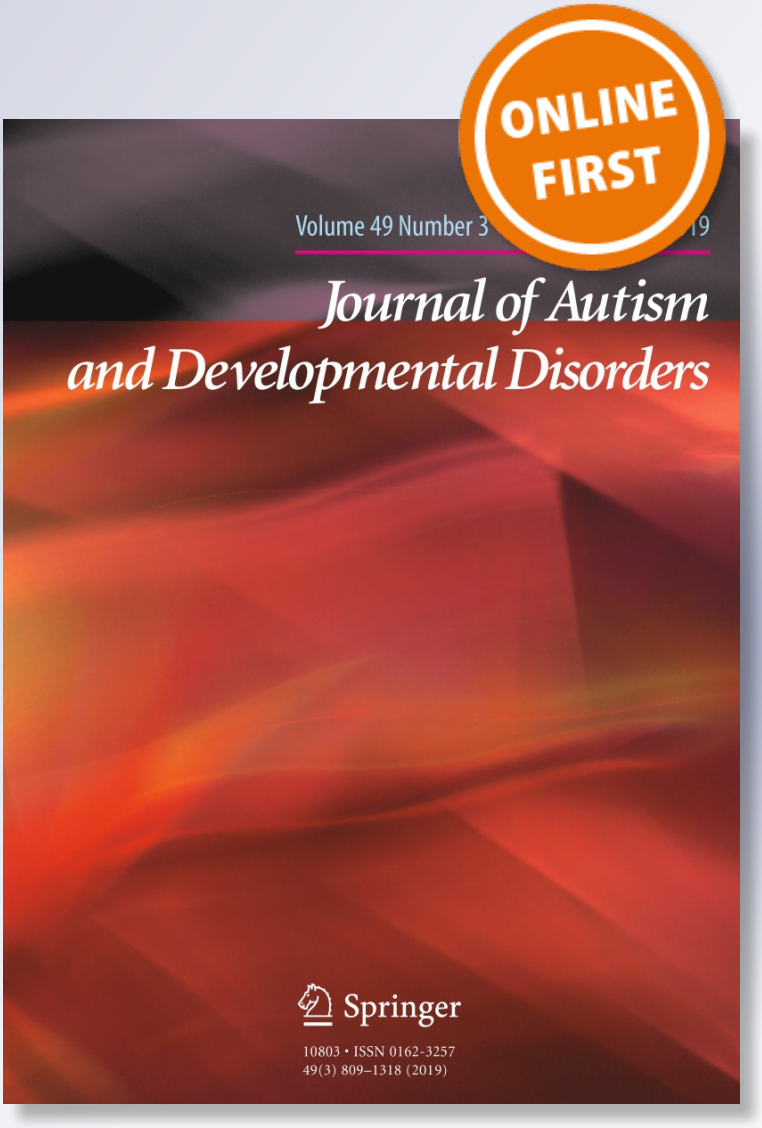
No Sex Differences in Cognitive Ability in Young Children with Autism Spectrum Disorder

Susanne W. Duvall, Lark Huang-Storms, Alison Presmanes Hill, Julianne Myers & Eric Fombonne

Journal of Autism and Developmental Disorders

ISSN 0162-3257

J Autism Dev Disord
DOI 10.1007/s10803-019-03933-1



Springer

Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-019-03933-1

ORIGINALPAPER

No Sex Differences in Cognitive Ability in Young Children with Autism Spectrum Disorder

Susanne W. Duvall[1,2] · Lark Huang-Storms[1] · Alison Presmanes Hill[1,3] · Julianne Myers[3,4] · Eric Fombonne[1,2]

Abstract
Inconsistent findings regarding sex differences in cognition have been found in people with autism spectrum disorder (ASD). This study evaluated sex differences in cognitive-developmental functioning in a large clinical sample of young children diagnosed with ASD. The sample included children 18–68 months of age who received the Mullen Scales of Early Learning (MSEL) through Autism Treatment Network (ATN) sites from 2007 to 2013 (N = 1587, 16.7% female). In this large clinically referred sample of young children with ASD in the United States, no significant differences were found between the sexes for the MSEL Early Learning Composite (ELC) standard score, domain T Scores or age equivalents. These findings persisted when examining different age ranges, cognitive levels and domain profiles.

Keywords Autism spectrum disorder · Cognition · Intellectual function · Sex differences · Female

Introduction

Autism spectrum disorder (ASD) is a multifaceted condition characterized by verbal and nonverbal social interaction and developmental communication deficits, as well as repetitive/restricted interests, behaviors and speech (DSM-5, American Psychiatric Association 2013). According to current estimates from the Centers for Disease Control and Prevention (CDC), approximately one in 59 children in the United States experiences ASD (Baio et al. 2018), males are approximately 4 times more commonly affected by ASD than females and even more frequently represented in higher functioning samples. In general, increased autism symptom severity has been linked with lower cognitive function, with systematic reviews showing approximately 50% of individuals with ASD scoring in the intellectual disability range (Hill et al. 2015; Myers et al. 2018). Moreover, it has been observed that females with ASD are more likely to be diagnosed with intellectual disability (ID) and global developmental delay than males, with relatively higher representation in lower functioning ASD samples (Baio 2012; Blumberg et al. 2013; Fombonne 2003; Giarelli et al. 2010; Loomes et al. 2017; Reinhardt et al. 2015). Thus, it is important to understand whether cognitive function differentially impacts the rate of diagnosis and/or the phenotypic presentation of ASD across the sexes. Examining sex differences could be important for genetic modeling studies if females are less commonly impacted, but when ASD is present they are more functionally impaired.

Despite documented sex differences in prevalence, research examining cognitive abilities across the sexes has been relatively limited and has resulted in inconsistent findings (Lai et al. 2015). In some more recent studies, males obtained higher scores on intelligence quotient (IQ) and cognitive tests (Ankenman et al. 2014; Banach et al. 2009; Giarelli et al. 2010; Postorino et al. 2015) whereas other contemporary studies showed few sex related cognitive differences (Howe et al. 2015; Koyama et al. 2009; Lai et al. 2012; Mussey et al. 2017). Discrepant findings across studies may

✉ Susanne W. Duvall
duvall@ohsu.edu

[1] Department of Pediatrics, Institute on Development and Disability, Oregon Health and Science University, CDRC, 707 SW Gaines Street, Portland, OR 97239-3098, USA

[2] Department of Psychiatry, Oregon Health and Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239-3098, USA

[3] Center for Spoken Language Understanding, Oregon Health amd Science University, CSLU GH 40, 3181 SW Sam Jackson Park Road, Portland, OR 97239-3098, USA

[4] Present Address: Department of Clinical Psychology, Seattle Pacific University, Seattle, WA, USA

Published online: 27 February 2019

reflect variability in sample characteristics such as clinically referred versus non-referred samples, the age of participants and the type of cognitive measure used. In particular, few studies focus solely on early childhood samples, and the inclusion of different age ranges across and within studies makes interpretation of these data challenging (Please refer to Table 1 for a summary of the reviewed literature.)

Community samples of individuals with ASD may differ from research-recruited samples in terms of representativeness, diversity and severity of presentation. Some research studies and meta-analyses suggest that females with ASD have lower mean IQ scores compared to males (Halladay et al. 2015; Frazier et al. 2014) whereas other studies of community-based clinics found similar rates of ID in males and females, with no differences observed in Full Scale (FSIQ), Verbal (VIQ) or Nonverbal IQ (NVIQ) standard scores (Mussey et al. 2017).

The use of various cognitive and developmental measures across studies may also explain inconsistencies in findings. For example, when examining NVIQ, males 4–17 years of age with ASD received higher scores on the Differential Ability Scales-Second Edition (DAS-II) than females (Ankenman et al. 2014). Similarly, young females 2–5 years of age obtained significantly lower scores on the Griffiths Mental Development Scale-Extended Revised (GMDS-ER; Postorino et al. 2015). In contrast, Koyama et al. (2009) found no sex differences on the Performance IQ of the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) in elementary school aged children with ASD.

The inclusion of large age ranges often results in different cognitive measures being utilized within the same study, precluding or clouding direct comparison of results across and within studies (Mandy et al. 2012). It has also been argued that studies examining very large age ranges may obscure or confound developmental effects (Lai et al. 2012) and, thus, use of a more restricted age range may be more informative. However, even among samples restricted to young children, inconsistencies persist as some studies have demonstrated sex differences in cognition (Carter et al. 2007; Lord et al. 1982) and others have found no sex differences in cognitive function (Hartley and Sikora 2009; Tsai and Beisler 1983).

The most common measure used in studies of young children with ASD is the Mullen Scales of Early Learning (MSEL; Mullen 1995). In several samples of toddlers, no sex differences were found for domain scores on the MSEL (Hartley and Sikora 2009; Reinhardt et al. 2015). In some prospective studies examining high-risk baby siblings (which may include young females with more subtle symptoms of ASD), females show a trend for stronger cognitive abilities with higher mean age equivalents on the MSEL fine motor domain (Zwaigenbaum et al. 2012), as well as higher, albeit non-significantly different, scores across other domain scales of the MSEL (Messinger et al. 2015; Reinhardt et al. 2015). In another sample of toddlers, girls and boys with ASD demonstrated different developmental profiles on the MSEL (Carter et al. 2007). Specifically, boys showed higher performance on expressive language and motor skills, especially after covarying for visual skills, while toddler girls scored higher on the visual reception domain after controlling for language ability.

