Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-018-3819-z

ORIGINAL PAPER

Gender Differences and Similarities: Autism Symptomatology and Developmental Functioning in Young Children

Maya Matheis[1] · **Johnny L. Matson**[1] · **Esther Hong**[1] · **Paige E. Cervantes**[2]

Abstract

A growing body of research suggests that symptoms of autism spectrum disorder (ASD) may present differently in males and females. This study examined gender differences in ASD symptoms and developmental functioning, using the *Baby and Infant Screen for Children with aUtism Traits, Part 1* (BISCUIT-Part 1) and the *Battelle Developmental Inventory, 2nd Edition* (BDI-2), amongst children aged 17–37 months meeting ASD diagnostic criteria ($n = 1317$). No gender differences were found in regards to overall symptom severity or symptom domains on the BISCUIT-Part 1 when gender groups were matched by cognitive ability. Females with ASD had greater motor deficits and less communication impairment compared to their male counterparts as measured by the BDI-2. Secondary analyses examining item endorsement patterns were also conducted. Implications of the findings are discussed.

Keywords Autism spectrum disorder · Gender differences · Autism symptomatology · Early identification

There is a well-documented male predominance in autism spectrum disorder (ASD), with an estimated male-to-female ratio of approximately 4:1 (Baio et al. 2018; Fombonne et al. 2011; Hill et al. 2016). While the skewed gender ratio may be related to underlying biological mechanisms, researchers have expressed concerns that females with ASD may be under-recognized and under-diagnosed (Gould and Ashton-Smith 2011; Haney 2015). Females with ASD have also been found to have a greater delay in diagnosis (Manning et al. 2011; Rosenberg et al. 2011; Shattuck et al. 2009) and to experience more difficulty in the diagnostic process compared to their male counterparts (Siklos and Kerns 2007). Such delays are of considerable concern given the association of early ASD-specific intervention with improved long-term prognosis (Zwaigenbaum et al. 2015). Further, estimates of the male-to-female ratio among individuals with ASD and accompanying cognitive impairment are substantially reduced, suggesting that females with high-functioning ASD may be particular risk for underdiagnosis (Begeer et al. 2012; Van Wijngaarden-Cremers et al. 2014).

The vast majority of research on autism symptomology has been based on a male phenotype, with many research studies not including enough female participants to accurately perform gender comparisons (Kirkovski et al. 2013; Rivet and Matson 2011). While this gender disparity in research may be largely due to the male predominance of ASD and the difficulty of recruiting female participants, it remains that the female phenotype of ASD is not well researched or understood. The male-bias in research has likely contributed to gendered diagnostic bias (Gould and Ashton-Smith 2011; Haney 2015). Commonly used diagnostic instruments such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 1999) and the Autism Diagnostic Interview-Revised (ADI-R; Rutter et al. 2003) have been found to produce significant gender differences in scores and may be less sensitive in capturing females with ASD (Adamou et al. 2018; Duvekot et al. 2017; Mussey et al. 2017).

Efforts to examine the female phenotype of ASD have indicated that girls with ASD demonstrate stronger social and play skills compared to boys with ASD (Attwood 2007; Bacon et al. 1998; Baron-Cohen et al. 2011; Bauminger et al. 2008; Brown and Dunn 1996; Koenig and Tsatsanis 2005; McClure 2000; Nydén et al. 2000; Rynkiewicz et al. 2016), with some research suggesting that girls with

✉ Maya Matheis
maya.matheis@gmail.com

[1] Department of Psychology, Louisiana State University, 236 Audubon Hall, Baton Rouge, LA 70803, USA

[2] Child Study Center, Department of Child and Adolescent Psychiatry, Hassenfeld Children's Hospital, NYU Langone Health, New York, NY, USA

Published online: 15 November 2018

high-functioning ASD exhibit compensatory behaviors (e.g., staying close to peers, joint engagement) that may mask their social deficits (Dean et al. 2017). However, research findings are mixed in regards to gender differences in restricted, repetitive behaviors and interests (RRBs). Several studies have reported no significant gender differences in RRBs (Banach et al. 2009; Holtmann et al. 2007; Szatmari et al. 2006). However, other studies have found that females with ASD exhibit less repetitive use of objects, preoccupation with parts of objects, and adherence to rituals in comparison to males with ASD (Hiller et al. 2014; Nicholas et al. 2008). Further, females with ASD, without cognitive impairment, have been found to have lower overall level of RRBs (Frazier and Hardan 2017), suggesting that cognitive functioning may be associated gender differences in symptomology.

Despite the strong focus within the field on early diagnosis and early intervention, our understanding of gender differences in ASD presentation in toddlers and preschool-aged children remains limited with mixed findings across the literature. For example, several studies have found no significant gender differences in core ASD symptom severity, cognitive ability, social communication, or RRBs (Andersson et al. 2013; Carter et al. 2007; Lawson et al. 2018; Postorino et al. 2015; Reinhardt et al. 2014), while others suggest that, compared to their male counterparts, young girls with ASD exhibit significantly fewer RRBs (Hartley and Sikora 2009), as well as more impairments in social affect (Lawson et al. 2018), motor development, adaptive behavior, and emotional development (Carter et al. 2007).

In a study of toddlers with ASD, Sipes et al. (2011) examined gender differences in ASD symptom endorsements while accounting for developmental functioning in a subsample of at-risk toddlers ($n = 390$) enrolled in a statewide early intervention program. This study evaluated gender differences in ASD symptom severity and symptom endorsement across ASD symptom domains, while stratifying by developmental quotient (DQ). Overall, findings indicated that there were no significant gender differences in ASD symptom severity, with the exception of the RRB domain, in which female toddlers with average DQ reported significantly fewer symptom endorsements compared to their male counterparts.

Although the area of gender differences in ASD has received increasing attention in research and clinical practice over the past several years, gender differences in the qualitative symptom presentation and diagnosis of ASD have not been sufficiently addressed (Koenig and Tsatsanis 2005; Rutter et al. 2003). Consequently, results from the existing literature in regards to young children are mixed and no firm conclusions have been made. In particular, there is a growing need for studies using large, representative samples with sizable female subsamples (Andersson et al. 2013; Lawson et al. 2018; Reinhardt et al. 2014). To address this gap in the literature, the current study replicated the Sipes et al. (2011) study by evaluating gender differences in ASD symptom endorsement while accounting for cognitive functioning, using a larger, population-level sample of toddlers enrolled in a statewide early intervention program. Stratifying groups based upon cognitive impairment is important given evidence of distinct gender differences found in high- versus low-functioning samples (Frazier and Hardan 2017; McLennan et al. 1993; Sipes et al. 2011). The present study further expanded upon the research conducted by Sipes et al. (2011) by investigating gender differences in developmental functioning among toddlers with and without ASD, in regards to overall developmental functioning and across developmental domains (i.e., adaptive, personal-social, communication, motor, cognitive). The use of a non-ASD sample allows for the evaluation of uniqueness of the results to the ASD population and is often lacking in this area of study (Lai et al. 2014).

Methods

Participants

Participants included in the current study were enrolled in EarlySteps, Louisiana's early intervention program under the Individuals with Disabilities Education Act, Part C. To be eligible for EarlySteps, children must be under the age of 36 months and have developmental delays or at risk for developmental delays (i.e., due to a medical condition). Participants were sampled from a pre-existing research database. The original sample included 17,838 children screened by EarlySteps between 2008 and 2017. Cases with missing relevant data or errors ($N = 349$) were removed from the sample, as were cases that declined the ASD screening ($N = 6562$). No differences in age or race/ethnicity were found between cases that declined or consented to ASD screening. The ratio of males:females (M:F) screened for ASD was found to be significantly higher amongst cases screened for ASD compared to cases that declined ASD screening, with females making up 36.5% of declined cases and 31.5% of screened cases, $X^2\ (1) = 55.367$, $p < .001$. The final sample consisted of 10,927 participants, including 1317 who met *DSM-5* criteria for ASD and 9610 who did not meet *DSM-5* criteria for ASD.

