Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jennifer Moriuchi Date

Clarifying sex differences in social disability in Autism Spectrum Disorder

By

Jennifer Moriuchi
Doctor of Philosophy

Clinical Psychology

_____
Ami Klin, Ph.D.
Advisor

_____
Linda Craighead, Ph.D.
Committee Member

_____
Jocelyne Bachevalier, Ph.D.
Committee Member

_____
Warren Jones, Ph.D.
Committee Member

_____
Celine Saulnier, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Clarifying sex differences in social disability in Autism Spectrum Disorder

By

Jennifer Moriuchi
B.A., Wellesley College, 2009
M.A., Emory University, 2013

Advisor: Ami Klin, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2017

Abstract

Clarifying sex differences in social disability in Autism Spectrum Disorder
By Jennifer Moriuchi

The diagnostic prevalence of Autism Spectrum Disorder (ASD) is significantly higher in males compared to females. Due to the relatively low frequency of affected females, many studies have excluded females or have included sample sizes too small to detect potential sex differences. As a result, understanding of sex-specific differences in ASD-related social behaviors and how those differences may relate to the etiology of the disorder remains limited. The current study leveraged multiple levels of behavioral metrics to identify similarities and differences in the manifestation of social disability across sexes (Study 1) and to determine shared and sex-specific predictors of positive functional outcomes (Study 2). In a sample of 259 school-age children, comprising 200 participants with ASD (136 male, 64 female) and 59 age-matched typically-developing peers (27 male, 32 female), we assessed the clinical phenotype of social disability based on parent report, clinician rating, and performance-based measures of social visual engagement. In Study 1, males and females with ASD did not differ based on parent report or clinician rating. Performance-based measures of social visual engagement were significantly less divergent from normative same-sex patterns in females with ASD compared to males. The factor structure of social disability across levels of measurement also differed in males and females. In Study 2, predictors of stronger adaptive social functioning did not differ across sexes. These results provide insight into the possible under-identification of females with ASD and emphasize the importance of considering sex-specific differences in typical development.

Clarifying sex differences in social disability in Autism Spectrum Disorder

By

Jennifer Moriuchi
B.A., Wellesley College, 2009
M.A., Emory University, 2013

Advisor: Ami Klin, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2017

Table of Contents

List of Tables and Figures

Clarifying sex differences in social disability in Autism Spectrum Disorder

In over 70 years of research since Leo Kanner's and Hans Asperger's initial case descriptions of Autism Spectrum Disorder (ASD) in the 1940s, the most consistent sex-related finding in ASD research is the disproportionate ratio of males relative to females diagnosed with ASD. Prevalence studies over time, across countries and cultures, and across evolving diagnostic definitions have reported a male-biased sex ratio of approximately 4 affected males for every 1 affected female (Christensen et al., 2016; Fombonne, 2009; Kim et al., 2011; Wing, 1981). Though not an epidemiological study, it is notable that of Kanner's original 11 case descriptions, 8 were male, and 3 were female (Kanner, 1943). Similarly, of Asperger's original 4 cases, all were male. Asperger specifically highlighted that "the autistic children we have seen are almost exclusively boys" (Asperger, 1944; Frith, 1991). At present, the most recent prevalence estimates in the United States across the Autism and Developmental Disabilities Monitoring Network indicate a sex ratio of 4.5 affected males for every 1 affected female (Christensen et al., 2016). Although this sex ratio is moderated by level of cognitive functioning, with a less discrepant sex ratio among children with comorbid cognitive impairments, prevalence is consistently lower in females relative to males across the spectrum (Christensen et al., 2016; Fombonne, 2009).

Several hypotheses have been proposed to explain the discrepancy in diagnostic frequency. One hypothesis is that prevalence estimates are skewed due to a bias against identifying females with ASD using current diagnostic approaches (e.g., Dworzynski, Ronald, Bolton, & Happé, 2012; Giarelli et al., 2010). Because diagnostic measures were developed in predominantly male samples (Koenig & Tsatsanis, 2005), they may not

have adequate sensitivity to the clinical phenotype of ASD in females. Other biologically-based hypotheses have suggested either additional risk factors for ASD in males (e.g., increased fetal testosterone exposure) or protective factors against ASD in females (e.g., requiring a higher burden of genetic variations). Consistent with these hypotheses, females with ASD carry a higher heritable mutational load than males in, in *de novo* cases, have higher rates of copy number and single nucleotide variants that disrupt a larger number of genes (Levy et al., 2011; Neale et al., 2012; Sanders et al., 2011).