In light of inconsistency in findings and the relatively small sample sizes utilized in previous studies of preschool children with ASD, we set to reexamine sex differences in cognitive functioning in a very large sample of young children with ASD ranging in age from 18 to 68 months using the MSEL. The specific questions addressed by this study include:

1. Are there sex differences in overall cognitive function in young males and females with ASD? 2. Are there sex differences in young males and females with ASD in specific cognitive domains?
2. Are there differences in male and females scores that are influenced by overall level of ability or chronological age?
3. Does the pattern of discrepancy between cognitive domains vary by sex?

Methods

Participants

The current study included children aged 18–68 months enrolled in the Autism Speaks Autism Treatment Network (ATN). The ATN is a collaboration among 18 academic health centers in North America, which was established to develop a model of comprehensive medical care for children and adolescents with ASD (Lajonchere et al. 2012). All participants in the ATN registry have a confirmed ASD diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev.; DSM-IV-TR; American Psychiatric Association 2000) criteria and supported by administration of the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 1999). Eligible families were identified through clinical visits and invited to participate in the registry, which included obtaining written consent and the collection of routine clinical data and behavior assessments. Data was entered by trained study coordinators and the Institutional Review Boards at each site approved the registry protocols.

All children enrolled in the ATN Registry (based on the date of consent) from 2007 to 2013 were included, and this current sample was obtained at the time of study conception and analysis. We included individuals who completed the MSEL within the standardized age range (0–68 months) as

Table 1 Summary of literature reviewed on possible sex differences in cognition in ASD

References	Age range in years, Mean (SD)†	Sex, n	Setting	ASD diagnostic criteria	Cognitive/IQ measure	Mean global IQ scores (SD)‡	Subdomains, additional IQ findings‡
Ankenman et al. (2014)	4–17:11 M + F = 8.8 (3.5) M: UNK F:UNK	M:1710 F: 244	Research recruited SSC	ADOS CPEA criteria ADI-R	DAS-II	M + F: 90.9 (19.9) M: UNK F:UNK	VIQ: M 89.6 (21.5), F 88.2 (23.3) NVIQ: M 93.8 (18.7), F 88.1 (20.4)* Discrepancies NVIQ > VIQ (M 28.7%, F 18.4%)* NVIQ = VIQ (M 58%, F 65.2%)* VIQ > NVIQ (M 13.4%, F 16.4%)
Banach et al. 2009)	Range UNK M: 9.4 (5.8) F: 9.1 (5.6)	Simplex M: 144 F: 50 Multiplex M: 152 F: 50	Research recruited	DSM-IV ADOS ADI-R	Leiter (only NVIQ)	Simplex*** M: 76.2 (29.5) F: 50.2 (22.8) Multiplex M: 68.9 (31.8) F: 74.8 (30.4)	Simplex percent of sample IQ 0–50, M 20.3%, F 54.8%* IQ 51–70, M 23.3%, F 23.8% IQ > 70, M 56.4%, F 21.4% Multiplex percent of sample: UNK
Carter et al. (2007)	18–33 mo M: 28.4 mo (3.5) F: 27.1 mo (4.8)	M: 68 F: 22	Research recruited	ADOS ADI-R	MSEL	ELC not reported	Similar MSEL AE scores across domains RL: M 14.3 (7.2), F 12.2 (6.4) EL: M 16.4 (8.1), F 13.3 (6.9) VR: M 20.3 (6.1), F 21.4 (5.8) GM: M 19.0 (3.3), F 17.1 (2.9)* FM: M 21.4 (3.7), F 19.6 (4.6) F > M on adjusted VR*[a] M > F on language composite* and motor skills composite**[b]
Frazier et al. (2014)	4–18 M: 9.01 (3.56) F: 9.32 (3.67)	M: 2114 F: 304	Research recruited SSC	ADOS ADI-R SRS	DAS-II WISC-IV	M: 82.56 (27.59) F:74.70 (27.59)	VIQ: M 79.2 (30.7), F 73.4 (32)** NVIQ: M 86 (25.8), F 77.4 (26.2)*** Discrepancies [VIQ–NVIQ] M 3.97 (16.3), F 6.79 (17.0)**
Giarelli et al. (2010)	8y index age SD UNK [Children 8 years old in 2002]	M: 2080 F: 487	Secondary analysis of population Surveillance data ADDM	Health and education data using DSM-IV-TR	UNK 75% of sample various IQ data	UNK	M + F: 42% IQ > 70, 33% IQ ≤ 70 More girls than boys had IQ scores ≤ 70 [X^2(df = 1, n = 1,919) = 12.4]**
Hartley and Sikora (2009)	18–47 mo M: 35.5 mo (7.1) F: 36 mo (7.3)	M: 157 F: 42	Clinically referred	DSM-IV Interview ADOS	MSEL	ELC not reported, Visual Reception (VR) used as cognition	Similar MSEL AE scores across domains VR: M 21.1 (0.6), F 20.6 (1.1) FM: M 22.5 (0.5), F 22.6 (0.9) RL: M 14.5 (0.8), F 12.9 (1.4) EL: M 15.3 (0.7), F 15.7 (1.3)

Table 1 (continued)

References	Age range in years, Mean (SD)†	Sex, n	Setting	ASD diagnostic criteria	Cognitive/IQ measure	Mean global IQ scores (SD)‡	Subdomains, additional IQ findings‡
Howe et al. (2015)	Over 5y, mean 8.6–9.7 [c] AGRE M: 9.4 (3.5) F: 9.1 (3.3) AC M: 10.1 (4.2) F: 9.7 (3.6) ATN M: 8.6 (3.0) F: 8.8 (2.9) SSC M: 9.6 (3.3) F: 9.9 (3.5)	M: 4851 F: 872	Research recruited SSC AGRE AC ATN	ADOS	SB-5 WPPSI DAS WASI WISC-IV	ADOS Mod 1 Mean M + F IQ ≤ 70 across all datasets	ADOS Mod 2 and 3 included below: ADOS Mod 2 ATN, M: 74 (23) F: 61 (15)** SSC, M: 69 (19) F: 62 (19)** AGRE, M: 68 (18) F: 66 (16) ADOS Mod 3 ATN, M: 92 (20) F: 92 (20) SSC, M: 95 (19) F: 91 (21)* AGRE, M: 93 (20) F: 9 (20)
Koyama et al. (2009)	Range UNK M: 9 (3.0) F: 8.2 (2.1)	M: 120 F: 22	Clinically referred	DSM-IV ICD-10	Japanese version WISC-III	HFA all FSIQ > 70 M: 96.0 (14.2) F: 97.9 (13.6)	VIQ: M 94.8 (18.0), F 99.4 (18.1) PIQ: M 98.2 (14.7), F 96.6 (12.4) PSI: M 92.5 (16.8), F 101.1 (13.1)*
Lai et al. (2012)	18–45 M: 27.0 (7.1) F: 26.9 (6.7)	M: 33 F: 29	Research recruited MRC AIMS	ADI-R ADOS, Mod 4	WASI	M:112.6 (16.3) F: 112.8 (15.7)	VIQ: M 111.5 (15.3), F 113.1 (15.4) PIQ: M 111.1 (16.4), F 109.5 (17.5)
Lord et al., (1982)	3–8 M: 5.29 (0.55) F: 5.31 (0.53)	M: 384 F: 91	Clinically referred	CARS	Leiter Bayley WISC-R MPSMT	UNK	NVIQ: M 43.62 (20.24), F 37.23 (16.88)*
Mandy et al. (2012)	3–18 M: 9.7 (3.1) F: 10.2 (3.5)	M: 273 F: 52	Clinically referred	3Di [d] ADOS	BPVS WASI WISC-III WISC-IV	UNK [“high functioning”]	VIQ: M 92.7 (19.5), F 92.5 (18.5) PIQ: M 94.8 (19.7), F 91.4 (19.6) No sex differences in proportion scoring < 70 on VIQ or PIQ
Messinger, et al. (2015)	Initial visit M + F 7.62 mo (3.58) Final visit M + F: 37.13 mo (2.06)	M: 193 F: 59	Research recruited, baby siblings BSRC	ADOS DSM-IV	MSEL	ELC not reported	Estimated MSEL AE[e] VR: M 33.0 (0.5), F 34.5 (0.9) FM: M 29.4 (0.5), F 30.66 (0.9) RL: M 28.5 (0.5), F 29.8 (0.9) EL: M 29.5 (0.5), F 30.4 (0.9)
Mussey et al. (2017)	M + F 1.75 − 5.63	M: 679 F: 113	Clinically referred	CARS ADOS	UNK, e.g., Wecshler scales	M: 85.6 (22.1) F: 86 (21.8)	VIQ: M 92.3 (20.8), F 90.7 (21.4) NVIQ: M 94.7 (19.8), F 89.7 (20.8) Equal rates of intellectual disability in males and females
Postorino et al., 2015)	2–5.4 M + F: 3.55 (0.9) [f]	M: 30 F: 30	Clinically referred	DSM-IV-TR ADOS	GMDS-ER	GDQ M: 70.0 ± 24.5 F: 58.2 ± 13.8*	Performance DQ:[g] M 82.2 (30.4) F 62.1 (20.3)** Eye–hand coordination DQ M 67.8 (27.5) F 52.6 (13.1)*