Participant characteristics of cases meeting ASD criteria are reported in Table 1 ($N = 1317$). The mean age of this subsample at the time of assessment was 26.01 months ($SD = 4.56$). The ASD subsample was 72.2% male and 27.8% female. In regard to race/ethnicity, 45.6% of the ASD subsample was identified as white, 41.8% as African American, and 10.1% as another racial/ethnic identity. Hispanic identity was included in the "other" racial/ethnic identity

Table 1 Demographic information for cases meeting ASD criteria ($N=1317$) and by group

	Total ASD ($N=1317$)	Male w/o Cog Delay ($N=462$, 35.1%)	Female w/o Cog Delay ($N=149$, 11.3%)	Male w/ Cog Delay ($N=542$, 41.2%)	Female w/ Cog Delay ($N=164$, 12.5%)
Gender [N (%)]					
Male	1004 (76.2%)	462 (100%)	0 (0%)	542 (100%)	0 (0%)
Female	313 (23.8%)	0 (0%)	149 (100%)	0 (0%)	164 (100%)
Age in months					
M (SD)	26.01 (4.56)	25.92 (4.29)	25.71 (4.56)	26.31 (4.75)	25.58 (4.67)
Range	17–37	17–36	17–35	17–37	17–35
Race/ethnicity [N (%)]					
White	600 (45.6%)	228 (49.4%)	62 (41.6%)	239 (44.1%)	71 (43.3%)
African American	550 (41.8%)	184 (39.8%)	68 (45.6%)	227 (41.9%)	71 (43.3%)
Other	133 (10.1%)	38 (8.2%)	15 (10.1%)	60 (11.1%)	20 (12.2%)
Missing data	34 (2.6%)	12 (2.6%)	4 (2.7%)	16 (2.9%)	2 (1.2%)

category due to low representation. The ASD cases were classified into groups based on cognitive domain scores from the *Battelle Developmental Inventory, 2nd Edition* (BDI-2) and gender. Stratification based on cognitive DQ was done in an effort to examine gender differences in symptom presentation between children with and without cognitive impairment. Participants with a cognitive DQ level of 70 or below were classified as having a cognitive delay and those with a cognitive DQ level of 71 or above were classified has having no cognitive delay. This resulted in four ASD groups: (1) males without cognitive delay (Males w/o Cog Delay), (2) males with cognitive delay (Males w/Cog Delay), (3) females without cognitive delay (Females w/o Cog Delay), and (4) females with cognitive delay (Females w/Cog Delay).

Participant characteristics of cases that did not meet ASD criteria are reported in Table 2 ($N=9610$). The non-ASD subsample was 67.3% male and 32.7% female. The mean age of this subsample at the time of assessment was 25.21 months ($SD=4.61$). In regards to race/ethnicity, 36.5% of this subsample were identified as white, 51.3% as African American, and 9.7% as another racial/ethnic identity.

Measures

The *Baby and Infant Screen for Children with aUtism Traits, Part 1* (BISCUIT-Part 1; Matson et al. 2007) is an informant-based assessment tool used to assess symptoms of ASD in children between the ages of 17–37 months. The BISCUIT-Part 1 is comprised of 62 items scored on a 3-point Likert scale. Parents and/or caregivers are instructed to rate each item with respect to how their child compares to a typically-developing child of the same age as: "0" (not different, no impairment); "1" (somewhat different; mild impairment); or "2" (very different; severe impairment). To help informants distinguish what typical development may look like, the BISCUIT includes an appendix that provides typical and atypical characteristics related to each item as well as examples of related behaviors. Each item score is summed to produce a total BISCUIT-Part 1 score. A total cut-off

Table 2 Demographic information for cases that did not meet ASD criteria ($N=9610$) by gender

	Total No ASD ($N=9610$)	Male ($N=6465$)	Female ($N=3145$)
Gender [N (%)]			
Male	6465 (67.3%)	6465 (100%)	0 (0%)
Female	3145 (32.7%)	0 (0%)	3145 (100%)
Age in months			
M (SD)	25.21 (4.61)	25.36 (4.62)	24.90 (4.58)
Range	17–37	17–37	17–36
Race/ethnicity [N (%)]			
White	3515 (36.5%)	2286 (35.4%)	1229 (39.1%)
African American	4929 (51.3%)	3364 (52.0%)	1565 (49.8%)
Other	925 (9.7%)	639 (9.9%)	289 (9.1%)
Missing data	241 (2.5%)	176 (2.7%)	65 (2.0%)

score of 17 has been found to differentiate toddlers with ASD from toddlers with atypical development and no ASD (Matson et al. 2010). Previous studies have demonstrated that the BISCUIT has good internal reliability ($r = .97$) and a classification rate of .88 (Matson et al. 2009, 2011). As an autism screen, convergent validity was established with the *Modified CHecklist for Autism in Toddlers* (M-CHAT) and was found to be high ($r = 0.80$; Matson et al. 2011). The BISCUIT also includes a demographic form that gathers information on age, gender, ethnicity, parental concerns regarding development, developmental milestones, medical history, and relevant informant/family information. For the current study, the demographic form was used to collect information regarding the child's gender, age, and race/ethnicity.

The *Battelle Developmental Inventory, 2nd Edition* (BDI-2; Newborg 2005) is a comprehensive assessment tool used to assess developmental skills in children aged birth to 7 years 11 months. The BDI-2 includes five developmental domains (i.e., Adaptive, Personal-Social, Communication, Motor, and Cognitive). The Adaptive domain includes Self-Care and Personal Responsibility subtests. The Personal-Social domain assesses social skills through Adult Interaction, Peer Interaction, and Self-Concept and Social Role subtests. The Communication domain is made up of Receptive and Expressive Communication subtests. The Motor domain evaluates Gross, Fine, and Perceptual Motor subdomains, and the Cognitive domain consists of Attention and Memory, Reasoning and Academic Skills, and Perception and Concepts subtests (Newborg 2005; Stone-MacDonald et al. 2018).

The BDI-2 offers flexibility in administration, which is beneficial for assessing young children with developmental delays and ASD; administration procedures (i.e., structured, observation, and interview procedures) vary across testing items with some opportunity to accommodate the demographic and clinical variables of the child being assessed (Stone-MacDonald et al. 2018). Evaluators in this study primarily used structured and interview procedures, as recommended (Stone-MacDonald et al. 2018). Within all procedures, each item is rated on a 3-point Likert scale as: "0" (no ability in the skill), "1" (emerging ability), or "2" (ability at the skill). The scaled scores of each of the five domains are used to establish domain DQs, which are then combined to provide a total DQ. Test–retest reliability for the BDI-2 was found to be > .80 for the total score and for each of the five domain scores (Newborg 2005). Notably, the BDI-2 was found to have moderate to strong correlations with scores from several well-established developmental scales, including the *Bayley Scales of Infant Development, 2nd Edition* (BSID-II), the *Denver Developmental Screening Test, 2nd Edition* (DDST-II), the *Preschool Language Scales, 4th Edition* (PLS-4), and the *Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition* (WPPSI-III; Newborg 2005). Further, the validity of the BDI-2 has been demonstrated for children with different types of disorders, such that the measure differentiates between children with ASD, developmental delays, speech and language delays, and motor delays (Newborg 2005).