These accounts all make the fundamental assumption that the clinical profile of social disability in ASD does not differ in males and females. Biological differences observed across sexes, such as distinct patterns of interhemispheric connectivity (Nordahl et al., 2015) or neural responses to social stimuli (Coffman, Anderson, Naples, & McPartland, 2013), have been attributed to sex-specific risk or protective factors. However, if the clinical presentation of ASD does differ across sexes, biological differences may relate to sex-specific features of social disability or sex-linked endophenotypes rather than directly revealing more about the core pathogenic mechanisms of ASD. Both possibilities advance the understanding of ASD, but the distinction is meaningful.

Based on the extant literature, it is unclear whether the assumption holds true that the clinical profile of social disability is identical across sexes. Despite the longstanding and consistent findings on sex-based discrepancies in ASD prevalence as well as the growing literature on genetic and neurobiological differences across sexes in ASD, research directly addressing sex differences in clinical phenotype has been limited. The

relatively low frequency of females with ASD has itself been a substantial hurdle; because there are fewer diagnosed females, many studies have simply excluded females from analyses or have lacked the power to detect anything less than large statistical effects in between-group differences. Nevertheless, there is emerging evidence that sex differences contribute to the vast phenotypic heterogeneity in ASD (e.g., Frazier, Georgiades, Bishop, & Hardan, 2014; Lai, Lombardo, Auyeung, Chakrabarti, & Baron-Cohen, 2014). This heterogeneity presents a major obstacle to research on the etiology of ASD, to the identification of diagnostic biomarkers, and to the implementation of successful, individualized treatments. Clarifying the differences and similarities across sexes is consequently critical to determining sex-specific risk or protective factors in ASD pathogenesis and to advancing clinical care for children with ASD.

Autism Spectrum Disorder

The diagnosis of ASD is defined by atypical reciprocal social interaction, including qualitative impairments in verbal and nonverbal social communication as well as the presence of repetitive and stereotyped patterns of behavior (American Psychiatric Association, 2013). The diagnostic category of ASD was introduced in the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* to encapsulate related neurodevelopmental disorders previously under the category of Pervasive Developmental Disorders, which had included Autistic Disorder, Asperger's Disorder, and Pervasive Developmental Disorder-Not Otherwise Specified (American Psychiatric Association, 2004). To meet diagnostic criteria for ASD, as a neurodevelopmental disorder, symptoms must be present from early in development. Diagnostic stability is high beginning at 18 to 24 months (Chawarska, Klin, Paul, Macari, & Volkmar, 2009; Lord et al., 2006; Ozonoff

et al., 2015). Early-emerging symptoms often include diminished eye contact, limited showing or pointing, limited vocalizations directed towards others, and unusual sensory examination of objects and are associated with difficulties in establishing and maintaining social relationships with peers (e.g., Chawarska et al., 2014; Zwaigenbaum et al., 2009).

Diagnostic evaluation of ASD. At present, the diagnosis of ASD is defined and determined on the basis of behavioral symptoms. Although the developmental pathogenesis of ASD has a biological basis, research to identify specific diagnostic biomarkers with translational utility in clinical practice is still emerging (McPartland, 2016). The marked phenotypic variability across the spectrum has been one fundamental obstacle in biomarker development (Volkmar, Lord, Bailey, Schultz, & Klin, 2004). The study of genetic diagnostic biomarkers has been complicated by the fact that ASD does not follow a simple, Mendelian genetic transmission pattern and instead includes a high prevalence of *de novo* mutations (Iossifov et al., 2014). Twin studies and studies of younger siblings of individuals with ASD have shown that ASD has a genetic basis (Folstein & Rutter, 1977; Hallmayer et al., 2011; Ozonoff et al., 2011), and in total, approximately 30 to 40% of ASD cases can be associated with a specific genetic liability (Schaefer & Mendelsohn, 2013). Nevertheless, approximately 70% of affected individuals in multiplex families do not share the same genetic liability (Yuen et al., 2015), and among the genes implicated in ASD, none has been shown to have more than 1 to 2% penetrance (Geschwind, 2011). The study of neural biomarkers has similarly been complicated by the limited diagnostic specificity of identified between-group differences at an individual level (Elsabbagh & Johnson, 2016; Kaiser et al., 2010). As a

result, ASD remains a behaviorally-defined disorder, and current diagnostic best-practice guidelines rely on subjective assessment of behavior as deviating from typical development.

Consensus clinician-assigned, best-estimate diagnosis is the widely accepted diagnostic gold standard for ASD (Lord & Corsello, 2005). Recommended best-practice parameters for ASD diagnosis emphasize gathering information using multiple modes of measurement to inform clinical judgment regarding each of the DSM diagnostic criteria. Practice parameters for diagnostic evaluation of ASD prescribe a comprehensive approach including a thorough developmental history, direct observation and interaction with the child by an experienced clinician, as well as standardized assessment of cognitive, adaptive, and language functioning (Filipek et al., 2000; Johnson, Myers, & the Council on Children with Disabilities, 2007; Klaiman, Fernandez-Carriba, Hall, & Saulnier, 2014).