Table 1 (continued)

References	Age range in years, Mean (SD)†	Sex, n	Setting	ASD diagnostic criteria	Cognitive/IQ measure	Mean global IQ scores (SD)‡	Subdomains, additional IQ findings‡
Reinhardt et al. (2015)	28.09 mo (12.30) M: 19.99 (2.25) F: 20.61 (2.32)	M: 234 F: 54	Research recruited	ADOS	MSEL CSBS	MSEL NVDQ [h] M: 26.07(9.09) F: 25.64 (8.85) MSEL VDQ [h] M:70.05(28.78) F: 22.67 (11.66) [CSBS UNK]	Similar MSEL AE scores across domains [i] VR: M 38.2(16.6), F 38.7(16.1) FM: M 36.8 (15.0), F 37.7(15.2) RL: M 34.8 (15.3), F 35.2 (15.4) EL: M 33.8 (14.1), F 36.6(15.2) Females scored higher* on CSBS Words subdomain; no sex differences across other domains [j]
Tsai and Beisler (1983)	M: 72.44 mo (29.7) F: 76.5 mo (35.6)	M: 52 F: 23	Research recruited	Kanner (1943) and Rutter (1978) criteria	Various measures [k]	M: 57.3 (26.01) F: 42.1 (26.78)*	Overall, 28% IQ > 70, 23% IQ 70 < 50 & 49% IQ < 50 More females than males with IQ < 50
Volkmar et al. (1993)	AUT M + F = 13.3 (7.6) PDD M + F = 7.5 (6.1) M: UNK F: UNK	M: 214 (156 AUT) F: 59 (43 AUT)	Clinically referred	ABC DSM-III DSM-III-R	Varied FSIQ & IQ estimate[z]	M + F = 55.7 (25.2) M: UNK F: UNK	PDD-NOS sample No differences on FSIQ (values UNK) IQ < 35, M 5.2%, F 6.3% IQ 35–69, M 39.7%, F 50% IQ > 70, M 55.2%, F 43.8% AUT Sample M 43.4 (24.3), F 36.5 (18.9)* on IQ [l] IQ < 35, M 33.2%, F 42.4%* IQ 35–69, M 27.1%, F 28.8% IQ > 70, M 12.6%, F 1.7%*
Zwaigenbaum et al. (2012)	36–42 mo M + F = 38.1 mo (2.5) M: UNK F: UNK	M: 57 F: 29	Research recruited, high risk baby siblings	DSM-IV-TR ADOS ADI-R	MSEL	M + F ELC = 83.4 (24.1) M: UNK F: UNK	MSEL T scores VR: M 46.0 (20.0), F 47.2 (16.1) FM: M 37.2 (16.2), F 42.8 (13.7) *** RL: M 41.9 (13.4), F 40.2 (13.0) EL: M 42.8 (12.9), F 42.1 (12.9)

Table 1 (continued)

ABC Autism Behavior Checklist, *AC* Autism Consortium, *ADDM* Autism and Developmental Disabilities Monitoring Network for ASD Surveillance, *AE* age equivalent in months, *AGRE* autism genetics resource exchange, *ASG* Asperger syndrome, AUT autistic disorder, *BPVS* British Picture Vocabulary Scale, *BRSC* Baby Siblings Research Consortium, *CARS* Childhood Autism Rating Scale, *CPEA* Collaborative Programs for Excellence in Autism, *CSBS-DP* Communication and Symbolic Behavior Scales Developmental Profile, *DAS-II* Differential ability scales—second edition, *DAWBA* development and well-being assessment, *DP* developmental profile, *DQ* developmental quotient, *EL* expressive language, *ELC* early learning composite (on MSEL), *F* Female, *FM* fine motor, *GDQ* general developmental quotient, *GM* gross motor, *GMDS-ER* Griffiths Mental Development Scale-Extend Revised, *HFA* high functioning autism, *ID* intellectual disability, *IQ* intelligence quotient, *K-ABC* Kaufman assessment battery for children, *Leiter* Leiter Performance Scales, *M* males, *mo* months, *MPSMT* Merrill-Palmer Scales of Mental Tests, *MRC AIMS* UK Medical Research Council Autism Imaging Multicentre Study, *NVIQ* Nonverbal IQ, *PDD-NOS* pervasive developmental disorder not otherwise specified, *PEP* psychoeducational profile, *PPVT* peabody picture vocabulary test, *PSI* processing speed index, *RBS-R* repetitive behavior scale, Revised, *RL* receptive language, *SB* Stanford Binet Scales of Intelligence, *SRS* social responsiveness scale; *SSC* Simons simplex collection (individuals without relatives with ASD), *UNK* unknown, *VABS* Vineland adaptive behavior scales, *VIQ* verbal IQ, *VR* visual reception, *WAIS* Wechsler adult intelligence scales, *WISC* Wechsler intelligence scales for children, *3Di* developmental, DIMENSIONAL AND DIAGNOSTIC INTERVIEW

Significant p-values are indicated as follows: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, [†]Chronological age in years unless otherwise noted; [‡]Scaled or standard IQ scores used unless otherwise noted

[a]Visual Reception when covarying by composite language level (see [b]) and age

[b]Language composite = z score calculated from raw scores of MSEL expressive and receptive AE and VABS expressive and receptive subdomain AE, motor composite = z score calculated from raw scores of MSEL GM and FM subscales and VABS gross and fine motor subdomain covaried on visual reception

[c]Each dataset was considered as an independent sample

[d]Interview applied to an algorithm to calculate ADI-R equivalent scores

[e]Estimated Marginal Means at 36 months are reported in the table, profile analysis for MSEL included data from three ages—18, 24, and 36 months and a similar pattern at each age was found on all subscales (FM, VR, RL, EL)

[f]Age-matched at initial ASD diagnosis

[g]Average of quotients resulting from 6 subscales (DQ = developmental age × 100/chronological age)

[h]Ratio DQ calculated based on AE scores/chronological age for VR and FM to reflect nonverbal DQ (NVDQ) and RL and EL to reflect verbal DQ (VDQ)

[I]Subjects with a nonverbal MSEL of less than 12 months were excluded

[j]Emotion, Eye Gaze, Communication, Gestures, Sounds, Understanding, Object Use

[k]Varied measures including the following: DP, MPSMT, SB, Bayley, WISC

[l]Varied measures used including the K-ABC, WISC, WAIS, SB, Leiter and the Uzgiris-Hunt Scales with nonverbal mental age and expressive language testing used to estimate FSIQ

part of their clinical visit (sample characteristics are summarized in Table 2). Of note, the ATN protocol includes the Stanford-Binet, Fifth Edition (SB-5; Roid 2003), Bayley Scales of Infant and Toddler Development (Bayley 2006) or MSEL as acceptable cognitive-developmental measures for this age range. Instrument decisions are based on clinical judgment per the presentation, age and language level of the child.

For young children with ASD the MSEL is commonly used over the SB-5 as it has a lower floor, can be used with younger children and it contains items suitable for a wide range of functioning (normed from birth to 68 months). An additional consideration is that within the ASD population where language delays and differences are common, verbal ability has been found to have an influence on scores, even within the nonverbal subtests of the SB-5 (Lennen et al. 2010). In the ATN sample, the MSEL was administered to the majority of young children in the database and thus was selected as the focus of this paper targeting the early childhood window. Fewer young children were administered the Bayley (n = 126 attempted, 46 with missing data). Pooling data from the full SB-5 with the MSEL would have been challenging due to the differences between the measures. Additionally, the ATN data mostly contained older children and only SB-5 abbreviated forms (utilizing only 2 subtests, n = 2567) which would have made it even more uncertain. However, data on the 892 children who received the full SB-5 were also analyzed to evaluate whether or not the MSEL findings held true with older children receiving a different cognitive test.

There were 1,923 individuals who received an MSEL administration. Individuals were excluded if birthdate was missing precluding calculation of the child's chronological age (N = 46). We also excluded individuals who were administered the MSEL but whose chronological age was older than 68 months of age (N = 125), precluding calculation of valid MSEL scores. In addition, we excluded individuals who had missing raw scores in the four primary domains evaluated [Visual Reception (VR), Fine Motor (FM), Expressive Language (EL), Receptive Language (RL); N = 165]; scores for gross motor skills were rarely present in the registry and thus were excluded from all analyses. After applying these exclusion criteria, the full sample included 1,587 children (265 female).