For the current study, the Cognitive DQ was used to determine if the participant had a cognitive delay using the recommended cutoff score of two standard deviations below the mean (≤ 70; Newborg 2005). While a strong correlation has been found between Cognitive DQ and full scale intellectual quotient (IQ) on the WPPSI-III (Moyal 2010; Newborg 2005), these and other measures administered in toddlerhood have been criticized for limited long term predictive value given the dynamic nature of early childhood development (Sonnander 2007). However, researchers have shown that individuals with significantly low scores on developmental measures (e.g., DQ ≤ 70) demonstrate less fluctuation in scores across time (Sonnander 2007); therefore, it is likely that the classification of "cognitive delay" in this study may be indicative of long-term prognoses.

Procedure

The use of EarlySteps records for research purposes was approved by the Louisiana State University Institutional Review Board and the State of Louisiana's Office for Citizens with Developmental Disabilities (OCDD). Institutional Review Boards determined that informed consent to participate in research was not required as the study utilized archival data, and as personal identifiers (e.g., name, date of birth) were removed from the database by OCDD prior to receipt.

All children enrolled in EarlySteps receive a standardized assessment battery, which includes the BDI-2. The assessment battery is administered upon entry into the program, annually thereafter, and prior to exiting the program at the age of 36 months. At each assessment timepoint, parents are given the option to undergo an ASD screening, which consists of the BISCUIT-Part 1. All assessments are conducted by EarlySteps service providers, who are trained to administer and score the assessment measures and who are experienced in the evaluation and treatment of young children with developmental delays. EarlySteps service providers hold appropriate degrees and certifications/licensures in various disciplines (e.g., social work, speech and language pathology, psychology, occupational therapy, speech therapy). All interviews and assessments were conducted one-on-one with the child and parent or caregiver, in the child's home or other private setting.

De-identified data from the BISCUIT-Part 1 and BDI-2 were shared for research purposes. EarlySteps assessment records were reviewed by a licensed psychologist with over 20 years' experience in the field. Diagnostic classifications

were made by mapping item results from the BISCUIT-Part 1 and subdomain scores from the BDI-2 onto *DSM-5* ASD diagnostic criteria, with consistent methods used across all cases.

Statistical Analyses

Statistical analyses were conducted using SPSS Statistics Software (Version 24). A Chi square test of independence was conducted to determine if the presence of cognitive delay differed significantly between genders. A priori analyses were used to determine if subsamples and groups differed significantly on demographic variables (i.e., race/ethnicity, age). One-way analysis of variance (ANOVA) was used to assess for differences in total DQ between genders in both subsamples, and to determine if ASD groups differed significantly in regard to mean BISCUIT–Part 1 total scores. One-way multivariate analyses of variance (MANOVAs) were used to determine differences in BDI-2 domain DQs (i.e., Adaptive, Personal-Social, Communication, Cognitive, and Motor DQ) across genders in both subsamples and between ASD groups. Of note, the non-ASD sample was used in these analyses to assess distinctiveness of any gender differences found to the ASD population.

A MANOVA was also used to test for differences in the mean values on the three BISCUIT-Part 1 factors (i.e., Socialization/Nonverbal Communication, Repetitive Behavior/Restricted Interests, Communication) between ASD groups. Post-hoc ANOVAs were used to further examine mean differences. Differences in BISCUIT-Part 1 item endorsement across ASD groups were examined using Chi square tests of independence. The non-ASD sample was not used as a reference group in these analyses given the nature of the symptom clusters assessed.

Results

A priori analyses were conducted to determine if ASD groups differed significantly in regards to gender, age, and race/ethnicity. A Chi square test of independence revealed no significant association between gender and cognitive delay amongst ASD cases, X^2 (1) = 0.242, $p = .623$, with 52.4% of girls and 54.0% of boys with Cognitive DQ ≤ 70. ASD cases were subsequently divided into the four groups previously described (i.e., Males w/o Cog Delay, Females w/o Cog Delay, Males w/Cog Delay, and Females w/Cog Delay). A priori analyses determined that the groups did not differ significantly in regards to race/ethnicity, X^2 (6) = 6.230, $p = .398$, or age at assessment, $F(3, 1316) = 1.537$, $p = .203$.

Further a priori analyses were conducted to compare the ASD and non-ASD subsamples on demographic variables. The non-ASD subsample was found to have a significantly lower M:F ratio (2.05:1) compared to the ASD sample (3.20:1), X^2 (1) = 42.989, $p < .001$. The non-ASD subsample was significantly younger, $F(1, 10925) = 35.468$, $p = < .001$. It also had a significantly different racial/ethnic composition compared to the ASD subsample, X^2 (1) = 16.169, $p = < .001$, in that it had a greater proportion of children identified as white.

Developmental Functioning

Gender differences in developmental functioning were tested by comparing BDI-2 Total DQ in both the ASD and non-ASD subsamples (Table 3). Differences in BDI-2 Total DQ were not examined between ASD groups (i.e., Males w/o Cog Delay, Females w/o Cog Delay, Males w/Cog Delay, and Females w/Cog Delay), as these groups were formed using BDI-2 Cognitive DQ as a determining variable, which is used in the calculation of Total DQ. Differences in BDI-2 domain DQs (i.e., Adaptive, Personal-Social,

Table 3 BDI-2 Total DQ and subdomain DQs by gender amongst ASD ($N = 1317$) and non-ASD cases, with statistical comparison

	ASD cases ($N = 1317$)				Non-ASD cases ($N = 9610$)			
	Total M (*SD*)	Male ($N = 1004$) M (*SD*)	Female ($N = 313$) M (*SD*)	*F*	Total M (*SD*)	Male ($N = 6465$) M (*SD*)	Female ($N = 3145$) M (*SD*)	*F*
BDI-2 total DQ	69.55 (12.04)	69.57 (11.57)	69.51 (13.44)	0.01	84.32 (18.27)	83.44 (18.13)	86.13 (18.41)	45.71***
Adaptive DQ	73.89 (13.05)	74.33 (12.90)	72.48 (13.44)	4.31	85.62 (18.75)	85.10 (20.27)	86.70 (15.08)	15.46***
Personal-social DQ	77.52 (11.79)	77.11 (11.61)	78.83 (12.28)	4.76	91.89 (18.66)	90.95 (17.08)	93.82 (91.89)	50.10***
Communication DQ	63.81 (12.48)	63.11 (11.96)	66.04 (13.81)	11.79**	76.63 (16.80)	75.08 (16.12)	79.84 (79.84)	172.96***
Motor DQ	87.71 (16.31)	88.53 (15.49)	85.05 (18.48)	8.95**	97.74 (23.97)	98.02 (26.43)	97.74 (23.97)	2.52
Cognitive DQ	72.01 (11.28)	72.00 (11.12)	72.04 (11.83)	0.00	82.27 (12.39)	81.67 (12.02)	83.52 (13.01)	47.34***

***$p < .001$; **$p < .01$; *$p < .05$

Communication, Cognitive, and Motor DQ) were examined between genders in both ASD and non-ASD subsamples and between the four ASD groups (Table 4).

Total DQ

Differences in the BDI-2 Total DQ between genders were tested in each subsample with a one-way ANOVA (Table 3). As the assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances ($p < .001$), Welch's ANOVA was used for both subsamples. Amongst cases meeting ASD criteria, males ($M = 69.55$, $SD = 12.04$) and females ($M = 69.51$, $SD = 13.44$) were not found to significantly differ in regards to mean BDI-2 Total DQ, Welch's F (1464.99) = .005, $p = .945$; partial $\eta^2 = .002$. In contrast, non-ASD males ($M = 83.44$, $SD = 18.13$) were found to have significantly lower mean BDI-2 Total DQ compared to non-ASD females ($M = 86.13$, $SD = 18.41$), Welch's F (1,6144.592) = 45.717, $p = < .001$; partial $\eta^2 = .005$.