Although a majority of diagnostic evaluations for ASD, approximately 70%, rely upon clinical judgement without the use of any ASD-specific diagnostic measures (Wiggins, Baio, & Rice, 2006), practice parameters recommend the use of standardized assessments based on direct observation of and interaction with a child. The current gold-standard direct observation measure, the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012), as well as the most commonly used observation-based diagnostic measure, the Child Autism Rating Scale (CARS; Schopler, Van Bourgondien, Wellman, & Love, 2010), rely on trained clinician judgement of observed ASD-related behaviors. Scoring is based on clinician judgement of the atypicality of behaviors relative to expectations based on the child's developmental level.

Parent report also plays a critical role in both screening and diagnostic assessment. The American Academy of Pediatrics recommends routine developmental surveillance and screening for ASD at all 18- and 24-month well-child visits with referrals for diagnostic evaluation for those who screen positive (Johnson et al., 2007). Commonly-used and empirically-supported screening measures, such as the ASD-specific Modified Checklist for Autism in Toddlers, Revised with Follow-up (M-CHAT-R/F; Robins, Fein, & Barton, 1999) and the broadband Ages and Stages Questionnaire (ASQ; Squires & Bricker, 2009), are questionnaires based on parent report. Although providers are advised to tailor questions during exam to best elicit potential developmental concerns and to utilize clinical judgment in assessing ASD-related concerns, such as limited verbal or nonverbal communication (Filipek et al., 2000), parent report is a crucial factor in determining whether a referral for diagnostic evaluation is warranted.

As part of the diagnostic evaluation, parent report is also essential to gather comprehensive developmental history and assess adaptive functioning. Gathering a thorough developmental medical history, beginning from prenatal and perinatal time periods and continuing through the present, is an important component to assess for early patterns of behavior that may be related to ASD and to assess for possible medical etiology or comorbidities (Filipek et al., 2000). The current gold-standard measure for gathering comprehensive developmental history of ASD symptomatology is the Autism Diagnostic Interview-Revised (ADI-R; Rutter, LeCouter, & Lord, 2003), a semi-structured parent interview administered by a trained clinician. Standardized scoring is based on the interviewer's trained judgement of parent report. However, given the

administration time (1.5 to 3 hours) and training required for standardized scoring, fewer than 0.1% of diagnostic evaluations utilize the ADI-R (Wiggins et al., 2006), instead utilizing clinical interview without a normed measure to gather parent report (Johnson et al., 2007).

In assessment of adaptive functioning, parents and caregivers have the requisite experience over time to describe a child's functional behavior in daily contexts and routines (Tassé et al., 2012). Commonly-used adaptive functioning measures either rely on direct parent report on questionnaires, such as the Adaptive Behavior Assessment System (ABAS; Harrison & Oakland, 2015), or an interviewer's trained judgment of parent report during semi-structured interview, such as the Vineland Adaptive Behavior Scales (Vineland; Sparrow & Cicchetti, 2005). Assessment of adaptive skills is a recommended component not only to assess the level of impairment in current daily functioning as part of differential diagnosis, but also to inform individually-tailored treatment recommendations.

Across these screening and diagnostic measures based on either clinician rating or parent report, no single measure necessarily validates or invalidates a diagnosis of ASD in best-practice parameters. Instead, these multiple assessment sources and methods are recommended to ensure that several lines of evidence converge to determine whether a diagnosis of ASD is warranted.

Long-term outcomes in ASD. Although there are a number of empirically-supported treatments for ASD and a small percentage of children with ASD go on to not meet diagnostic criteria later in life (Dawson & Bernier, 2013; Fein et al., 2013), the vast majority of individuals with ASD continue to require support into older adolescence and

adulthood regardless of whether they have a comorbid Intellectual Disability (Gerhardt, Zawacki, & Satriale, 2008; Howlin, Moss, Savage, & Rutter, 2013). Consequently, the economic cost of ASD is high, estimated at $1.4 million per individual with ASD without cognitive impairment and $2.4 million per individual with ASD and Intellectual Disability (Buescher, Cidav, Knapp, & Mandell, 2014). Given that the overall prevalence of ASD has increased over time, particularly since the 1990s, and currently is estimated at 1 in 68 children (Christensen et al., 2016), the potential economic burden of ASD is growing. On an individual level, ASD is associated with lower quality of life and higher rates of anxiety, depression, and suicidality into adulthood (Croen et al., 2015; Gotham, Brunwasser, & Lord, 2015; Howlin et al., 2013).