Measures

Sociodemographic Information

Parent report provided child sex, age, race/ethnicity, and parent(s) education levels. Age was grouped into three distinct age ranges 18–35 months, 36–47 months, and 48–68 months. Race was categorized as White, Black/African American, Asian, Native American or Alaskan Native, Native Hawaiian or Pacific Islander, or more than one race. Due to the small numbers in some categories, these were collapsed into three categories for analyses: White, Black/African American and All Other Races. Ethnicity was categorized as either Hispanic/Latino origin or Not Hispanic/Latino. Parent education level was classified as the maximum educational level of the child's primary or secondary parent. For analyses, these were grouped as follows: high school graduate or less, some college, or college graduate or higher (see Table 2).

Cognitive Functioning

The current study consisted of individuals who had completed the Mullen Scales of Early Learning (MSEL; Mullen 1995). Raw scores were used to calculate domain T Scores and global standard scores as described in the manual. Individual domain scores were represented as T Scores (mean of 50 and standard deviation of 10) and included: Visual Reception (VR), Expressive Language (EL), Receptive Language (RL) and Fine Motor (FM). As described in the manual, over 99% of T Scores fall in the 20–80 range. Scores lower than 3 standard deviations below the mean based upon raw scores in each domain are termed "extremely rare" (Mullen 1995). In this sample, we analyzed children who obtained scores lower than 3 standard deviations below the mean as a "T Scores of < 20" category. The MSEL Early Learning Composite (ELC) Standard Score has a mean of 100 and a standard deviation of 15; on the Mullen ELC scores are permitted to range from 49 to 155, with 99% of scores falling between 55 and 145. The ELC score was used as an estimate of overall developmental cognitive level. The ELC score is calculated by summing the four domain T Scores to obtain a Cognitive T Score Sum which is used in the manual (Table C.3) to calculate the ELC score. The minimum Cognitive T Score Sum reported in the manual is 80 which corresponded to an ELC score of 49. In preliminary analyses, the ELC scores significantly deviated from the normal distribution. Inspection of the data revealed a distribution with significant positive skew. In this sample if the child obtained all four T Scores of < 20, we analyzed children as an "ELC of < 49" category. We created developmental categories derived from the ELC as follows: average to below average range (≥ 70), mild impairment range (49–69) and moderate to severe impairment range (< 49). Additionally, as this is commonly reported in the literature, we also calculated MSEL age-equivalent (AE) scores for each domain. These are not based on chronological age but rather are calculated by converting the raw scores obtained on each domain into age equivalents that are reported in months as detailed in the manual (Table C.4) (Mullen 1995). Age-equivalent scores may be more sensitive to lower performance, which can be commonplace in ASD samples (Carter et al. 2007; Hartley and Sikora 2009; Messinger et al. 2015).

The measure used for testing older and/or more able children was the Stanford-Binet, Fifth Edition (SB-5; Roid 2003). The SB-5 yields a full Scale IQ score (FSIQ) with a mean of 100 and a standard deviation of 15. It is normed on individuals from ages 2 to 85+ years of age.

Data Processing

In the ATN database, some errors were detected in the ELC and T Scores for several children and were corrected. Specifically, 210 children were missing ELC data, 465 children were missing at least one T Score and 63 children were missing at least one age equivalent score despite having raw scores for all four domains in the database. Thus, scores were recalculated and checked using the raw scores as described in the MSEL manual for all subjects in the sample (Mullen 1995). We used R scripts, based on the child's raw domain scores and chronological age as described in the Mullen procedures.

Missingness

To address missingness, we examined associations between missingness and sex for all sociodemographic and cognitive variables; no associations were significant. Analyses reported are based on complete case analysis for each variable; we report the sample sizes for each analysis separately.

Data Analysis

Pearson's χ^2 tests were used to analyze sex differences for categorical variables across sex. When very low frequencies occurred in some cells, the χ^2 statistic was computed using 2,000 simulated Monte-Carlo replicates (Hope 1968). Due to negatively skewed distributions, MSEL T Scores were compared using Wilcoxon-Mann-Whitney (WMW) nonparametric analyses. For MSEL AE scores, we used Welch's t tests assuming unequal variances. Developmental quotients (DQs) were calculated by dividing the subscale age-equivalent score by the child's chronological age and multiplying by 100, as recommended in the literature in order to avoid possible floor and ceiling effects (Messinger et al. 2013). Finally, in order to examine the effects of sex and age on cognitive scores, we used linear regression models with AE scores for each domain as the dependent variables.

Results

Sex Differences in Sociodemographic, Age, and Diagnostic Measures

Table 2 displays the characteristics of the sample. Sex was not associated with any sociodemographic measures or age (see Table 2). In this sample, 1,209 children (195 female) were diagnosed with Autistic Disorder, 261 children (53 female) were diagnosed with Asperger Syndrome or Pervasive Developmental Disorder (AS/PDD-NOS), and 117 children (17 female) did not have a recorded subtype. There was no association between sex and autism subtype [$\chi^2(1) = 2.67$, $p = 0.10$].

Sex Differences in Cognition

Are There sex Differences in Overall Cognitive Function?

To evaluate sex differences in cognition across the full sample (N = 1,587; 265 female), we first examined ELC scores as a categorical variable (< 49, 49–69, ≥ 70). Almost all subjects (98.3% of males and 98.2% of females) scored in the impaired range (ELC < 70). In fact, 17.1% of boys and 17.4% of girls tested fell in the lowest category (termed < 49). In the 1,315 children who had an ELC score of 49 or greater, scores on the MSEL composite were very low, with a median score lower than three standard deviations below the normative mean of 100 (median = 53, full range 49–138, interquartile range 49–65). Due to low frequencies of individuals with an ELC score ≥ 70, a χ^2 test with 2,000 simulated Monte-Carlo replicates was used. There was no association between ELC scores and sex ($\chi^2 = 0.08$, $p = 0.96$).

Are There Sex Differences in Specific Cognitive Domains?

When T Scores for individual domains were analyzed as categorical variables (< 20, ≥ 20), there were no associations with sex (see Table 3). For those children who received T Scores ≥ 20 (N = 622; 102 female), we compared the distributions of T Scores between sexes separately for each of the four domains. To assist in visualizing the distribution and probability density, violin plots (box plots with a rotated kernel density plot on each side) are presented in Fig. 1. Wilcoxon-Mann-Whitney tests revealed no effect of sex for any domain (VR: Z = − 1.08, $p = 0.28$; FM: Z = − 0.20, $p = 0.84$; RL: Z = − 1.70, $p = 0.09$; EL: Z = − 0.60, $p = 0.55$).

Due to a concern regarding floor effects and the large number of lower functioning children in this sample, we posited that it might be more appropriate to use age equivalent scores. Age equivalent scores are commonly reported in the literature, but unlike T scores, AE scores do not account for chronological age. Welch's t tests (assuming unequal variances) were used to compare the distributions of AE-scores between sexes (see Fig. 2). Three of the four comparisons yielded nonsignificant results [VR: t(398.4) = 2.42, $p = 0.02$; FM: t(380.59) = 1.61, $p = 0.11$; RL: t(377.41) = 1.58, $p = 0.11$; EL: t(371.22) = 0.63, $p = 0.53$] across the full sample (N = 1587; 265 female). The difference between boys

Table 2 Sociodemographic characteristics for the total sample and by sex

	Total sample ($N=1,587$)				Sex differences	
	n	%	Girls ($n=265$)	Boys ($n=1,322$)	χ^2	*p*
Age					2.90	0.24
18–35 months	646	40.7	45.3% (120)	39.8% (526)		
36–47 months	551	34.7	32.8% (87)	35.1% (464)		
48–68 months	390	24.6	21.9% (58)	25.1% (332)		
Race					5.43	0.07
White	1,141	71.9	67.9% (180)	72.7% (961)		
Black/African American	123	7.8	9.4% (25)	7.4% (98)		
All other races	228	14.4	18.5% (49)	13.5% (179)		
No data	95	6.0	4.2% (11)	6.4% (84)		
Ethnicity					1.73	0.19
Hispanic or Latino origin	181	11.4	14.0% (37)	10.9% (144)		
Not of Hispanic or Latino origin	1,307	82.4	81.5% (216)	82.5% (1091)		
No data	99	6.2	4.5% (12)	6.6% (87)		
Caregiver(s)' highest level of education					0.14	0.93
High school or less	262	16.5	17.4% (46)	16.3% (216)		
Some college	428	27.0	26.8% (71)	27.0% (357)		
College graduate or more	752	47.4	47.2% (125)	47.4% (627)		
No data	145	9.1	8.7% (23)	9.2% (122)		

All percentages within sociodemographic groupings are listed for the total sample in the first column and the second and third columns contain percentages broken down by sex

Table 3 Mullen scales of early learning domain T scores, for the total sample and by sex

