INSERT TABLE 7 HERE]

Males and females did not differ in the types of friendships they reported, nor in reported levels of social interest, in either cohort. Previous research demonstrating gender differences in social relationships in autism has focused on the quality of relationships (Sedgewick et al., 2018); this is not explicitly assessed in the ADOS and so could not be measured in the GABS. We have suggested a probe which could be added to the ADOS to elicit descriptions of friendship quality in Supplementary Table 1. Females in Cohort 1 reported greater emotional influence of social interactions – acceptance or rejection by others – than males, supporting previous findings of greater social motivation (Dean et al., 2014; Head et al., 2014). However, further research is needed to confirm whether this non-traditional presentation of autism can be measured using the GABS, as this difference was not found in Cohort 2.

Evidence for gender differences in internalizing and externalizing, as measured by the GABS, was found in both cohorts. In Cohort 1, females reported significantly more internalizing difficulties than males. In Cohort 2, males reported significantly more externalizing difficulties than females. In combination, these findings support previous conclusions that emotional difficulties in autistic individuals may be expressed differently depending on gender (Hiller et al., 2014; Mandy et al., 2012; Oswald et al., 2016). The GABS can be used to identify these differences and so may be useful in assessing possible gender-specific variations in psychiatric comorbidities as part of the diagnostic process. Considering autistic individuals often have multiple comorbid psychiatric

diagnoses (Lai et al., 2019), the GABS could provide a useful framework for assessing psychiatric comorbidities in comparison to patient self-report or direct clinical observation.

Comparison of focused interests in both cohorts revealed different, but complimentary results. In Cohort 1, females reported more relational interests (focused on people and animals) than males, with no difference in intensity. However, in Cohort 2 males reported greater intensity of interests than females but no difference in type was found. One explanation for the difference in findings might be the extent to which interests were discussed in ADOS interviews across each cohort. Males in Cohort 1 provided conversational leads related to their interests, whereas females did not. In contrast, all Cohort 2 participants provided leads for questions about their interests during the ADOS. We have suggested a probe to explicitly query about interests, in Supplementary Table 1.

The GABS was designed as a measure of non-traditional autistic characteristics in females and males; higher scores do not necessarily represent a more ‘female’ presentation of autism, but greater presentation of autistic characteristics which are not currently measured on the ADOS. The present analyses examined whether females in both cohorts would demonstrate higher overall scores on this measure. In Cohort 1 this was confirmed; however, in Cohort 2 there was no significant gender difference in total GABS score. The GABS may be a better measure in younger individuals, as included in Cohort 1. The ages of participants in Cohort 2 ranged from early childhood to middle adulthood, thus age differences may have outweighed gender differences in Cohort 2. It should also be noted that gender differences observed in the present study were

generally of a moderate effect size (interpreted following Cohen, 1988); post-hoc power analysis (using Gpower 3.1.9.2; Faul et al., 2007) revealed that both Cohort 1 and Cohort 2 were underpowered to detect differences smaller than d = .80. It is therefore possible that gender differences on the GABS may exist but were not captured, whether fully or at all, in the present study. Given the small *n* and exploratory nature of these analyses, additional work in larger cohorts is necessary to determine whether the GABS composite score is an accurate measure of a non-traditional phenotype that may be prevalent in females with autism.

Limitations and Future Directions

Importantly, this study demonstrated that the GABS can be used by raters of various levels of training and familiarity with autism, as well as across cultural contexts. However, additional work in larger samples is needed to examine the psychometric properties of the GABS in detail. Specifically, only one of the raters in the current study was blind to the purpose of the GABS; future studies exclusively using blind raters are needed to confirm the GABS' sensitivity in detecting gender differences. Care was taken to design GABS items that were distinct from existing ADOS items, and this was supported by the negative or non-significant correlations between GABS subscale and ADOS domain scores identified [in Cohort 1], suggesting that the two measures do capture different aspects of autistic presentations. However some GABS items (such as B1) showed little variation and therefore should be further examined and potentially removed following additional testing.

Given the possibility that the ADOS and other commonly used assessment measures are more sensitive to a presentation of autism in males, it is possible that

there would have been more gender differences in GABS scores had both cohorts included individuals that were suspected of having autism but did not have an official diagnosis of ASD and/or did not meet criteria on the ADOS. However, because this study was conceived as a proof of concept and pilot of the GABS assessment, the authors felt it more methodologically sound to measure the feasibility of the GABS in individuals with definitive diagnoses of autism. The GABS was not designed to capture the full range of autistic characteristics, and so participants who already had demonstrated autistic characteristics on the ADOS were included with the aim that the GABS might identify individual differences in additional characteristics not captured by the ADOS.