Notably, long-term adult outcomes are generally poorer for females with ASD, including in employment attainment and in severity of anxiety and depression symptoms (Gotham et al., 2015; Howlin et al., 2013). There is emerging evidence that potential differences in the clinical phenotype of ASD across sexes may influence the effectiveness of vocational rehabilitation services and employment outcomes for adults with ASD (Wang, Sung, Sa, Leahy, & Sa, 2015). In addition, females are, on average, diagnosed with ASD later than males (Begeer et al., 2013; Shattuck et al., 2009). Because earlier diagnosis provides earlier access to intervention, which is in turn associated with more positive outcomes (Lord, 1995), females with ASD also begin at a disadvantage relative to their male peers. Clarifying similarities and differences in the clinical phenotype of social disability in ASD across sexes can help to evaluate how current diagnostic practices are identifying females with ASD and promote more individualized interventions.

Sex differences in the clinical phenotype of ASD

Extant studies directly examining similarities and differences in the clinical phenotype of ASD across sexes in individuals with an identified diagnosis have produced equivocal results. As a result, it has been difficult to draw conclusions with translational significance about the manifestation of social disability in males and females.

Cognitive functioning. The primary, most consistently identified difference across males and females in the clinical phenotype of ASD is in level of cognitive functioning. Female children and adults with ASD have, on average, lower cognitive abilities and are more likely to have a comorbid intellectual disability (e.g., Fombonne, 2009). Beginning with early epidemiological studies of ASD in the 1960s, females with ASD were more likely to have an IQ score below 55 compared to male peers (Lotter, 1966). Additional studies over time have replicated the result across different IQ cut-off scores. Females with ASD were 8.8 times more likely to exhibit cognitive functioning in the intellectual disability range (IQ below 70) compared to males with ASD (Volkmar, Szatmari, & Sparrow, 1993). Females with ASD were also more likely to have IQ scores below 80 (Frazier et al., 2014), suggesting borderline to impaired cognitive functioning; below 50 (Banach et al., 2009; Tsai & Beisler, 1983), in the moderate intellectual disability range; and below 34 (Lord & Schopler, 1985), in the severe intellectual disability range. Mean differences in IQ across sexes have been attributed to the larger proportion of females than males with cognitive functioning in the borderline to impaired range (Frazier et al., 2014). When restricting samples to children and adults with higher cognitive functioning and without intellectual disability, multiple studies have found that

males and females with ASD did not differ in verbal or nonverbal cognitive functioning (Koyama, Kamio, Inada, & Kurita, 2009; Lai et al., 2012; Mandy et al., 2012).

As a result of the overall difference across sexes in cognitive functioning, the male-to-female ratio in ASD diagnosis is significantly moderated by level of cognitive functioning. A previous review of epidemiological studies found a median ratio of 5.5 affected males for every 1 affected female among individuals with intact cognitive functioning compared to 1.95 affected males for every 1 affected male among individuals with ASD and comorbid moderate to severe intellectual disability (Fombonne, 2003). The most recent epidemiological estimates in the United States across the Autism and Developmental Disabilities Monitoring Network have similarly found a greater discrepancy in prevalence between males and females with higher cognitive functioning relative to those with lower cognitive functioning (Christensen et al., 2016). The estimated prevalence ratio is 5.1 affected males for every 1 affected female among children with ASD without intellectual disability and 3.7 to 1 among children with both ASD and comorbid intellectual disability.

Given the consistency of findings on sex differences in cognitive functioning in ASD as well as the relationship between cognitive functioning and ASD symptomatology in several past studies (e.g., Frazier et al., 2013; Havdahl et al., 2016), it is important to consider level of cognitive functioning when evaluating profiles of social disability across sexes.

Social communication and interaction. Results of past studies examining sex differences in ASD symptomatology within the domain of social communication and social interaction have been inconsistent. Studies have alternately found greater symptom

severity in females (e.g., Frazier et al., 2014), lower symptom severity in females (e.g., Holtmann, Bölte, & Poustka, 2007), or no differences between males and females with ASD (e.g., Mandy et al., 2012). These inconsistent sex difference findings, representing the three possible outcomes in between-group analysis, have been noted across methods of social disability assessment.

On parent-report measures in the large Simons Simplex Collection sample, females were rated with more severe ASD symptomatology and more severe social interaction impairment (Frazier et al., 2014). In smaller studies using both parent-interview (e.g., ADI-R) and parent-report measures, females with ASD were reported to present as more socially immature, have more difficulty engaging in reciprocal play, and have more difficulty maintaining age-appropriate friendships compared to toddler, school-age, and adult male peers (Carter et al., 2007; Holtmann et al., 2007; Kopp & Gillberg, 2011; McLennan, Lord, & Schopler, 1993). In contrast, retrospective parent report suggested that females exhibited less social impairment than males with ASD at 4 to 5 years old (McLennan et al., 1993). School-age and adult females with ASD were also reported to exhibit relative strengths compared to males in specific social communication behaviors, such as increased frequency of showing, directing others' attention, and utilizing an appropriate range of facial expressions (Holtmann et al., 2007).