	Total sample ($N=1587$)				Sex differences	
	n	%	Girls ($n=265$)	Boys ($n=1,322$)	χ^2	*p*
Visual reception					0.11	0.74
T score < 20	356	22.4	23.4% (62)	22.2% (294)		
T score ≥ 20	1231	77.6	76.6% (203)	77.8% (1028)		
Fine motor					0.09	0.76
T score < 20	380	23.9	23.0% (61)	24.1% (319)		
T score ≥ 20	1,207	76.1	77.0% (204)	75.9% (1003)		
Receptive language					0.61	0.44
T score < 20	921	58.0	60.4% (160)	57.6% (761)		
T score ≥ 20	666	42.0	39.6% (105)	42.4% (561)		
Expressive language					0.04	0.85
T score < 20	965	60.8	61.5% (163)	60.7% (802)		
T score ≥ 20	622	39.2	38.5% (102)	39.3% (520)		

All percentages within domains are listed for the total sample in the first column and the second and third columns contain percentages broken down by T score grouping by sex

and girls in the VR domain remained significant after applying an adjusted Bonferroni threshold (nominal $p=0.0125$). However, when linear regression models included both sex and age (in months) as predictor variables, sex was no longer a significant predictor of AE scores for VR and the other 3 domains as well (see Table 4). By contrast, age was systematically associated with AE scores on each MSEL domain. Thus, when chronological age was accounted for in AE scores, no effect of sex was discernable. Likewise, when a developmental quotient was calculated, no sex differences in mean domain scores were found across the domains [VR: $t(385.03)=1.27$, $p=0.20$; FM: $t(370.77)=0.11$, $p=0.91$; RL: $t(366.3)=0.87$, $p=0.38$; EL: $t(358.53)=-0.26$, $p=0.79$].

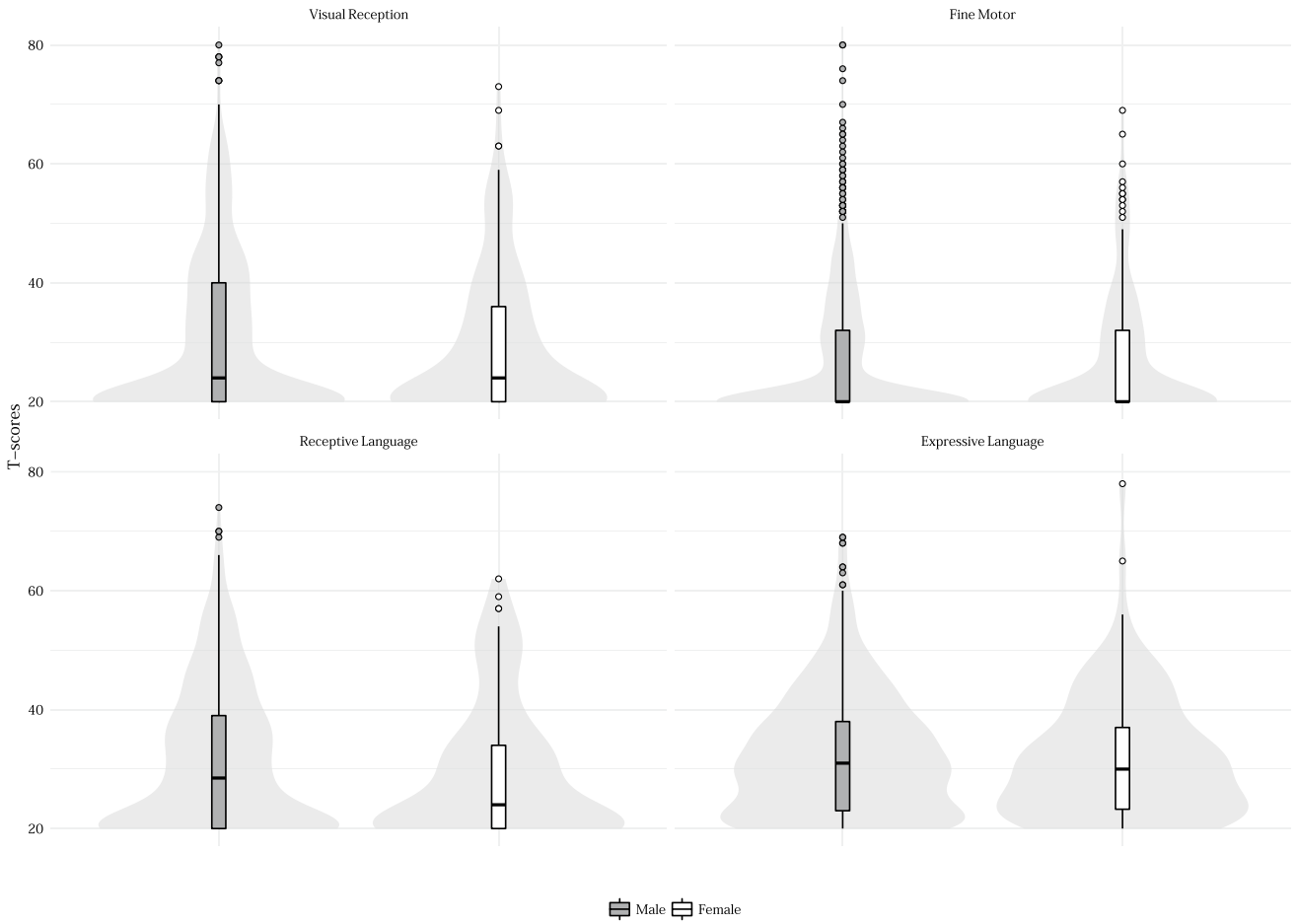


Fig. 1 Distributions of T scores by Sex for each MSEL subscale

Could Differences be Influenced by Overall Level of Ability or Chronological Age?

We further examined sex differences among children who received 0–4 categorical T Scores < 20 across domains. The proportions of subjects were as follows: 31.3% (417 males, 80 females) received zero T Scores < 20, 15.4% (205 males, 40 females) received only one T Score < 20; 27.1% (355 males, 75 females) received two out of four T Scores < 20; 9.0% (119 males, 24 females) received three out of four domain scores of T Scores < 20 and 17.1% (226 males, 46 females) received T Scores < 20 on all four domains. There was no association between the number of T Scores < 20 and sex [$\chi^2(4) = 0.33$, $p = 0.99$]. Follow-up analyses focusing on domain-specific T Scores (< 20, ≥ 20) also showed no associations with sex for any domain evaluated [VR: $\chi^2(1) = 0.11$, $p = 0.74$; FM: $\chi^2(1) = 0.09$, $p = 0.76$; RL: $\chi^2(1) = 0.61$, $p = 0.44$; EL: $\chi^2(1) = 0.04$, $p = 0.85$).

To assess whether sex differences in cognition were present only within more restricted age ranges, we analyzed associations between ELC scores as a categorical variable (< 49, 49–69, ≥ 70) and sex, within three distinct age ranges (18–35 months: n = 646; 36–47 months: n = 550; 48–68 months: n = 390). Due to low frequencies in individual cells, χ^2 tests with 2,000 simulated Monte-Carlo replicates were used; these analyses consistently showed no associations between sex and ELC scores for any of the three age ranges (see Table 5).

Does the Pattern of Discrepancy Between Cognitive Domains Vary by Sex?

Similar patterns were seen in terms of relative strengths and weaknesses across the sexes. Both males and females demonstrated strengths in the domains of Visual Reception and Fine Motor with over 75% of the sample obtaining T Scores ≥ 20. For both sexes, similar areas of weakness were seen in the Receptive Language and Expressive Language domains where only approximately 40% of the sample was able to obtain a T Score > 20 (See Table 5 for additional details).