Developmental Domains

One-way MANOVAs were used to test for gender differences in regards to Adaptive DQ, Personal-Social DQ, Communication DQ, Cognitive DQ, and Motor DQ in the ASD and non-ASD subsamples (Table 3). As the assumption of variance–covariance matrices was violated, as assessed by Box's M test ($p = .013$), Pillai's trace was used in both subsamples.

In the non-ASD subsample, the difference between genders on the combined dependent variables was statistically significant, $F(5, 9582) = 41.101$, $p < .001$; Pillai's Trace = .021; partial $\eta^2 = .021$. Follow-up univariate ANOVAs using Bonferroni adjustments for multiple comparisons revealed significant differences between non-ASD males and females in regards to mean Adaptive DQ ($p < .001$, partial $\eta^2 = .002$), Personal-Social DQ ($p < .001$, $\eta^2 = .005$), Communication DQ ($p < .001$, $\eta^2 = .018$), and Cognitive DQ ($p < .001$, $\eta^2 = .005$), with females having higher scores in each developmental domain. No significant gender differences in the non-ASD subsample were found in mean Motor DQ ($p = .112$, $\eta^2 = .000$).

In the ASD subsample, the difference between genders on the combined dependent variables was statistically significant, $F(5, 1308) = 12.685$, $p < .001$; Pillai's Trace = .046; partial $\eta^2 = .046$. Follow-up univariate ANOVAs using Bonferroni adjustments for multiple comparisons did not reveal significant differences between males and females in regards to Adaptive DQ, Personal-Social DQ, and Cognitive DQ. In regard to the Motor domain, males were found to have a significantly higher functioning ($M = 88.53$, $SD = 15.49$) compared to females ($M = 85.05$, $SD = 18.48$), Welch's $F(1, 457.30) = 8.95$, $p = .003$; partial $\eta^2 = .008$. Conversely, in regard to the Communication domain, males were found to have a significantly lower functioning ($M = 63.11$, $SD = 11.96$) compared to females ($M = 66.04$, $SD = 13.81$), Welch's $F(1, 467.20) = 11.795$, $p = .001$; partial $\eta^2 = .010$.

A follow-up MANOVA (Table 4) was conducted to test for differences across ASD groups (i.e., Males w/o Cog Delay, Females w/o Cog Delay, Males w/Cog Delay, Females w/Cog Delay) in regards to Adaptive DQ, Personal-Social DQ, Communication DQ, and Motor DQ (i.e., omitting Cognitive DQ). As the assumption of variance–covariance matrices was violated, as assessed by Box's M test ($p < .001$), Pillai's trace was used. The differences between ASD groups on the combined dependent variables was statistically significant, $F(4, 3927) = 95.846$, $p < .001$; Pillai's Trace = .68, partial $\eta^2 = .227$. Post-hoc Games-Howell analyses to univariate ANOVAs revealed significant differences

Table 4 BISCUIT (Part 1 total scores and subscale scores) and BDI-2 (Adaptive, Personal-Social, Communication, and Motor DQs) scores by ASD group, with statistical comparison

	Total ($n = 1317$) M (*SD*)	Male w/o Cog Delay ($N = 462$) M (*SD*)	Female w/o Cog Delay ($N = 149$) M (*SD*)	Male w/ Cog Delay ($N = 542$) M (*SD*)	Female w/ Cog Delay ($N = 164$) M (*SD*)	*F*
BDI Subdomains						
Adaptive DQ	73.89 (13.05)	81.83 (12.08)	80.79 (12.13)	67.94 (9.75)	64.98 (9.63)	188.18***
Personal-social DQ	77.52 (11.79)	85.06 (9.82)	86.94 (9.58)	70.34 (8.24)	71.51 (9.58)	290.07***
Communication DQ	63.81 (12.48)	70.54 (13.64)	75.66 (14.00)	56.78 (4.41)	57.36 (5.23)	222.42***
Motor DQ	87.71 (16.31)	99.51 (9.28)	98.27 (9.67)	79.18 (13.45)	73.13 (16.30)	358.81***
BISCUIT						
Part 1 total score	52.88 (20.51)	44.42 (17.02)	44.39 (18.76)	59.85 (19.34)	61.44 (22.64)	77.06***
Socialization	22.35 (10.58)	18.06 (8.86)	17.67 (8.71)	25.77 (19.19)	27.41 (11.48)	78.63***
Communication	11.00 (3.07)	9.80 (3.43)	9.59 (3.27)	12.16 (2.14)	11.78 (2.66)	72.24***
RRB	15.35 (8.43)	13.00 (7.15)	13.04 (8.48)	17.36 (8.48)	17.41 (9.36)	32.18***

Note: ***$p < .001$, ** $p < .01$, * $p < .05$

between ASD groups with and without cognitive delays across all domains, as expected based on group classification. Additional significant group differences were found across the Adaptive, Communication, and Motor domains. In regards to Adaptive DQ, Females w/Cog Delay had significantly lower Adaptive DQ Scores ($M = 64.98$, $SD = 9.32$) compared to Males w/Cog Delay ($M = 67.94$, $SD = 9.75$), $p = .004$. Females w/o Cog Delay ($M = 75.66$, $SD = 14.00$) were found to have significantly higher Communication DQ compared to Males w/o Cog Delay ($M = 70.54$, $SD = 13.64$), $p = .001$. In regards to Motor DQ, Males w/Cog Delay ($M = 79.18$, $SD = 13.45$) were found to have significantly higher scores compared to Females w/Cog Delay ($M = 73.13$, $SD = 16.30$), $p < .001$.

ASD Symptom Severity

Differences in the severity of ASD symptoms across ASD groups were tested with a one-way ANOVA using the BISCUIT-Part 1 total score as a dependent variable (Table 4). As the assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances ($p < .001$), Welch's ANOVA was used. Results indicated significant differences in mean BISCUIT-Part 1 total score across the four ASD groups, Welch's $F\ (3, 420.03) = 77.064$, $p < .001$; partial $\eta^2 = .148$. Games-Howell post hoc analysis revealed that significant differences were found between ASD groups with and without cognitive delay: between Males w/o Cog Delay and Males w/Cog Delay, $p < .001$; between Males w/o Cog Delay and Females w/Cog Delay, $p < .001$; between Males w/Cog Delay and Females w/o Cog Delay, $p < .001$; and between Females w/o Cog Delay and Females w/Cog Delay, $p < .001$. Significantly lower total BISCUIT-Part 1 total scores were found amongst ASD groups without cognitive delay compared to those with cognitive delay. Significant differences were not found between ASD groups matched in regards to cognitive ability.

A one-way MANOVA was used to test for differences across ASD groups in regards to ASD symptom domains, as measured by the three BISCUIT-Part 1 subscales of Socialization, Communication, and RRB (Table 4). As the assumption of variance–covariance matrices was violated, as assessed by Box's M test ($p < .001$), Pillai's trace was used. The differences between the ASD groups on the combined dependent variables was statistically significant, $F(9, 3190.78) = 32.865$, $p < .001$; Pillai's Trace = .210; partial $\eta^2 = .070$. Follow-up univariate ANOVAs using Bonferroni adjustments for multiple comparisons found significant differences between ASD groups on all three subscales: Socialization, $F(3, 429.504) = 78.639$, $p < .001$; partial $\eta^2 = .151$; Communication, Welch's $F(3, 408.078) = 72.242$ $p < .001$; partial $\eta^2 = .144$; and RRB, Welch's $F(3, 418.749) = 32.181$, $p < .001$; partial $\eta^2 = .067$. Games-Howell post hoc analysis revealed the same pattern of group differences found previously for the BISCUIT-Part 1 total score for all three BISCUIT-Part 1 subscales: significant differences between Males w/o Cog Delay and Males w/Cog Delay, $p < .001$; between Males w/o Cog Delay and Females w/Cog Delay, $p < .001$; between Males w/Cog Delay and Females w/o Cog Delay, $p < .001$; and between Females w/o Cog Delay and Females w/Cog Delay, $p < .001$. Significantly lower total BISCUIT-Part 1 subscale scores were found amongst ASD groups without cognitive delay compared to those with cognitive delay. As with the analyses assessing for differences in BISCUIT-Part 1 total score across groups, significant differences were not found between ASD groups matched in regards to cognitive ability.