On clinician-rating measures based on clinical judgment of direct observation (e.g., ADOS), females with ASD exhibited significantly more severe symptomatology in the domain of social communication in the toddler to school-age years (Carter et al., 2007; Frazier et al., 2014; Hartley & Sikora, 2009; Tsai & Beisler, 1983). Similarly, in analysis of symptomatology profiles using DSM-5 diagnostic criteria for ASD, school-

age females with ASD were more likely than male peers to exhibit diagnostic impairments in integrating verbal and nonverbal behaviors, initiating interactions, and maintaining reciprocal conversations (Hiller, Young, & Weber, 2014). In adulthood though, females were rated to exhibit less severe symptomatology compared to male peers (Lai et al., 2011). Notably, 79.3% of adult females with an existing ASD diagnosis received scores on the ADOS in the non-ASD range, a significantly higher proportion than in adult males with ASD.

At the same time, several other studies using the same parent-report and parent-interview measures (e.g., Mandy et al., 2012; Szatmari et al., 2012) and the same clinician-rating measures (e.g., Carter et al., 2007; Mandy et al., 2012; Reinhardt, Wetherby, Schatschneider, & Lord, 2014) in different samples have found no differences in social disability across males and females with ASD. The inconsistent results are not based only on relatively small samples of females with ASD included in many studies, as similar discrepancies in sex difference results have been identified across large, multisite national datasets (e.g., Simons Simplex Collection, Autism Treatment Network; Howe et al., 2015), particularly among participants with higher language functioning.

When controlling for identified sex differences in cognitive functioning though, most studies have reported that males and females with ASD exhibit similar levels of social impairment across age groups and across parent-report and clinician-rating measures (Holtmann et al., 2007; Pilowsky, Yirmiya, Shulman, & Dover, 1998; Tsai & Beisler, 1983; Volkmar et al., 1993). In the Simons Simplex Complex dataset, females' lower cognitive functioning mediated more severe social ASD symptomatology and accounted for measured differences in level of social disability (Thomas W. Frazier et al.,

2014). Accounting for level of cognitive and language functioning also led to increased agreement in sex difference results across the Autism Genetics Resource Exchange, Autism Treatment Network, and Simons Simplex Collection datasets (Howe et al., 2015). Consistent with these findings, recent meta-analysis showed that males and females with ASD exhibited similar levels of social impairment across ages (Van Wijngaarden-Cremers et al., 2014).

Restricted, repetitive behaviors and interests. In contrast to the equivocal findings on sex differences in the social communication and interaction domain, findings on sex differences in the restricted, repetitive behaviors and interests (RRBI) domain have been relatively consistent. Few studies have reported more severe RRBI in females with ASD; Tsai and Beisler (1983) found that females with ASD exhibited more atypical, repetitive motor movements compared to males, but this difference was no longer significant when controlling for level cognitive functioning. Instead, most studies assessing RRBI by parent report or clinician rating across age groups have found either no difference across sexes (Carter et al., 2007; Holtmann et al., 2007; Mandy et al., 2012; McLennan et al., 1993; Reinhardt et al., 2014) or lower RRBI in females (Bölte, Duketis, Poustka, & Holtmann, 2011; Hartley & Sikora, 2009; Hiller et al., 2014; Mandy et al., 2012). Results have generally been consistent across mode of measurement. As a rare exception, in one clinic sample of school-age children with ASD, parent ratings of RRBI did not differ across sexes, but clinician-rated RRBI was significantly lower in females (Mandy et al., 2012). Recent meta-analysis found that females with ASD exhibited lower levels of RRBI compared to males, particularly in school-age children, adolescents, and adults (Van Wijngaarden-Cremers et al., 2014). Females with ASD exhibited fewer

unusual sensory interests, fewer stereotyped mannerisms, and less routinized play (Lord, Schopler, & Revicki, 1982; Nicholas et al., 2008; Szatmari et al., 2012).