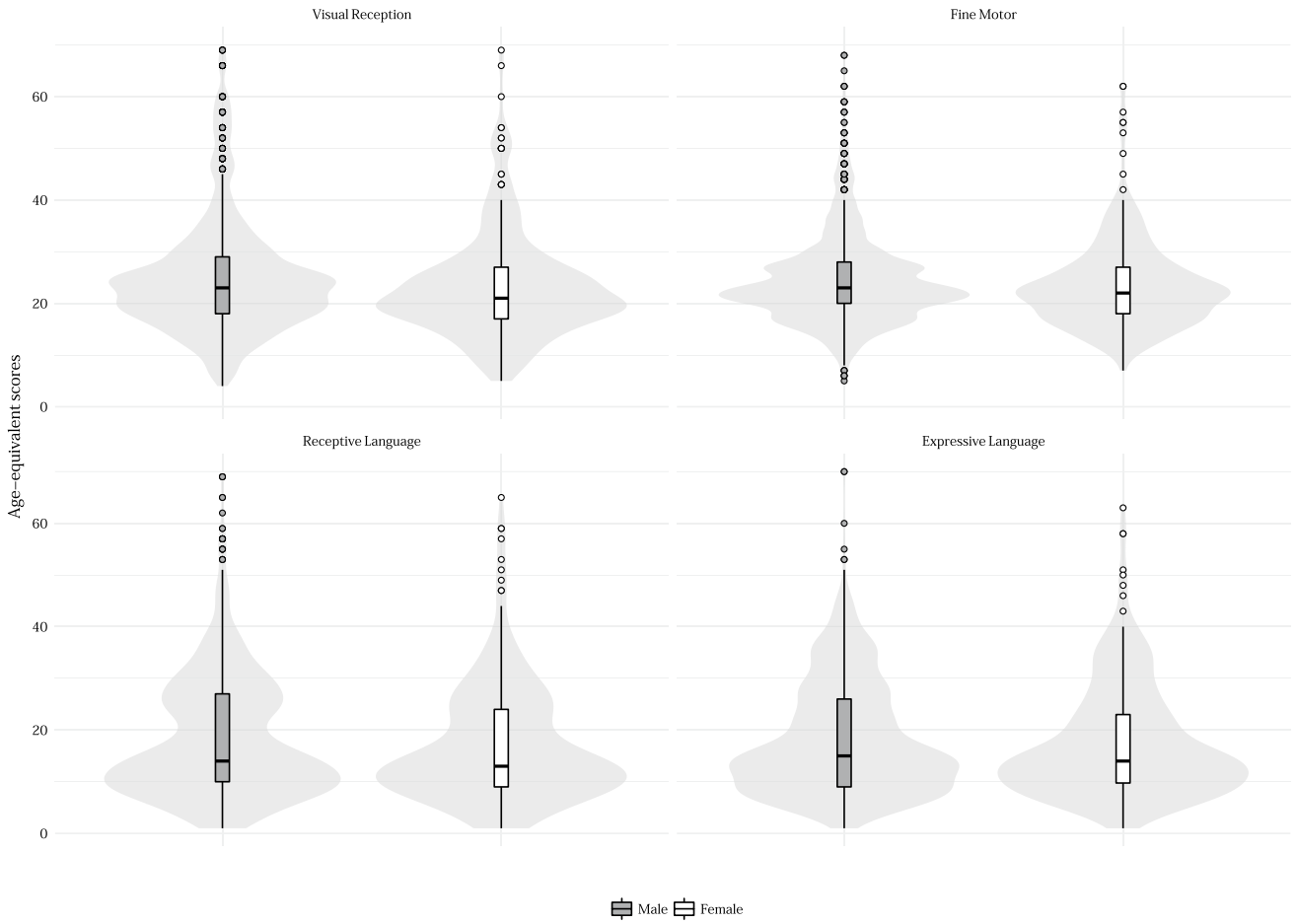


Fig. 2 Distributions of age equivalent scores by sex for each MSEL subscale

Table 4 Linear regression of mullen scales of early learning age equivalent scores on sex and age

	B (SE)	t value	*p*
Visual reception			
Sex	− 1.11 (0.70)	− 1.59	0.113
Age (months)	0.41 (0.02)	16.97	0.001
Fine motor			
Sex	− 0.43 (0.54)	− 0.80	0.426
Age (months)	0.36 (0.02)	19.47	0.001
Receptive language			
Sex	− 0.70 (0.75)	− 0.93	0.356
Age (months)	0.37 (0.03)	14.48	0.001
Expressive language			
Sex	0.02 (0.69)	0.02	0.991
Age (months)	0.33 (0.02)	13.68	0.001

Finally, we examined the strength of the correlations between domain scores within males and females and compared it across the sexes. When comparing correlations of age equivalent scores between males and females on the domain scores, no statistically significant differences were found [VR × FM: Z = 0.55, $p = 0.58$; VR × RL: Z = 0.74, $p = 0.46$; VR × EL: Z = 1.16, $p = 0.25$; FM × RL: Z = – 0.29, $p = 0.78$; FM × EL: Z = 0.26, $p = 0.79$; RL × EL: Z = 1.28, $p = 0.20$], see Fig. 3.

Is a Similar Pattern Seen in Individuals Who Received the SB-5?

Among the 892 children tested with the full SB-5 (10 subtests), there were 759 males [mean age = 91.7 months (SD = 40.2); mean FSIQ = 76.7 (SD = 22.2)] and 133 females [mean age = 94.0 months (42.2); mean FSIQ = 75.8 (25.0)]. There was no difference between males and females for age [t (891) = -0.63, p = 0.53] or FSIQ [t (891) = – 0.85, $p = 0.39$]. This finding of no difference in cognition between sexes was also seen in children who were given only the abbreviated (2 subtest) screening measure from the SB-5. When compared to children evaluated with the Mullen, children tested with the full SB-5 were older [F (3, 2201) = 4.14,

Table 5 Mullen scales early learning composite (ELC) score, by sex and age band

Age range	Total sample ($N=1,587$)		Sex differences			
	n	%	Girls ($n=265$)	Boys ($n=1322$)	χ^2	*p*
18–35 months			*n=120*	*n=526*	0.39	0.74
ELC score < 49	8	1.2	1.7% (2)	1.1% (6)		
ELC score 49–69	624	96.6	96.7% (116)	96.6% (508)		
ELC score 70 +	14	2.2	1.7% (2)	2.3% (12)		
36–47 months			*n=87*	*n=464*	1.64	0.38
ELC score < 49	85	15.4	19.5% (17)	14.7% (68)		
ELC score 49–69	462	83.8	79.3% (69)	84.7% (393)		
ELC score 70 +	4	0.7	1.1% (1)	0.6% (3)		
48–68 months			*n=58*	*n=332*	0.43	0.82
ELC score < 49	179	45.9	46.6% (27)	45.8% (152)		
ELC score 49–69	202	51.8	50.0% (29)	52.1% (173)		
ELC score 70 +	9	2.3	3.4% (2)	2.1% (7)		

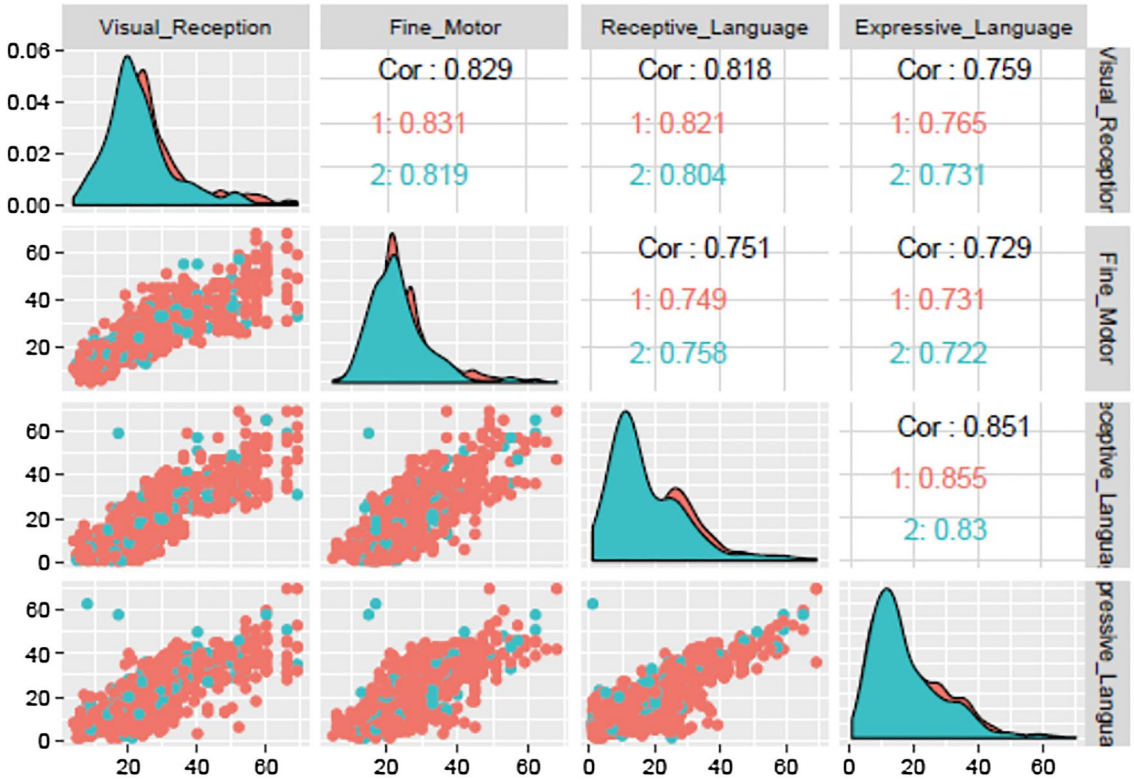


Fig. 3 Correlations among age-equivalents by domain in males and females

$p=0.04$] and had higher IQ scores [F (3, 2201) = 261.03, p < .001]; however, these differences were unrelated to sex [F (3, 2201) = 1.22, p = 0.27]. Thus, no sex difference in cognitive functioning was found in older children tested with a different measure, indicating that our findings are unlikely to be restricted to a particular age group, level of functioning or cognitive test.