BISCUIT-Part 1 Item Endorsement

Chi square tests of independence were used to examine differences in frequency of BISCUIT-Part 1 item endorsement between ASD groups (Table 5). For these analyses, an item was considered endorsed if scored either a 1 or a 2 on the three-point Likert scale (i.e., in comparison to other children his/her age, 0 = no different, 1 = different, or 2 = very different). Due to the number of statistical comparisons involved in these analyses, the Benjamin-Hochberg procedure was used to control for false discovery rates at the .05 and .01 levels (Benjamini and Hochberg 1995). As previous analyses examining differences in BISCUIT-Part 1 total score and subscale scores indicated that significant group differences were found between ASD groups with and without cognitive delay, comparisons were made between ASD gender groups with equivalent cognitive ability (i.e., Males w/o Cog Delay compared to Females w/o Cog Delay, Males w/Cog Delay compared to Females w/Cog Delay).

Significant differences in frequency of item endorsement were found in nine BISCUIT-Part 1 items between males and females without cognitive delay. Compared to Males w/o Cog Delay, Females w/o Cog Delay had significantly higher frequency of endorsement of the following items: (2) intellectual abilities; (5) verbal communication; (9) use of language to communicate; (11) reactions to normal, everyday sounds; (54) clumsiness; and (60) respect for others' personal space. Males w/o Cog Delay were found to have significantly higher frequency of endorsement when compared to Females w/o Cog Delay on (34) abnormal preoccupation with the parts of an object or objects; (56) imitation of an adult or child model; and (42) abnormal fascination with the movement of spinning objects.

In regards to children with cognitive delay, significant differences in frequency of item endorsement between males and females were found in eight BISCUIT-Part 1 items. Females w/Cog Delay were found to have significantly higher endorsement on the following items compared

Table 5 BISCUIT-Part 1 item frequency endorsement and results of Chi square analyses (items with significant differences between ASD groups in bold)

BISCUIT-Part 1 item # and description	Male w/o Cog Delay (%)	Female w/o Cog Delay (%)	χ^2	Male w/ Cog Delay (%)	Female w/ Cog Delay (%)	χ^2
1. Communication skills	92.9	94.6	.566	98.2	97.5	.239
2. Intellectual abilities	**54.1**	**65.1**	**5.543***	74.5	81.0	2.854
3. Age appropriate self-help and adaptive skills	60.2	59.1	.058	81.2	87.2	3.166
4. Engages in repetitive motor movements for no reason	60.7	55.4	1.321	78.2	73.8	1.415
5. Verbal communication	**93.7**	**98.0**	**4.127***	98.5	98.2	0.108
6. Prefers foods of a certain texture or smell	46.5	43.0	.583	57.9	60.1	0.247
7. Ability to recognize the emotions of others	50.9	47.7	.466	70.2	70.6	0.008
8. Maintains eye contact	68.8	68.5	.007	73.9	67.5	2.570
9. Use of language to communicate	**89.8**	**96.0**	**5.373***	**98.7**	**96.3**	**3.903***
10. Social interactions with others his/her age	88.1	87.2	.076	90.8	90.9	0.001
11. Reactions to normal, everyday sounds	**41.6**	**53.0**	**5.898***	48.9	51.2	0.273
12. Response to others' social cues	43.7	46.6	.381	68.5	64.0	1.123
13. Reaction to normal, everyday lights	14.9	14.8	.003	20.8	17.1	1.104
14. Peer relationships	81.0	79.9	.085	85.6	84.8	0.073
15. Rhythm of speaking	16.7	18.1	.152	18.2	12.8	2.614
16. Use of language in conversation with others	87.9	90.6	.839	**96.9**	**93.3**	**4.215***
17. Shares enjoyment, interest, or achievements with others	48.8	50.7	.157	75.2	70.7	1.330
18. Ability to keep and make friends	63.6	63.1	.011	68.7	73.2	1.190
19. Interest in participating in social games, sports, and activities	45.8	39.0	2.036	62.8	64.0	0.075
20. Interest in another person's side of the conversation	57.8	50.7	2.324	72.2	73.8	0.154
21. Able to understand the subtle cues or gestures of others	44.6	43.9	.019	65.0	70.1	1.478
22. Use of too few or too many social gestures	36.4	35.1	.083	56.6	62.8	1.964
23. Body posture and/or gestures	22.1	22.8	.031	**44.0**	**56.7**	**8.169****
24. Communicates effectively	76.6	72.5	1.047	93.5	95.1	.549
25. Likes affection	27.1	33.8	2.480	32.3	35.4	.539
26. Displays a range of socially appropriate facial expressions	25.3	24.2	.081	**43.9**	**53.0**	**4.231***
27. Restricted interests and activities	40.7	38.3	.279	59.4	63.8	1.012
28. Motivated to please others	54.9	49.0	1.569	70.7	73.8	.599
29. Eye-to-eye gaze	48.9	45.0	.705	63.0	54.9	3.521
30. Reaction to sounds and sights	36.1	43.9	2.873	44.6	42.7	.197
31. Awareness of the unwritten or unspoken rules of social play	57.1	61.7	.981	67.5	75.0	3.301
32. Facial expression corresponds to environmental events	28.4	26.8	.127	47.4	55.5	3.286
33. Sticking to odd routines or rituals that don't have purpose or make a difference	37.4	40.3	.380	46.3	43.9	.294
34. Abnormal preoccupation with parts of an object or objects	**49.0**	**32.2**	**12.859****	57.0	50.0	2.505
35. Plays appropriately with others	78.1	76.5	.173	83.9	82.3	.244
36. Reads nonverbal cues	43.5	38.9	.955	58.7	59.6	.041
37. Speaks in monotone	5.4	4.7	.119	8.7	13.5	3.277
38. Expects others to know their thoughts, experiences, and opinions without communicating them	44.1	38.3	1.588	49.7	52.4	.371
39. Interest in a highly restricted set of activities	36.0	31.5	.988	52.0	53.4	.090
40. Talking to others in a social context	31.5	36.2	.278	26.4	28.0	.178
41. Use of facial expression	23.8	30.9	2.956	**45.7**	**56.4**	**5.836***
42. Abnormal fascination with the movement of spinning objects	**44.3**	**32.2**	**6.729****	**56.9**	**38.9**	**16.284****
43. Curiosity with surroundings	18.6	15.4	.777	**29.0**	**39.6**	**6.646****
44. Saying words and phrases repetitively	20.2	24.8	1.457	11.6	12.8	.168

Table 5 (continued)