Measurement of social disability

One factor that may contribute to the equivocal results on sex differences in social disability, particularly in the social communication and interaction domain, is the fact that extant studies have primarily relied upon between-group analysis of summary scores from standardized measures. The most frequently used measures are the ADI-R (Rutter, LeCouter, et al., 2003), the current gold-standard diagnostic parent-interview measure, and the ADOS (Lord et al., 2012), the current gold-standard diagnostic clinician-rating measure based on direct observation, which have each been used in approximately 20% of studies on sex differences in ASD (Van Wijngaarden-Cremers et al., 2014). Other key instruments used include the parent-rating measure SRS (Constantino & Gruber, 2005) and the clinician-rating measure CARS (Schopler et al., 2010). All of these measures are also commonly used in clinical practice for diagnostic screening and assessment.

The SRS developed sex-specific norms because the distribution of ASD symptomatology in the general population skewed towards lower, non-pathologic scores in females compared to males (Constantino, 2011), a recently replicated finding (Frazier, Ratliff, et al., 2013a). In contrast, none of the other parent-report or clinician-rating measures of social disability directly addressed measurement equivalence across sexes during development. Reflecting the male-biased skew in diagnostic rates, it is estimated that the standardization samples for instruments measuring social disability were disproportionately male with a ratio of approximately 3 affected males for every 1 affected female (Koenig & Tsatsanis, 2005), approximately matching the average

population-level prevalence ratio of 4 affected males for every 1 affected female (Fombonne, 2009). As a result of this male bias in standardization, assessment criteria of most diagnostic instruments assessing social disability are based predominantly on the social behavior of males. If the underlying measurement structure differs across sexes, between-group mean comparisons may not accurately reflect similarities and differences in symptoms.

Several recent studies have directly addressed measurement equivalence across sexes in ASD at both the item and summary score level on the SRS (Frazier & Hardan, 2016; Frazier, Ratliff, et al., 2013b), ADI-R (Duku et al., 2013; Frazier & Hardan, 2016; Frazier et al., 2014), and ADOS (Frazier et al., 2014). In all cases, results supported measurement equivalence across sexes. Across modes of social disability measurement, scores for males and females with ASD reflected similar factor structures that mapped onto the DSM-5 symptom domains, providing evidence that any identified group mean differences are not simply attributable to measurement bias.

Although past studies assessing measurement equivalence have shown that males' and females' performance on instruments assessing social disability map onto the underlying factor structure of ASD symptomatology in the same way, less is known about how that latent construct of social disability compares across mode of measurement. Measurement equivalence analyses focus on single measures. Nonetheless, clinical practice parameters recommend the use of multiple standardized assessments in diagnostic assessment for ASD (Filipek et al., 2000; Johnson et al., 2007). These multiple measures must converge as multiple lines of evidence to support a diagnosis of ASD.

Information source has been cited repeatedly as a factor impacting measurement of social disability (e.g., Lai et al., 2016; Pilowsky et al., 1998), since different sources, such as parents and clinicians, each with their own experience and biases, have differential lengths of behavioral observation available when rating the child. Evaluation of concurrent validity across information source generally is not included in the standardization process (e.g., Lord et al., 2012). Extant studies have suggested low correlations between parent-report and clinician-rated measures of social disability in ASD (Hus, Bishop, Gotham, Huerta, & Lord, 2013; Reszka, Boyd, McBee, Hume, & Odom, 2013). There is also evidence that females with ASD exhibit increased discrepancy across methods of measuring social disability, potentially reflecting a difference in the latent construct. In two studies, while all participants met diagnostic criteria for ASD based on experienced clinical judgment, several females failed to meet cut-off criteria for a diagnosis based on either parent report or clinician rating of social disability (Lai et al., 2011; Pilowsky et al., 1998). Potential differences in the measurement of social disability across information source may contribute not only to the difference in diagnostic rate, but also to the disparity in age of diagnosis between males and females (Begeer et al., 2013; Shattuck et al., 2009).

Current study

The overarching goal of the current study was to clarify sex differences in the manifestation of social disability in school-age children with ASD by leveraging multiple levels of behavioral metrics. Extant studies comparing social behavior in males and females with ASD have largely found limited sex differences on parent-report and clinician-rating measures (Van Wijngaarden-Cremers et al., 2014). In addition to parent-

report and clinician-rating measures of social disability, we evaluated quantitative analyses of the moment-by-moment deployment of visual attention as a direct, performance-based measure of social behavior. Assessing how social disability manifested across information source was not meant to suggest a ranking of utility, but instead was intended to delineate factors influencing sex differences in ASD that may help in addressing existing clinical disparities in diagnostic prevalence and age of diagnosis.