Future research is needed to elucidate whether the GABS may be used to help describe autism features in those who might otherwise be missed by current assessment measures. Specifically, future research should test the GABS in males and females who are suspected of having autism but have not yet been formally assessed, or individuals who are judged by expert clinicians to meet the diagnostic criteria for autism despite a negative (i.e., false negative) finding on a diagnostic assessment tool and/or scores below ADOS thresholds. This would determine whether the GABS can be used more broadly to measure non-traditional autism presentations, and therefore improve the sensitivity of autism assessments when combined with the ADOS. We note here that we are not suggesting that the GABS is used alone at this stage of development, but that it might identify additional characteristics which could be incorporated into future refinements of autism diagnostic tools such as the ADOS.

The greater age on average of participants in Cohort 2 also suggests some items in the GABS should be adapted depending on participant age. Romantic relationships,

and gendered differences in the experience of these, came up multiple times during adult assessments in Cohort 2, but were not included in the GABS coding frame. Gender differences in experience of and interest in romantic relationships have been previously reported in autistic adults (Strunz et al., 2016), suggesting codes measuring these could provide additional means of assessing the non-traditional phenotype in older individuals. Future adaptations of the GABS could add age-appropriate codes to capture the non-traditional phenotype across development.

Implications

This study demonstrates that the GABS can be used as a tool for secondary coding of ADOS data, where ADOS assessments have been video recorded and there is permission to share them. The GABS therefore increases the research utility of existing data, and allows for the testing of hypotheses around gender differences in autistic behaviours, using already-collected data. Training on the GABS has been demonstrated to be reliable even when performed remotely, meaning research teams across different countries can collaborate and share knowledge. The GABS therefore enables research to be conducted more easily, efficiently, and cost-effectively.

In addition, the GABS has potential to be used as part of autism diagnostic assessments, potentially identifying expressions of autistic behaviours which may not be assessed using traditional tools. Importantly, this can be done without requiring additional assessments in what is already a lengthy diagnostic process. Our preliminary evidence suggests that the GABS is suitable for assessments with children and adults, although further research is needed to pilot the GABS across a range of ages, to ensure all items are age-appropriate.

Acknowledgements

The authors would like to thank Hannah Lee for her hard work coding the GABS.

This work was supported in part by the Simons Foundation Autism Research Initiative (grant #'s 286756 and 454555), the Robert J. and Nancy D. Carney Institute for Brain Sciences at Brown University, the Norman Prince Neurosciences Institute at Lifespan, and the Hassenfeld Child Health Innovation Institute at Brown University. Part of this work was conducted by RL when she was funded by Health Education England.

Conflict of interest

The authors declare they have no conflicts of interest.

Availability of data and materials

The datasets used during the current study are not publicly available due to patient confidentiality as participants did not consent to data to be shared outside the research team.

References

American Psychiatric Association. (2013). *DSM 5*. American Psychiatric Association.

Bargiela, S., Steward, R., & Mandy, W. (2016). The experiences of late-diagnosed women with Autism Spectrum Conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders*, *46*(10), 3281–3294. https://doi.org/10.1007/s10803-016-2872-8

Begeer, S., Mandell, D., Wijnker-Holmes, B., Venderbosch, S., Rem, D., Stekelenburg, F., & Koot, H. M. (2013). Sex differences in the timing of identification among children and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*, 1151–1156.

Beggiato, A., Peyre, H., Maruani, A., Scheid, I., Rastam, M., Amsellem, F., Gillberg, C. I.,

Leboyer, M., Bourgeron, T., Gillberg, C., & Delorme, R. (2017). Gender differences in autism spectrum disorders: Divergence among specific core symptoms. *Autism Research*, *10*(4), 680–689. https://doi.org/10.1002/aur.1715

Bishop, S. L., Farmer, C., & Thurm, A. (2015). Measurement of Nonverbal IQ in Autism Spectrum Disorder: Scores in Young Adulthood Compared to Early Childhood. *Journal of Autism and Developmental Disorders*, *45*(4), 966–974. https://doi.org/10.1007/s10803-014-2250-3

Cage, E., & Troxell-Whitman, Z. (2019). Understanding the Reasons, Contexts and Costs of Camouflaging for Autistic Adults. *Journal of Autism and Developmental Disorders*, *49*(5), 1899–1911. https://doi.org/10.1007/s10803-018-03878-x