Discussion

In this large sample of young children with ASD assessed with the MSEL, no sex differences were found in overall cognitive ability, specific domain scores or patterns of cognitive performance. These results persisted after stratifying by developmental level. Closely similar proportions of males

and females were seen in different levels of cognitive function (ELC of < 49, 49– 69 and 70+). When examining function by domain, we also saw no sex differences in Visual Reception, Fine Motor, Receptive Language or Expressive Language domain age equivalent scores or T Scores. Nor did we find sex specific patterns of strengths and weaknesses across domains. When age equivalents were utilized in a follow-up analysis, males with ASD were found to score marginally higher on the Visual Reception subtest, although this result did not retain significance when chronological age was accounted for. For both sexes, the age equivalents in the domains of Visual Reception and Fine Motor were areas of relative strength, which is in line with other research utilizing the MSEL in young ages (Carter et al. 2007; Hartley and Sikora 2009; Zwaigenbaum et al. 2012). In sum, a very robust lack of sex differences was found throughout the various comparisons performed in this study, which is consistent with more recent research in young children utilizing the same measure (Hartley and Sikora 2009; Reinhardt et al. 2015).

The current findings are in contrast with those from historical studies, which indicated that females with ASD were more likely to present with lower IQ and associated intellectual disability (IQ less than 70) than their male counterparts (Bryson et al. 1988; Lord et al. 1982; Wing 1981). Our findings also do not seem to align with the contemporary theory of "female protective effect" (e.g., Jacquemont et al. 2014; Werling 2016). This variation of a multiple threshold liability model for ASD risk posits that males and females have different minimum variant loads at which they present an ASD phenotype. Affected females would be expected to carry a greater risk variant load, on average, than affected males which would result in more pronounced cognitive deficits in girls with ASD. Several factors may explain the present study's contrasting findings. Cognitive differences noted in earlier research (males outperforming females) may be related to evolving diagnostic criteria, measurement tool heterogeneity, and age effects, as well as, selection biases, especially regarding the underrepresentation of higher functioning females in previous studies (Halladay et al. 2015). Additionally, and unlike prior studies that excluded impaired children, our sample included young children who performed very poorly on developmental testing (T Score < 20).

Evolving diagnostic criteria have allowed for increasing identification of more subtle ASD symptoms. Historically, females formally diagnosed with ASD tended to have more severe core symptoms and increased ASD symptoms have been associated with increased cognitive impairment (Giarelli et al. 2010; Joseph et al. 2002; Lai et al. 2015). Therefore, earlier findings may have been influenced by poorer identification of ASD among higher functioning females (Lord and Schopler 1985; Robinson et al. 2013). Additionally, older findings may be reflective of higher functioning females being diagnosed later, and thus being under-represented in analyses of young children with ASD (Dworzynski et al. 2012; Reinhardt et al. 2015; Werling and Geschwind 2013). By contrast, our findings of no sex differences were obtained in a large, recently and systematically ascertained sample; they are generally consistent with more recent research in young children utilizing the MSEL (Hartley and Sikora 2009; Reinhardt et al. 2015).

The possibility exists that the current results are related specifically to the testing instrument utilized. When examining the literature, one must consider whether or not a floor effect and the specific scoring conventions of the MSEL may have obscured sex differences in performance. In studies of similar populations tested with the MSEL, participants falling more than 3 standard deviations below the mean were not separately described nor specifically examined (Carter et al. 2007; Hartley and Sikora 2009). Moreover, some studies excluded the lowest functioning children (AE < 12 months) from analyses (Reinhardt et al. 2015). To address this potential pitfall, we performed in-depth analysis of children who obtained the lowest scores on the MSEL. Yet, no sex differences were found amongst the most cognitively delayed children (see Table 3 for additional information based on age bands). Likewise, similar proportions of males and females obtained the lowest categorical scores, either alone or in combination, across the Visual Reception, Fine Motor, Receptive Language and Expressive Language domains (T Score < 20), further supporting no sex differences in this lowest functioning group.

Another potential concern is that our sample tested with the MSEL may have over-represented older children with marked developmental delays who may not have been able to complete other more rigorous IQ testing. Children with more age typical presentation and language may have more frequently completed other ATN protocol options, (i.e., the SB-5 or Bayley). However, sex differences were also absent when we examined MSEL ELC and domain scores after stratification on age groups. Furthermore, no sex differences were found among children with higher MSEL scores defined as either ELC ≥ 49 or ELC ≥ 70. Finally, we performed additional checks and showed that older and more able children who received the SB-5 equally showed no sex differences in FSIQ scores.

Sample composition achieved through multiple referral pathways may also have impacted study findings. Our sample was drawn from a clinically referred population, selected in specialty diagnostic and research settings associated with the ATN. Research recruited samples may have higher cognitive function than those described in clinically referred samples; in young children from clinical samples, co-occurring intellectual disability is common (Hill et al. 2015; Myers et al. 2018). High-risk baby sibling research may be more likely to include children with non-impaired cognitive function given the prospective nature of the sample and the frequent

assessments (Messinger et al. 2015; Zwaigenbaum et al. 2012). More severely affected children are also likely underrepresented in research studies due to behavioral challenges that can arise in working with these groups. Additionally, some research samples recruit only high functioning children (IQ > 70), which may limit the generalizability of findings (Koyama et al. 2009). Research recruited samples based on simplex or multiplex status add further heterogeneity. For example, males have outperformed females in simplex families on nonverbal IQ testing, but this pattern has not held for multiplex families (Banach et al. 2009).

Age heterogeneity is another source of variance between studies. As noted above, females with ASD tended to be diagnosed later, and, as a result, they may have been excluded from earlier studies of young children (Giarelli et al. 2010). Within our preschool sample no sex differences were found in our study even when narrow age bands were examined; it is possible that earlier diagnosis of females with less severe autism symptoms "washed out" the historical sex difference findings in the literature. It is also possible that sex differences in cognition emerge as children grow. Although our study design does not permit us to test in full this hypothesis, it is noteworthy that no trend for emerging sex differences with increasing age could be detected between ages 18 and 68 months in our sample.

This study has several strengths, most notably its large sample size of young children with ASD and the inclusion of all participants in data analyses regardless of cognitive level. The clinical implications of our findings are that customizing behavioral interventions to cognitive level will continue to be important; however, intervention planning need not be varied by sex. In assessment settings, intellectual disability and global developmental delay are common differential diagnoses for young children presenting with concerns of ASD. Our findings suggest that sex does not influence the probability of co-occurrence of intellectual disability with ASD, and should therefore prevent sex bias and stereotypes in the context of those assessments.

This study also has several limitations. All participants were clinically referred to specialty diagnostic centers, thus reflecting individuals for whom ASD concerns were noted at a young age. This sample may not be representative of all individuals with ASD in the general population, especially those with milder ASD symptoms, less language impairment and less functional impairment. It is possible that females in the broader community may need to exhibit more significant core ASD symptoms to meet threshold for referral as compared to males (Solomon et al. 2012). If so, our results may have failed to detect a female cognitive advantage that only studies of older age groups, or studies of truly representative populations of ASD cases, may be able to document. Given the extent of global developmental delays, and especially expressive and receptive language deficits, in our sample, bias could have occurred in our study if girls with ASD have relatively preserved language skills. Likewise, the sample may have under included boys with mild symptomatology and preserved language/cognitive development which, in turn, could have obscured findings. Nevertheless, the fact that no sex differences in either direction could be found in the portion of our sample with cognitive level in the normal range (ELC > 70) makes it less likely that the aforementioned possible issues have affected our results. Other data are needed to further examine these possibilities.

Additional limitations included the restricted age at inclusion, which prevents extrapolation of results to children outside of the 18–68 month age range. It is possible that children tested at young ages may show more variability in developmental assessment results. Additionally, we only examined children who were administered the MSEL and, therefore, our results are limited to the domains and constructs assessed by this measure. Finally, the ATN dataset consists of data collected across several clinical sites, each with testers with unknown inter-rater reliability. While this may result in increased measurement error in test scores, the large sample size should have compensated for this source of variance. Moreover, there is no reason to postulate that measurement imprecision would differentially apply to males and females, therefore providing confidence in our sex difference results.

Future directions may consider examining the relationship between severity of ASD symptoms and developmental functioning, as increased autism symptom severity has been linked with lower cognitive function (Joseph et al. 2002). Following children longitudinally may help enhance our understanding of developmental cognitive function within the ASD population. As maturation allows for the use of richer cognitive and intelligence tests and the MSEL is only a developmental assessment, it will be important to examine if the lack of differences between sexes is maintained over time. More comprehensive neuropsychological and/or cognitive tests in follow-up longitudinal studies may reveal cognitive differences between the sexes that we could not observe in this very young and cross-sectional sample. Comprehensive and systematic investigation into phenotypically distinct groups (e.g., higher/lower functional self-care and adaptive skills and developmental functioning), as well as population based representative samples across different age groups is required to more fully understand whether meaningful cognitive differences exist across males and females with ASD over the lifespan (Lai et al. 2015).

The current findings of no cognitive differences between young boys and girls with ASD are of importance due to the large sample size, the more narrowly defined age range, and reliance on a single cognitive measure. Although our findings align with some smaller studies of preschoolers assessed with the MSEL (Hartley and Sikora 2009; Reinhardt et al.