BISCUIT-Part 1 item # and description	Male w/o Cog Delay (%)	Female w/o Cog Delay (%)	χ^2	Male w/ Cog Delay (%)	Female w/ Cog Delay (%)	χ^2
45. Make-believe or pretend play	56.1	52.7	.511	76.4	78.7	.367
46. Understanding age appropriate jokes, figures of speech, or sayings	25.4	31.1	1.819	45.3	48.8	.618
47. Gives subtle cues or gestures when communicating with others	30.0	23.6	2.244	47.3	48.8	.109
48. Becomes upset if there is a chance in routine	57.6	65.8	3.143	51.5	52.4	.047
49. Needs reassurance, especially if events don't go as planned	50.6	49.0	.124	42.5	44.5	.205
50. Language development	95.0	96.6	.678	98.7	98.2	.258
51. Responds to others' distress	52.1	44.6	2.498	71.2	70.7	.011
52. Socializes with other children	82.9	81.9	.082	85.1	87.8	.776
53. Use of nonverbal communication	51.7	51.0	.024	78.7	76.8	.271
54. Clumsiness	**41.8**	**61.1**	**16.868****	44.1	50.0	1.754
55. Limited number of interests	40.0	34.9	1.256	54.8	59.8	1.256
56. Imitation of an adult or child model	**32.8**	**20.8**	**7.681****	62.2	62.2	.000
57. Abnormal, repetitive hand or arm movements	32.3	29.5	.386	51.6	55.2	.667
58. Abnormal, repetitive motor movements involving entire body	37.0	35.6	.101	51.8	50.6	.066
59. Development of social relationships	81.3	81.9	.025	84.1	87.2	.918
60. Respect for others' personal space	**34.6**	**45.0**	**5.153***	37.2	32.3	1.312
61. Needs reassurance, especially if events don't go as planned	58.4	58.4	.000	**66.2**	**50.6**	**13.132****
62. Participation in games or other social activities	46.5	49.0	.273	68.2	68.7	.015

*$p < .05$; **$p < .01$ using Benjamini-Hochberg procedure

to their male counterparts: (23) body posture and/or gesture; (26) displays a range of socially appropriate facial expressions; (41) use of facial expressions; and (43) curiosity with surroundings. Compared to Females w/ Cog Delay, Males with Cog Delay had significantly higher endorsement on: (9) use of language to communicate; (16) use of language in conversation with others; (42) abnormal fascination with the movement of spinning objects; and (61) needs reassurance, especially if events don't go as planned. Only two items were found to have significant gender differences in endorsement frequency between those with and without cognitive delay, with males having higher endorsement: (9) use of language to communicate and (42) abnormal fascination with the movement of spinning objects.

Discussion

This study examined gender differences in ASD symptoms and developmental functioning among young children by replicating and extending a study previously conducted by Sipes et al. (2011). Consistent with the previous study (Sipes et al. 2011), significant differences in overall ASD symptom severity and severity of symptoms in each domain were identified in the current study between groups of young children with ASD with and without cognitive delay. However, when groups were matched on cognitive ability, no gender differences were observed in either total score or subscale scores. These findings together suggest that overall severity of ASD symptoms in young children may be associated with cognitive functioning and not with gender. While results support previous findings demonstrating a lack of gender differences in ASD symptomology during young childhood (Andersson et al. 2013; Carter et al. 2007; Lawson et al. 2018; Reinhardt et al. 2014), this study is among the first to replicate these findings within a population-level sample. As mentioned, previous studies on this topic acknowledged limitations due to sample size and encouraged further research on large, representative samples to improve strength of conclusions and elucidate mixed findings (Andersson et al. 2013; Lawson et al. 2018). Population-level samples, such as the one used in the current study, are advantageous due to the representativeness of the children in the sample and thus the increased generalizability of results (Murdoch and Detsky 2013).

Despite findings suggesting a lack of differentiation in ASD symptom severity, item endorsement patterns were compared to identify subtle differences in how symptoms present across genders and developmental levels in the ASD sample. Distinct relationships were found. First, a greater number of specific symptom presentation differences were

found among males and females without cognitive delay than those with cognitive delay. Consistent with previous research (Kreiser and White 2014; Postorino et al. 2015; Sipes et al. 2011), these results suggest that there are fewer gender differences in ASD presentation within lower functioning ASD populations. Interestingly, one of the two symptoms that demonstrated significant gender differences across cognitive levels was "abnormal fascination with the movement of spinning objects;" male participants were more likely to demonstrate this behavior than females. Of note, this item is one of the most discrete on the RRB domain of the BISCUIT-Part 1. Because researchers have found that females are more likely to display atypical fixations (Hiller et al. 2014), it could be that demonstrating abnormal fascination with spinning objects is a more classic RRB seen in ASD and therefore is more commonly found in males. This possibility highlights the potential male-gender bias in diagnostic instruments; as such, practitioners should be encouraged to examine both quantitative and qualitative data across multiple methods of assessment when evaluating females for ASD.

In regards to children without cognitive delay, caregivers of female participants were more likely to endorse "clumsiness." This finding is consistent with the gender differences identified related to motor functioning on the BDI-2. Caregivers of males were also significantly more likely to endorse certain RRBs, which again aligns with previous findings (Frazier et al. 2014; Frazier and Hardan 2017; Lai et al. 2014). Notably, females received significantly higher endorsements for items related to impairments in verbal language (5 and 9) when compared to males, indicating greater deficits than male participants according to parent report.

To evaluate gender differences in developmental functioning within children with ASD, a non-ASD sample was used to determine uniqueness of the results to the ASD population. Though effect sizes were small, notable gender differences were observed in both non-ASD and ASD samples, and distinctions were found between the results of gender comparisons in the samples. First, whereas Total DQ did not differ between males and females with ASD, males had significantly lower scores than females in the non-ASD sample. Specifically, females in the non-ASD sample outperformed males in Adaptive, Personal-Social, Communication, and Cognitive domains, and no gender differences were found in the Motor domain.

Some similar results were found in the ASD sample when examining gender differences within developmental domains; females were observed to have significantly less impairment in the area of communication. Subsequent comparisons suggest that this differentiation is most true of females with ASD without cognitive delay compared to males with ASD without cognitive delay. This is surprising given that these females received higher endorsements of language impairment than their male counterparts on the BISCUIT-Part 1. Because females were found to have greater communication skills on the BDI-2 (an observational measure) and provided that BISCUIT items are endorsed in comparison to same-aged peers, these results may suggest that caregivers have higher expectations for females when evaluating verbal communication. This supports the theory that social and cultural influences may affect ASD presentation and identification in females (Kreiser and White 2014). Continued research is required to evaluate the magnitude of this influence. As gender expectations evolve with age and differ by culture, cross-cultural comparisons and examination of longitudinal trends would be informative.

In line with previous findings (Carter et al. 2007), females with ASD were also found to have more impairment than males in the area of motor skills. Following analyses controlling for cognitive level, these differences in motor skills appear only for those children with cognitive delay. Further, females with ASD and cognitive delay were found to have greater impairments in adaptive skills compared to males with ASD and cognitive delay; this suggests females with ASD and cognitive impairment may be more affected in the defining areas of intellectual disability than their male peers. Notably, given the small effect sizes, it is difficult to discern the clinical relevance of these developmental differences; continued research in this area should occur to clarify these results. Further, longitudinal studies would be helpful in demonstrating the persistence and course of observed differences in developmental functioning.

This study is not without limitations. As the current sample was pulled from the state's early intervention system, the children studied were quite young. Researchers have consistently found that females with ASD receive their diagnosis significantly later than males, even when symptom severity is held constant (Postorino et al. 2015). Further, children without cognitive delays, and females without cognitive delays specifically, are more likely to receive a delayed diagnosis or no diagnosis at all (Dworzynski et al. 2012; Mandell et al. 2005). Therefore, females with ASD, especially those without cognitive delay, are likely underrepresented; it may be the case that a portion of unscreened females or females in the non-ASD sample are found to meet criteria later in life. This hypothesis is supported by the finding that significantly more parents of female compared to male children in the current study declined the ASD screen. Future research should aim to include more female representative populations to address inadequate and potentially biased diagnostic procedures. Researchers may also consider including young females with subthreshold ASD symptomology to capture those at risk for a later diagnosis.