Cassidy, S. A., Gould, K., Townsend, E., Pelton, M., Robertson, A. E., & Rodgers, J. (2019). Is Camouflaging Autistic Traits Associated with Suicidal Thoughts and Behaviours? Expanding the Interpersonal Psychological Theory of Suicide in an Undergraduate Student Sample. *Journal of Autism and Developmental Disorders*, *50*(10), 3638–3648. https://doi.org/10.1007/s10803-019-04323-3

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220. https://doi.org/https://doi.org/10.1037/h0026256

Dean, M., Kasari, C., Shih, W., Frankel, F., Whitney, R., Landa, R., Lord, C., Orlich, F., King, B., & Harwood, R. (2014). The peer relationships of girls with ASD at school: comparison to boys and girls with and without ASD. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *55*(11), 1218–1225. https://doi.org/10.1111/jcpp.12242

Duvekot, J., Van der Ende, J., Verhulst, F. C., Slappendel, G., Van Daalen, E., Maras, A., & Greaves-Lord, K. (2017). Factors influencing the probability of a diagnosis of Autism Spectrum Disorder in girls versus boys. *Autism*, *21*(6), 646–658. https://doi.org/10.1177/1362361316672178

Dworzynski, K., Ronald, A., Bolton, P., & Happé, F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(8), 788–797. https://doi.org/10.1016/j.jaac.2012.05.018

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and cognitive characteristics of females and males with autism in the Simons Simplex Collection. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(3), 329-40.e1-3. https://doi.org/10.1016/j.jaac.2013.12.004

Frazier, T. W., & Hardan, A. Y. (2017). Equivalence of symptom dimensions in females and males with autism. *Autism*, *21*(6), 749–759. https://doi.org/10.1177/1362361316660066

Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the behavioral sciences* (10th ed.). Cengage Learning.

Grove, R., Hoekstra, R. A., Wierda, M., & Begeer, S. (2018). Special interests and subjective wellbeing in autistic adults. *Autism Research*, *11*(5), 766–775. https://doi.org/10.1002/aur.1931

Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: an examination of developmental functioning, autistic symptoms, and coexisting behavior problems in toddlers. *Journal of Autism and Developmental Disorders*, *39*(12), 1715–1722. https://doi.org/10.1007/s10803-009-0810-8

Head, A. M., McGillivray, J. A., & Stokes, M. A. (2014). Gender differences in emotionality and sociability in children with autism spectrum disorders. *Molecular Autism*, *5*(19). https://doi.org/doi:10.1186/2040-2392-5-19

Hiller, R. M., Young, R. L., & Weber, N. (2014). Sex differences in autism spectrum

disorder based on DSM-5 criteria: evidence from clinician and teacher reporting. *Journal of Abnormal Child Psychology*, *42*(8), 1381–1393. https://doi.org/10.1007/s10802-014-9881-x

Holliday Willey, L. (2015). *Pretending to be Normal: Living with Asperger's Syndrome* (Expanded E). Jessica Kingsley Publishers. https://books.google.it/books?hl=it&lr=&id=Q7RNBAAAQBAJ&oi=fnd&pg=PA4&dq=imaginary+friends+and+autism+spectrum+disorders&ots=PNAqWTY99t&sig=d_A74_xG4eXQociMMGlRy9IFY7A#v=onepage&q=imaginary friends and autism spectrum disorders&f=false

Hull, L., Mandy, W., & Petrides, K. V. (2017). Behavioural and cognitive sex/gender differences in autism spectrum condition and typically developing males and females. *Autism*, *21*(6), 706–727. https://doi.org/10.1177/1362361316669087

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., & Mandy, W. (2017). "Putting on My Best Normal": Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, *47*(8), 2519–2534. https://doi.org/10.1007/s10803-017-3166-5

Hull, L, Lai, M.-C., Baron-Cohen, S., Allison, C., Smith, P., Petrides, K. V, & Mandy, W. (2020). Gender differences in self-reported camouflaging in autistic and non-autistic adults. *Autism*, *24*(2), 352–363. https://doi.org/10.1177/1362361319864804

Hull, Laura, Mandy, W., Lai, M. C., Baron-Cohen, S., Allison, C., Smith, P., & Petrides, K. V. (2018). Development and Validation of the Camouflaging Autistic Traits Questionnaire (CAT-Q). *Journal of Autism and Developmental Disorders*, *49*(3), 819–833. https://doi.org/10.1007/s10803-018-3792-6