Categorical sex differences in social disability. We sought to examine categorical between-group differences in the clinical phenotype of social disability in ASD across sexes using multiple modes of measurement, including parent report of social impairment in daily contexts (e.g., SRS-2), clinician rating of social impairment from direct observation within a semi-structured context (e.g., ADOS-2), and performance-based assessment of social visual engagement during an unstructured task. We hypothesized that males and females with ASD would not differ on standardized parent-report or clinician-rating measures of social disability, consistent with past sex differences research. We hypothesized that performance-based measures of social visual engagement would have greater sensitivity to potential sex differences. Collected using eye-tracking technology during unstructured, naturalistic viewing of social scenes, the performance-based measures allowed for direct analysis of social behavior. In addition to summary measures of the distribution of visual fixations on specific regions-of-interest (e.g., eyes), which have generally revealed significant group differences between children with ASD and typically-developing peers (e.g., Jones, Carr, & Klin, 2008; Klin, Jones, Schultz, Volkmar, & Cohen, 2002), we utilized more temporally- and spatially-sensitive

quantification of social visual engagement, derived as the deviation from normative behavior (Jones & Klin, n.d.), to assess how children engaged with and responded to dynamically changing social environments.

Rather than focusing on only group mean differences across sexes, we also defined and compared the latent construct of social disability across modes of measurement. We hypothesized that the latent factor structure of social disability would differ across sexes, reflecting emerging evidence that measures of social disability have lower concurrent validity for females compared to males with ASD.

Dimensional association with adaptive social ability across sexes. As a secondary aim, we sought to identify factors associated with adaptive social ability in daily contexts in order to further elucidate potential sex differences across modes of social disability measurement in ASD. Counterintuitively, social ability and social disability are not consistently related across methods of assessment; in other words, decreased symptomatology does not necessarily translate to greater social competency in ASD. Parent report of social disability is negatively correlated with social ability (Bölte, Poustka, & Constantino, 2008), but there is no significant relationship between clinician rating of social disability and adaptive social ability (Kanne et al., 2011; Klin et al., 2007; Saulnier & Klin, 2007). Nevertheless, long-term follow-up studies have found that adaptive functioning, rather than cognitive profile or level of social disability, is the strongest predictor of positive adult outcomes (Farley et al., 2009). Particularly given existing evidence that adult outcomes are poorer for females relative to males with ASD (Howlin et al., 2013), it is consequently important for the advancement of individualized interventions to identify factors influencing adaptive ability.

Similar to past research on social disability, studies examining the dimensional association between social disability and social ability have used predominantly male samples (Bölte et al., 2008; Kanne et al., 2011) or have excluded females altogether (Klin et al., 2007; Saulnier & Klin, 2007). No published studies have addressed the effect of sex. Potential differences in the relationship between social disability and social ability in daily environments with peers may be an additional factor affecting the manifestation and assessment of social behaviors across sexes in ASD. We hypothesized that sex would moderate the correlation between social disability and adaptive social ability across parent-report, clinician-rating, and performance-based measures.

Method

Participants

A total of 259 children participated, all with the written informed consent of their parents and/or legal guardians. The research protocol was approved as non-significant risk by the Human Investigations Committee at Yale University School of Medicine and the Institutional Review Board at Emory University School of Medicine. Recruitment and data collection occurred through the Autism Program of the Yale Child Study Center in New Haven, CT (n = 182) and through the Marcus Autism Center in Atlanta, GA (n = 77). Recruitment for participants with ASD was conducted through clinic referrals. Recruitment for both participants with ASD and typically-developing participants was also conducted through advertisements at local ASD fundraisers, resource fairs, children's groups, and schools. Families were free to withdraw from the study at any time.

The 259 participants included 200 children with ASD (136 male, 64 female) and 59 typically-developing children (27 male, 32 female). Table 1 provides participant characterization data and statistical comparisons. The sample represented a wide range of ages (±1 SD, ~7-13 years) and level of cognitive functioning (±1 SD, Full Scale IQ of ~72-119). The ASD group also represented a wide range of symptomatology severity (±1 SD, ADOS-2 Calibrated Severity Score of ~5 to 9). Participants were matched across diagnostic group and sex based on age ($F_{3,255} = 0.89$, $p = 0.44$). The same sample was used for Studies 1 and 2.

Participant characterization

Cognitive functioning assessment. Both participants with ASD and typically-developing participants were administered a cognitive functioning assessment as part of the clinical evaluation. A variety of cognitive functioning measures were used over the time frame of the study, reflecting changes in clinic practice over time rather than individualized measure selection. Cognitive functioning assessment measures used included: the Differential Ability Scales, Second Edition (DAS-II; Elliott, 2006; n = 157); the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999; n = 51); the Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991; n = 19); the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler, 2004; n = 29); and the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Wechsler, 2002; n = 1). The DAS-II and Wechsler measures have been shown to have sufficiently high concurrent validity (Elliott, 2006). In addition, the Mullen Scales of Early Learning (Mullen; Mullen, 1995; n = 2) was administered to derive ratio IQ scores as an estimate of cognitive functioning if a participant's

developmental level fell below 2 years, 6 months. There was no difference in the distribution of cognitive functioning measures used across sexes ($X^2 = 8.402, p = 0.14$).