Further, while clinically interesting, the ASD and non-ASD groups significantly differed from one another on a number of demographic factors. The non-ASD subsample

had a lower M:F ratio, were younger, and a greater proportion identified as white; therefore, comparisons between results of the ASD versus non-ASD groups were not made on matched samples. Also, given the nature of this study and this sample, ASD diagnoses were assigned for purposes of these comparisons through review of clinical records using *DSM-5* criteria rather than individualized clinical evaluations. While this procedure is common in evaluating population-level data, such as in the CDC's Autism and Developmental Disabilities Monitoring (ADDM) Network and other studies (Fombonne et al. 2004; Huerta et al. 2012; Kim et al. 2014), there are drawbacks to this method. For instance, there is a recognizable tradeoff between precision and power when examining population-level data; with increased sample size and generalizability comes reduced detail and control in evaluation. Lastly, ASD symptom presentation was measured based upon parent report in the current study. Given the discrepancies found between the BISCUIT-Part 1 and the BDI-2 noted above, parental biases may be at play. Future research should look to find a balance between informant report and observational measures.

Despite these limitations, the current study adds to the research base and contributes to our understanding of the early autism phenotype in males and females. The current comparison of males and females with ASD who demonstrate identifiable autism symptomology early in development is valuable and should be utilized to track how discrepancies between genders change across time and between diverse age samples. While severity discrepancies were more attributable to cognitive level than to gender, subtle presentation differences were observed between male and female participants. This was particularly true for those without cognitive delay. While research in this area is ongoing, the information necessary to improve our understanding of this gender discrepancy in diagnosis will not be available if research samples continue to reflect low gender ratios. Therefore, continued efforts are required to discern true gender differences and improve clinical care for a variety of phenotypes.

Author Contributions MM conceived of the study, participated in its design, performed the statistical analysis, participated in the interpretation of the data, and helped to draft the manuscript; JM participated in the design and coordination of the study; EH helped to draft the manuscript; PC participated in the interpretation of the data and helped to draft the manuscript. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of interest Mrs. Deann Matson, Dr. Johnny L. Matson's wife, is the sole owner of the *Baby and Infant Screen for Children with aUtism Traits (BISCUIT)* and sells the scale. Ms. Maya Matheis, Ms. Esther Hong, and Dr. Paige E. Cervantes declares that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent For this type of study, formal consent is not required.

References

Adamou, M., Johnson, M., & Alty, B. (2018). Autism Diagnostic Observation Schedule (ADOS) scores in males and females diagnosed with Autism: A naturalistic study. *Advances in Autism*. https://doi.org/10.1108/AIA-01-2018-0003.

Andersson, G. W., Gillberg, C., & Miniscalco, C. (2013). Pre-school children with suspected autism spectrum disorders: Do girls and boys have the same profiles? *Research in Developmental Disabilities, 34*(1), 413–422. https://doi.org/10.1016/j.ridd.2012.08.025.

Attwood, T. (2007). *The complete guide to Asperger's syndrome*. London: Jessica Kingsley Publishers.

Bacon, A. L., Fein, D., Morris, R., Waterhouse, L., & Allen, D. (1998). The responses of autistic children to the distress of others. *Journal Of Autism And Developmental Disorders, 28*(2), 129–142.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., et al. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2014. *MMWR. Surveillance Summaries, 67*(SS-6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1.

Banach, R., Thompson, A., Szatmari, P., Goldberg, J., Tuff, L., Zwaigenbaum, L., & Mahoney, W. (2009). Brief report: Relationship between non-verbal IQ and gender in autism. *Journal of Autism and Developmental Disorders, 39*(1), 188–193. https://doi.org/10.1007/s10803-008-0612-4.

Baron-Cohen, S., Lombardo, M. V., Auyeung, B., Ashwin, E., Chakrabarti, B., & Knickmeyer, R. (2011). Why are autism spectrum conditions more prevalent in males? *PLOS Biology, 9*(6), e1001081. https://doi.org/10.1371/journal.pbio.1001081.

Bauminger, N., Solomon, M., Aviezer, A., Heung, K., Brown, J., & Rogers, S. J. (2008). Friendship in high-functioning children with autism spectrum disorder: Mixed and non-mixed dyads. *Journal Of Autism And Developmental Disorders, 38*(7), 1211–1229.

Begeer, S., Mandell, D., Wijnker-Holmes, B., Venderbosch, S., Rem, D., Stekelenburg, F., & Koot, H. (2012). Sex differences in the timing of identification among children and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-012-1656-z.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289–300.

Brown, J. R., & Dunn, J. (1996). Continuities in emotion understanding from 3 to 6 yrs. *Child Development, 67*(3), 789–802. https://doi.org/10.2307/1131861.

Carter, A. S., Black, D. O., Tewani, S., Connolly, C. E., Kadlec, M. B., & Tager-Flusberg, H. (2007). Sex differences in toddlers with autism spectrum disorders. *Journal Of Autism And Developmental Disorders, 37*(1), 86–97.

Dean, M., Harwood, R., & Kasari, C. (2017). The art of camouflage: Gender differences in the social behaviors of girls and boys with autism spectrum disorder. *Autism, 21*(6), 678–689. https://doi.org/10.1177/1362361316671845.

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 & Adolescent Psychiatry, 51*(8), 788–797. https://doi.org/10.1016/j.jaac.2012.05.018.

Fombonne, E., Heavey, L., Smeeth, L., Rodrigues, L. C., Cook, C., Smith, P. G., et al. (2004). Validation of the diagnosis of autism in general practitioner records. *BMC Public Health, 4*(1), 1.

Fombonne, E., Quirke, S., & Hagen, A. (2011). Epidemiology of pervasive developmental disorders. In D. Amaral, D. Geschwind & G. Dawson (Eds.), *Autism spectrum disorders* (pp. 90–111). New York: Oxford University Press.

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 & Adolescent Psychiatry, 53*(3), 329–340.e3. 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.

Gould, J., & Ashton-Smith, J. (2011). Missed diagnosis or misdiagnosis? Girls and women on the autism spectrum. *Good Autism Practice (GAP), 12*(1), 34–41.

Haney, J. L. (2015). Autism, females, and the DSM-5: Gender bias in autism diagnosis. *Social Work in Mental Health*. https://doi.org/10.1080/15332985.2015.1031858.

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.

Hill, A. P., Zuckerman, K., & Fombonne, E. (2016). Epidemiology of autism spectrum disorder. In C. McDougle (Ed.), *Autism Spectrum Disorder* (pp. 181–204). New York: Oxford University Press.

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.

Holtmann, M., Bölte, S., & Poustka, F. (2007). Autism spectrum disorders: Sex differences in autistic behaviour domains and coexisting psychopathology. *Developmental Medicine & Child Neurology, 49*(5), 361–366. https://doi.org/10.1111/j.1469-8749.2007.00361.x.

Huerta, M., Bishop, S. L., Duncan, A., Hus, V., & Lord, C. (2012). Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *American Journal of Psychiatry, 169*(10), 1056–1064. https://doi.org/10.1176/appi.ajp.2012.12020276.

Kim, Y. S., Fombonne, E., Koh, Y.-J., Kim, S.-J., Cheon, K.-A., & Leventhal, B. L. (2014). A comparison of DSM-IV pervasive developmental disorder and DSM-5 autism spectrum disorder prevalence in an epidemiologic sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(5), 500–508. https://doi.org/10.1016/j.jaac.2013.12.021.