Jamison, R., Bishop, S. L., Huerta, M., & Halladay, A. K. (2017). The clinician perspective on sex differences in autism spectrum disorders. *Autism*, *21*(6), 772–784. https://doi.org/10.1177/1362361316681481

Kanfiszer, L., Davies, F., & Collins, S. (2017). "I was just so different": The experiences of

women diagnosed with an autism spectrum disorder in adulthood in relation to gender and social relationships. *Autism : The International Journal of Research and Practice*, *21*(6), 661–669. https://doi.org/10.1177/1362361316687987

Kopp, S., & Gillberg, C. (1992). Girls with social deficits and learning problems: Autism, atypical Asperger syndrome or a variant of these conditions. *European Child & Adolescent Psychiatry*, *1*(2), 89–99. https://doi.org/10.1007/BF02091791

Lai, M.-C., Kassee, C., Besney, R., Bonato, S., Hull, L., Mandy, W., Szatmari, P., & Ameis, S. H. (2019). Prevalence of co-occurring mental health diagnoses in the autism population: A systematic review and meta-analysis. *Lancet Psychiatry*, *6*(10), 819–829. https://doi.org/10.1016/S2215-0366(19)30289-5

Lai, M.-C., Lombardo, M. V., Ruigrok, A. N. V., Chakrabarti, B., Auyeung, B., Szatmari, P., Happé, F., Baron-Cohen, S., & Consortium, M. A. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism*, *21*(6), 690–702. https://doi.org/10.1177/1362361316671012

Lai, M.-C., Lombardo, M. V, Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/gender differences and autism: setting the scene for future research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(1), 11–24. https://doi.org/10.1016/j.jaac.2014.10.003

Lai, M.-C., Lombardo, M. V, Pasco, G., Ruigrok, A. N. V, Wheelwright, S. J., Sadek, S. A., Chakrabarti, B., & Baron-Cohen, S. (2011). A behavioral comparison of male and female adults with high functioning autism spectrum conditions. *PloS One*, *6*(6), e20835. https://doi.org/10.1371/journal.pone.0020835

Little, L. M., Wallisch, A., Salley, B., & Jamison, R. (2017). Do early caregiver concerns differ for girls with autism spectrum disorders? *Autism*, *21*(6), 728–732. https://doi.org/10.1177/1362361316664188

Livingston, L. A., Shah, P., Milner, V., & Happé, F. (2020). Quantifying compensatory

strategies in adults with and without diagnosed autism. *Molecular Autism*, 7–10.

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6). https://doi.org/10.1016/j.jaac.2017.03.013

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule Second Edition (ADOS-2) Manual (Part 1): Modules 1–4.* Western Psychological Services.

Lord, Catherine, Risi, S., Lambrecht, L., Cook, E. H. J., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders*, *42*(7), 1304–1313. https://doi.org/10.1007/s10803-011-1356-0

May, T., Cornish, K., & Rinehart, N. J. (2012). Gender Profiles of Behavioral Attention in Children With Autism Spectrum Disorder. *Journal of Attention Disorders*, *20*(7), 627–635. https://doi.org/10.1177/1087054712455502

McCormick, C. E. B., Kavanaugh, B. C., Sipsock, D., Righi, G., Oberman, L. M., Moreno De Luca, D., Gamsiz Uzun, E. D., Best, C. R., Jerskey, B. A., Quinn, J. G., Jewel, S. B., Wu, P. C., McLean, R. L., Levine, T. P., Tokadjian, H., Perkins, K. A., Clarke, E. B., Dunn, B., Gerber, A. H., … Morrow, E. M. (2020). Autism Heterogeneity in a Densely Sampled U.S. Population: Results From the First 1,000 Participants in the RI-CART Study. *Autism Research*, *13*(3), 474–488. https://doi.org/10.1002/aur.2261

Oswald, T. M., Winter-Messiers, M. A., Gibson, B., Schmidt, A. M., Herr, C. M., & Solomon, M. (2016). Sex Differences in Internalizing Problems During Adolescence in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *46*(2), 624–636. https://doi.org/10.1007/s10803-015-2608-1

Rogers, C. L., Goddard, L., Hill, E. L., Henry, L. A., & Crane, L. (2016). Experiences of diagnosing autism spectrum disorder: A survey of professionals in the United Kingdom. *Autism*, *20*(7), 820–831. https://doi.org/10.1177/1362361315611109

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. In *Journal of Clinical Child and Adolescent Psychology* (Vol. 37, Issue 1). https://doi.org/10.1080/15374410701817808

Schuck, R. K., Flores, R. E., & Fung, L. K. (2019). Brief Report: Sex/Gender Differences in Symptomology and Camouflaging in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*, 2597–2604. https://doi.org/10.1007/s10803-019-03998-y

Sedgewick, F., Crane, L., Hill, V., & Pellicano, E. (2018). Friends and lovers: the relationships of autistic and neurotypical women. *Autism in Adulthood*, *1*(2).