2015), our results are discrepant from other studies which suggest a possible female advantage (Zwaigenbaum et al. 2012) or sex specific patterns of strengths and weaknesses (Carter et al. 2007). Further research is needed to clarify these differences and systematic oversampling of girls in ASD research is indicated.

Acknowledgments This research was conducted as part of the Autism Treatment Network (ATN). The ATN is funded by Autism Speaks and a cooperative agreement (UA3 MC 11054) from the Health Resources and Services Administration to Massachusetts General Hospital. We acknowledge the members of the ATN for use of the data and the families who participated in the Registry. We acknowledge Naomi Knoble for her contributions. A version of this paper was presented as a poster in 2014 at the International Meeting for Autism Research (IMFAR), Atlanta, Georgia.

Author Contributions Authors SWD and LHS contributed to the literature review and writing of the manuscript. Authors SWD, LHS, EF and APH contributed to the study design and editing of the manuscript. Authors APH and JM contributed to the statistical analyses.

Funding This research was conducted as part of the Autism Treatment Network (ATN). The ATN is funded by Autism Speaks and a cooperative agreement (UA3 MC 11054) from the Health Resources and Services Administration to Massachusetts General Hospital.

Compliance with Ethical Standards

Conflict of interest Susanne W. Duvall, Ph.D. declares that she has no conflict of interest. Lark Huang-Storms, Ph.D. declares that she has no conflict of interest. Alison Presmanes Hill Ph.D. declares that she has no conflict of interest. Julianne Myers, BA declares that she has no conflict of interest. Eric Fombonne, MD declares that he has 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 Informed consent was obtained from all individual participants included in the study.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th edn.). text rev). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th edn.). Washington, DC: American Psychiatric Association.

Ankenman, K., Elgin, J., Sullivan, K., Vincent, L., & Bernier, R. (2014). Nonverbal and verbal cognitive discrepancy profiles in autism spectrum disorders: Influence of age and gender. *American Journal on Intellectual and Developmental Disabilities, 119*(1), 84–99. https://doi.org/10.1352/1944-7558-119.1.84.

Baio, J. (2012). Prevalence of autism spectrum disorders: autism and developmental disabilities monitoring network, 14 Sites, United States, 2008. In: *Morbidity and mortality weekly report. Surveillance summaries*. Volume 61, Number 3. *Centers for Disease Control and Prevention*.

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*(6), 1.

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.

Bayley, N. (2006). *Bayley scales of infant and toddler development* (3rd edn.). TX: Hartcourt Assessments, Inc., San Antonio.

Blumberg, S. J., Bramlett, M. D., Kogan, M. D., Schieve, L. A., Jones, J. R., & Lu, M. C. (2013). Changes in prevalence of parent-reported autism spectrum disorder in school-aged US children: 2007 to 2011–2012. National Center for Health Statistics Reports. Number 65. National Center for Health Statistics.

Bryson, S. E., Clark, B. S., & Smith, I. M. (1988). First report of a Canadian epidemiological study of autistic syndromes. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 29*(4), 433–445.

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. https://doi.org/10.1007/s10803-006-0331-7.

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. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders, 33*(4), 365–382. https://doi.org/10.1023/A:1025054610557.

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, 329–340.

Giarelli, E., Wiggins, L. D., Rice, C. E., Levy, S. E., Kirby, R. S., Pinto-Martin, J., & Mandell, D. (2010). Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disability and Health Journal, 3*(2), 107–116. https://doi.org/10.1016/j.dhjo.2009.07.001.

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., & Taylor, J. L. (2015). Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. *Molecular Autism, 6*(1), 36.

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.

Hill, A. P., Zuckerman, K., & Fombonne, E. (2015). Epidemiology of autism spectrum disorders. F. R. Volkmar & R. Paul (Eds.), *Handbook of autism spectrum disorders* (5th edn.). New York: Wiley.

Hope, A. C. (1968). A simplified Monte Carlo Significance test procedure. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30 (3), 582–598. JSTOR, http://www.jstor.org/stable/2984263.

Howe, Y. J., O'Rourke, J. A., Yatchmink, Y., Viscidi, E. W., Jones, R. N., & Morrow, E. M. (2015). Female autism Phenotypes investigated at different levels of language and developmental abilities. *Journal of Autism and Developmental Disorders, 45*(11), 3537–3549. https://doi.org/10.1007/s10803-015-2501-y.

Jacquemont, S., Coe, B. P., Hersch, M., Duyzend, M. H., Krumm, N., Bergmann, S., & Eichler, E. E. (2014). A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. *The American Journal of Human Genetics, 94*(3), 415–425.

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social- communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43*(6), 807–821.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*(3), 217–250.

Koyama, T., Kamio, Y., Inada, N., & Kurita, H. (2009). Sex differences in WISC-III profiles of children with high-functioning pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 39*(1), 135–141. https://doi.org/10.1007/s10803-008-0610-6.

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 & Adolescent Psychiatry, 54*(1), 11–24.

Lai, M. C., Lombardo, M. V., Ruigrok, A. N., Chakrabarti, B., Wheelwright, S. J., Auyeung, B., et al. (2012). Cognition in males and females with autism: Similarities and differences. *PLoS One, 7*(10), e47198.

Lajonchere, C., Jones, N., Coury, D. L., & Perrin, J. M. (2012). Leadership in health care, research, and quality improvement for children and adolescents with autism spectrum disorders: Autism Treatment Network and Autism Intervention Research Network on Physical Health. *Pediatrics, 130*(Suppl 2), S62–S68.

Lennen, D. T., Lamb, G. D., Dunagan, B. J., & Hall, T. A. (2010). Verbal prowess equals higher IQ: Implications for evaluating autism. *Research in Autism Spectrum Disorders, 4*(1), 95–101.

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 & Adolescent Psychiatry, 56*(6), 466–474.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic observation schedule—WPS (ADOS-WPS)*. Los Angeles: Western Psychological Services.

Lord, C., & Schopler, E. (1985). Differences in sex ratios in autism as a function of measured intelligence. *Journal of Autism and Developmental Disorders, 15*(2), 185–193.

Lord, C., Schopler, E., & Revicki, D. (1982). Sex differences in autism. *Journal of Autism and Developmental Disorders, 12*(4), 317–330.

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.

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., et al. (2013). Beyond autism: a baby siblings research consortium study of high-risk children at three years of age. *Journal of the American Academy of Child & Adolescent Psychiatry, 52*(3), 300–308.

Messinger, D. S., Young, G. S., Webb, S. J., Ozonoff, S., Bryson, S. E., Carter, A., et al. (2015). Early sex differences are not autism-specific: A Baby Siblings Research Consortium (BSRC) study. *Molecular autism, 6*(1), 32.

Myers, S. M., Voigt, R. G., Colligan, R. C., Weaver, A. L., Storlie, C. B., Stoeckel, R. E., et al. (2018). Autism spectrum disorder: Incidence and time trends over two decades in a population-based birth cohort. *Journal of Autism and Developmental Disorders,* 1–20.

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines: AGS.

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.

Postorino, V., Fatta, L. M., De Peppo, L., Giovagnoli, G., Armando, M., Viari, 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.

Reinhardt, V. P., Wetherby, A. M., Schatschneider, C., & Lord, C. (2015). Examination of sex differences in a large sample of young children with autism spectrum disorder and typical development. *Journal of Autism and Developmental Disorders, 45*(3), 697–706.

Robinson, E. B., Lichtenstein, P., Anckarsäter, H., Happé, F., & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences, 110*(13), 5258–5262.

Roid, G. H. (2003). *Stanford-binet intelligence scale*. Rolling Meadows: Riverside Publishing Company, Itasca, IL.

Rutter, M. (1978). Diagnosis and definition of childhood autism. *Journal of Autism and Childhood Schizophrenia, 8*(2), 139–161.

Solomon, M., Miller, M., Taylor, S. L., Hinshaw, S. P., & Carter, C. S. (2012). Autism symptoms and internalizing psychopathology in girls and boys with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(1), 48–59.

Tsai, L. Y., & Beisler, J. M. (1983). The development of sex differences in infantile autism. *The British Journal of Psychiatry, 142*(4), 373–378.

Volkmar, F. R., Szatmari, P., & Sparrow, S. S. (1993). Sex differences in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 23*(4), 579–591.

Werling, D. M. (2016). The role of sex-differential biology in risk for autism spectrum disorder. *Biology of Sex Differences, 7*(1), 58.

Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology, 26*(2), 146–153.

Wing, L. (1981). Sex ratios in early childhood autism and related conditions. *Psychiatry Research, 5*(2), 129–137. https://doi.org/10.1016/0165-1781(81)90043-3.

Zwaigenbaum, L., Bryson, S. E., Szatmari, P., Brian, J., Smith, I. M., Roberts, W., et al. (2012). Sex differences in children with autism spectrum disorder identified within a high-risk infant cohort. *Journal of Autism and Developmental Disorders, 42*(12), 2585–2596.