Kirkovski, M., Enticott, P. G., & Fitzgerald, P. B. (2013). A review of the role of female gender in autism spectrum disorders. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-013-1811-1.

Koenig, K., & Tsatsanis, K. D. (2005). Pervasive developmental disorders in girls. In D. J. Bell, S. L. Foster, & E. J. Mash (Eds.), *Handbook of Behavioral and Emotional Problems in Girls* (pp. 211–237). New York: Springer.

Kreiser, N. L., & White, S. W. (2014). ASD in females: Are we overstating the gender difference in diagnosis? *Clinical Child and Family Psychology Review, 17*(1), 67–84. https://doi.org/10.1007/s10567-013-0148-9.

Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2014). Sex/gender differences and autism: Setting the scene for future research. *Journal of the American Academy of Child and Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2014.10.003.

Lawson, L. P., Joshi, R., Barbaro, J., & Dissanayake, C. (2018). Gender differences during toddlerhood in autism spectrum disorder: A prospective community-based longitudinal follow-up study. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-018-3516-y.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation schedule: Manual*. Los Angeles: Western Psychological Services.

Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics, 116*(6), 1480–1486. https://doi.org/10.1542/peds.2005-0185.

Manning, S. E., Davin, C. A., Barfield, W. D., Kotelchuck, M., Clements, K., Diop, H., et al. (2011). Early diagnoses of autism spectrum disorders in Massachusetts birth cohorts, 2001–2005. *Pediatrics, 127*(6), 1043–1051. https://doi.org/10.1542/peds.2010-2943.

Matson, J. L., Boisjoli, J. A., Hess, J. A., & Wilkins, J. (2010). Factor structure and diagnostic fidelity of the Baby and Infant Screen for Children with aUtIsm Traits-Part 1 (BISCUIT–part 1). *Developmental Neurorehabilitation, 13*(2), 72–79. https://doi.org/10.3109/17518420903213576.

Matson, J. L., Boisjoli, J. A., & Wilkins, J. (2007). *The baby and infant screen for children with autism traits (BISCUIT)*. Baton Rouge: Disability Consultants, LLC.

Matson, J. L., Wilkins, J., & Fodstad, J. C. (2011). The validity of the Baby and Infant Screen for Children with aUtIsm Traits: Part 1 (BISCUIT: Part 1). *Journal of Autism and Developmental Disorders, 41*(9), 1139–1146. https://doi.org/10.1007/s10803-010-0973-3.

Matson, J. L., Wilkins, J., Sharp, B., Knight, C., Sevin, J. A., & Boisjoli, J. A. (2009). Sensitivity and specificity of the Baby and Infant Screen for Children with aUtIsm Traits (BISCUIT): Validity and cutoff scores for autism and PDD-NOS in toddlers. *Research in Autism Spectrum Disorders, 3*(4), 924–930. https://doi.org/10.1016/j.rasd.2009.04.001.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin, 126*(3), 424–453. https://doi.org/10.1037/0033-2909.126.3.424.

McLennan, J. D., Lord, C., & Schopler, E. (1993). Sex differences in higher functioning people with autism. *Journal of Autism and Developmental Disorders, 23*(2), 217–227. https://doi.org/10.1007/BF01046216.

Moyal, N. (2010). *The Battelle Developmental Inventory, 2nd Edition: A study of concurrent validity and stability in young children with known disabilities*. New Jersey: Fairleigh Dickinson University.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *JAMA, 309*(13), 1351–1352. https://doi.org/10.1001/jama.2013.393.

Mussey, J. L., Ginn, N. C., & Klinger, L. G. (2017). Are males and females with autism spectrum disorder more similar than we thought? *Autism, 21*(6), 733–737. https://doi.org/10.1177/1362361316682621.

Newborg, J. (2005). *Battelle Developmental Inventory* (*2nd edn*). Itasca: Riverside.

Nicholas, J. S., Charles, J. M., Carpenter, L. A., King, L. B., Jenner, W., & Spratt, E. G. (2008). Prevalence and characteristics of children with autism-spectrum disorders. *Annals of Epidemiology, 18*(2), 130–136.

Nydén, A., Hjelmquist, E., & Gillberg, C. (2000). Autism spectrum and attention-deficit disorders in girls. Some neuropsychological aspects. *European Child & Adolescent Psychiatry, 9*(3), 180–185.

Postorino, V., Fatta, L. M., De Peppo, L., Giovagnoli, G., Armando, M., Vicari, S., & Mazzone, L. (2015). Longitudinal comparison between male and female preschool children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(7), 2046–2055. https://doi.org/10.1007/s10803-015-2366-0.

Reinhardt, V. P., Wetherby, A. M., Schatschneider, C., & Lord, C. (2014). Examination of sex differences in a large sample of young children with autism spectrum disorder and typical development. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-014-2223-6.

Rivet, T. T., & Matson, J. L. (2011). Review of gender differences in core symptomatology in autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*(3), 957–976. https://doi.org/10.1016/j.rasd.2010.12.003.

Rosenberg, R. E., Landa, R., Law, J. K., Stuart, E. A., & Law, P. A. (2011). Factors affecting age at initial autism spectrum disorder diagnosis in a national survey. *Autism Research & Treatment*. https://doi.org/10.1155/2011/874619.

Rutter, M., Caspi, A., & Moffitt, T. E. (2003). Using sex differences in psychopathology to study causal mechanisms: Unifying issues and research strategies. *Journal of Child Psychology and Psychiatry, 44*(8), 1092–1115. https://doi.org/10.1111/1469-7610.00194.

Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised (ADI-R)*. Los Angeles: Western Psychological Services.

Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., & Baron-Cohen, S. (2016). An investigation of the 'female camouflage effect' in autism using a computerized ADOS-2 and a test of sex/gender differences. *Molecular Autism*. https://doi.org/10.1186/s13229-016-0073-0.

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., et al. (2009). Timing of identification among children with an autism spectrum disorder: Findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*(5), 474–483. https://doi.org/10.1097/CHI.0b013e31819b3848.

Siklos, S., & Kerns, K. A. (2007). Assessing the diagnostic experiences of a small sample of parents of children with autism spectrum disorders. *Research in Developmental Disabilities, 28*(1), 9–22. https://doi.org/10.1016/j.ridd.2005.09.003.

Sipes, M., Matson, J. L., Worley, J. A., & Kozlowski, A. M. (2011). Gender differences in symptoms of autism spectrum disorders in toddlers. *Research in Autism Spectrum Disorders, 5*(4), 1465–1470. https://doi.org/10.1016/j.rasd.2011.02.007.

Sonnander, K. (2007). Early identification of children with developmental disabilities. *Acta Paediatrica, 89*(s434), 17–23. https://doi.org/10.1111/j.1651-2227.2000.tb03091.x.

Stone-MacDonald, A., Pizzo, L., & Feldman, N. (2018). Assessment in early intervention: Using the Battelle developmental inventory. In A. Stone-MacDonald, L. Pizzo & N. Feldman (Eds.), *Fidelity of implementation in assessment of infants and toddlers: evaluating developmental milestones and outcomes*, 2nd edn, (pp. 29–45). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-74618-0_3.

Szatmari, P., Georgiades, S., Bryson, S., Zwaigenbaum, L., Roberts, W., Mahoney, W., et al. (2006). Investigating the structure of the restricted, repetitive behaviours and interests domain of autism. *Journal Of Child Psychology And Psychiatry, And Allied Disciplines, 47*(6), 582–590.

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.

Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., et al. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. *Pediatrics, 136*(Supplement), S60–S81. https://doi.org/10.1542/peds.2014-3667E.