Sedgewick, F., Hill, V., Yates, R., Pickering, L., & Pellicano, E. (2015). Gender Differences in the Social Motivation and Friendship Experiences of Autistic and Non-autistic Adolescents. *Journal of Autism and Developmental Disorders*, *46*(4), 1297–1306. https://doi.org/10.1007/s10803-015-2669-1

Simone, R. (2010). *Aspergirls: Empowering females with Asperger syndrome*. Jessica Kingsley Publishers.

Sparrow, S. S., & Cicchetti, D. V. (1985). Diagnostic uses of the vineland adaptive behavior scales. *Journal of Pediatric Psychology*, *10*(2), 215–225. https://doi.org/10.1093/jpepsy/10.2.215

Strunz, S., Schermuck, C., Ballerstein, S., Ahlers, C. J., Dziobek, I., & Roepke, S. (2016).

Romantic Relationships and Relationship Satisfaction Among Adults With Asperger Syndrome and High-Functioning Autism. *Journal of Clinical Psychology*, *73*(1), 113–125. https://doi.org/10.1002/jclp.22319

Van Wijngaarden-Cremers, P. J. M., van Eeten, E., Groen, W. B., Van Deurzen, P. A., Oosterling, I. J., & Van der Gaag, R. J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: a systematic review and meta-analysis. *Journal of Autism and Developmental Disorders*, *44*(3), 627–635. https://doi.org/10.1007/s10803-013-1913-9

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children (WISC-IV)*. Psychological Corporation.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine*, *11*, 115–129. https://doi.org/10.1017/S0033291700053332

Wood-Downie, H., Wong, B., Kovshoff, H., Mandy, W., Hull, L., & Hadwin, J. A. (2020). Sex/Gender Differences in Camouflaging in Children and Adolescents with Autism. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-020-04615-z

World Health Organization. (2018). *The ICD-11. International Classification of Diseases for Mortality and Morbidity Statistics Eleventh Revision*. https://doi.org/https://icd.who.int/

Zeidner, M., Matthews, G., Roberts, R. D., & MacCann, C. (2003). Development of emotional intelligence : Towards a multi-level investment model. *Human Development*, *46*, 2–3.

GENDERED AUTISM BEHAVIORAL SCALE

Table 2. *Categories and descriptions of items included in the GABS*

Broad Category	Items	Description
A. Social Adaptation	A1. Camouflaging (masking)	Reports that autistic traits are actively hidden or suppressed. This item is unique to the GABS and does not have a clear corollary on the ADOS.
	A2. Camouflaging (compensating)	Reports copying or imitating others so as not to appear different from peers. This item is unique to the GABS and does not have a clear corollary on the ADOS.

	A3. Self-reflection on social behavior	Shows awareness of the impact of their behavior as well as differences between self and others. Like item B6 of ADOS Module 3 and item B7 of Module 4 (Insight into Typical Social Situations and Relationships), this item assesses the ability to reflect on social interactions with others. Additionally, this item assesses respondents' insight into their *own* social behavior in comparison to the social behavior of others.
B. Social Relationships	B1. Reported friendships	Reports one or more friendships. Whereas item B6 of ADOS Module 3 and item B7 of Module 4 (Insight into Typical Social Situations and Relationships) codes for respondents' *understanding* of friendships, this item codes for whether a respondent reports having friendships.
	B2. Understanding of friendships	Shows superficial or meaningful understanding of components of friendships. Like item B6 of ADOS Module 3 and item B7 of Module 4 (Insight into Typical Social Situations and Relationships), this item assesses respondents' insight into social relationships. To receive full credit for this item, respondents must display a higher degree of insight into friendships than is needed on the ADOS.

	B3. Quality of reported friendships	Evidence of mutual affection in reported friendships via shared interests, experiences, regular contact, etc. Though this item has some content overlap with item B6 of ADOS Module 3 and item B7 of Module 4 (Insight into Typical Social Situations and Relationships), respondents cannot receive credit for this item unless they report having clear friendships (i.e. beyond listing names of coworkers or schoolmates).
	B4. Age of friends	Reports friends that are not markedly older or younger. This item is unique to the GABS and does not have a clear corollary on the ADOS.
	B5. Nature of friendships	Reports having 1) a mixture of friendships, 2) one or two intense friendships, 3) being an "outsider." This item is unique to the GABS and does not have a clear corollary on the ADOS.

	B6. Responding to conversational cues	Responds appropriately to conversational cues (i.e. cliffhangers) provided by the examiner. Similar to Item A6 on ADOS Modules 3 & 4 (Asks for Information), this item assesses respondents' spontaneous inquiries into the examiner's thoughts, feelings, and experiences. It also assesses for respondents' ability to appropriately pick up on social cues from the examiner.
	B7. Social interest	Actively seeks and explicitly states enjoyment in social activities. This item has some parallels to item B4 on ADOS Modules 3 & 4 (Shared Enjoyment in Interaction). However, rather than assessing the respondents' enjoyment while interacting with the examiner, this item codes for offers of information and other indicators that the respondent enjoys interacting with others.

Table 2. *Categories and items included in the GABS (continued)*

Broad Category	Items	Description

C. Managing Emotions	C1. Internalizing difficulties	Evidence of depressive or anxious symptoms during the assessment. This item overlaps somewhat with item E3 (Anxiety) on ADOS Modules 3 & 4; in addition to assessing anxiety, this item also codes for indicators of depression.
	C2. Externalizing difficulties	Evidence of aggressive or hyperactive symptoms during the assessment Like items E1 (Overactivity/Agitation) and E2 (Tantrums, Aggression, Negative or Disruptive Behavior) on ADOS Modules 3 & 4, this item also codes for indicators of externalizing symptoms.
	C3. Communicating about emotions	Describes various emotional states in detail and without visible difficulty or discomfort. In addition to assessing insight into emotions—like item B5 on ADOS Module 3 (Comments on Others' Emotions/Empathy) and Module 4 (Communication of Own Affect)—this item also captures indicators of discomfort in respondents while discussing this subject matter with the examiner.

	C4. Influence of social acceptance or rejection on emotions	Describes being emotionally impacted by one or more instances of social acceptance or rejection. This item is unique to the GABS and does not have a clear corollary on the ADOS.
	C5. Violations of sameness or rigidness on emotions	Describes being emotionally impacted by one or more instances of violations of sameness. This item is unique to the GABS and does not have a clear corollary on the ADOS.
D. Interests & Hobbies	D1. Reported interests (frequency/intensity)	Reports one or more interests with appropriate or inappropriate frequency and intensity. Like ADOS Module 3 & 4 item D4 (Excessive Interest in or References to Unusual or Highly Specific Topics or Objects or Repetitive Behaviors), this item codes for the presence of unusual preoccupations and/or circumscribed interests. However, unlike the ADOS, the coding for this item also captures appropriate expressions of interest in topics and hobbies.

	D2. Type of interests (quality/nature)	Majority of reported interests are social in nature or are non-social in nature. This item is unique to the GABS and does not have a clear corollary on the ADOS.

Table 7. *Probes for ADOS observations/tasks based on the GABS. Note all ADOS-2 items are numbered and described based on Module 4 of the ADOS-2 (Hus & Lord, 2014).*

ADOS-2 Observation	Suggested probe	Corresponding GABS item(s)
6. Social Difficulties & Annoyance	*"Have you ever tried to change anything about yourself to fit in with other people?"*	Social Adaptation A1-3
6. Social Difficulties & Annoyance	If reports being bullied or teased: *"How did you feel when that happened?"*	Managing Emotions C4
12. Friends, Relationships & Marriage	When describing friendships: *"Would you say that you have a best friend or friends? Is there anyone you are friendly with, but not that close to?"*	Social Relationships B5
7. Emotions	To follow 'happiness' question: *"Do you have any hobbies or interests that make you feel happy?"*	Interests & Hobbies D1-2
7. Emotions	To follow 'frightened or anxious' question: *"How often do you feel that way?"*	Managing Emotions C1
7. Emotions	To follow 'sad' question: *"How often do you feel that way?"* or *"Do you ever feel hopeless, or that good things don't ever happen to you?"*	Managing Emotions C1