Social disability assessment. Participants with ASD were administered the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al., 2012), and parents of children with ASD were interviewed using the Autism Diagnostic Interview-Revised (ADI-R; Rutter, LeCouter, et al., 2003). Parents of both participants with ASD and typically-developing participants completed the Social Communication Questionnaire, Lifetime Form (SCQ; Rutter, Bailey, et al., 2003) and Social Responsiveness Scale, Second Edition (SRS-2; Constantino, 2012). Results of the ADOS-2 and SRS were used in analyses as measures of social disability in participants with ASD. The ADI-R and SCQ were administered as part of the clinical evaluation to aid in diagnostic clarification.

The ADOS-2 is a semi-structured, clinician-based assessment of social communication behaviors through direct observation. The child is placed in naturalistic social situations that probe for specific social-communicative responses. Though the protocol follows a standard administration, the situations themselves are unstructured to allow for a sample of the child's naturalistic social behavior. The ADOS-2 includes 4 modules that are designated based on language ability. The appropriate module was determined based on clinical judgment. The child's behaviors in the domains of Social Affect and Restricted and Repetitive Behavior are scored based on frequency and the degree of abnormality or severity. The diagnostic algorithm of the ADOS-2, which is based on the combined domain totals, has good discriminative validity in identifying clinically-diagnosed children with ASD (Lord et al., 2012). All clinicians were trained

and reliable (i.e., >80% exact agreement with the measure authors or certified trainers across all item and domain scores) on the administration and scoring of all modules of the ADOS-2. Although the ADOS-2 was only recently published, the revised scoring algorithms for Modules 1 to 3 have been used with permission in research for several years and the revised scoring algorithm for Module 4 was published separately (Hus & Lord, 2014), so all participants in the sample had scores based on the ADOS-2 algorithm. Given the wide range of ADOS-2 modules used in this study's large sample (Modules 1 through 4), we also calculated the Calibrated Severity Score for each participant, a measure standardized across modules with higher scores denoting more severe ASD symptomatology (Gotham, Pickles, & Lord, 2009; Hus & Lord, 2014).

The SRS-2 is a parent-report questionnaire that provides a quantitative, continuous measure of severity of social impairment. The total score across all items, presented as a Likert-type scale based on behavior frequency in the past 6 months, is converted to a T-score based on sex-specific norms, with higher scores indicating more severe ASD symptomatology. The instrument has high convergent validity with the ADI-R and discriminant validity in differentiating ASD from other child psychiatric disorders (Constantino & Gruber, 2005). Based on the evaluation date, parents completed either the first edition of the SRS or the more recently published second edition. Because both editions include the same items, all participants were scored based on the updated SRS-2 norms. The manual recommends use of raw total scores in research (Constantino & Gruber, 2005), though T-scores based on sex-specific norms have also used in recent studies (e.g., Constantino, 2011; Frazier et al., 2013). Both raw total scores and T-scores were used as parent-report measures of social disability in ASD.

The SCQ is a parent-report screening questionnaire assessing social development and ASD symptomatology. The total score across all items is assessed relative to a cut-off score for likely ASD diagnosis. The instrument has high convergent validity with the ADI-R and high discriminant ability in differentiating clinically-diagnosed individuals with ASD from those without ASD, including individuals with intellectual disability (Rutter, Bailey, et al., 2003). The total score on the SCQ, in reference to the validated cut-off score, was used as a screening measure for ASD symptomatology and likelihood of ASD diagnosis for both participants with ASD and typically-developing participants.

The ADI-R is a comprehensive, semi-structured parent interview assessing most developmental and behavioral aspects of ASD. Using a parent or caregiver's report, a clinician scores items based on behavior severity, abnormality, and frequency. The diagnostic algorithm of the ADI-R, which utilizes scores from a subset of items, has good discriminative validity in identifying clinically-diagnosed children with ASD relative to children with non-ASD language impairment (Rutter, LeCouter, et al., 2003). Clinicians were trained and reliable (i.e., >80% exact agreement with the measure authors or certified trainers across all item and domain scores) in the administration and scoring of the ADI-R. The diagnostic algorithm of the ADI-R as well as the extensive clinical information collected during the interview was used in diagnostic classification.

Adaptive functioning assessment. Parents of both participants with ASD and typically-developing participants completed the Vineland Adaptive Behavior Scales to assess the frequency and independence of adaptive behaviors across the domains of Communication, Daily Living, Socialization, and Motor Skills. Parents of a majority of participants completed the Vineland Adaptive Behavior Scales, Second Edition, Survey

Interview Form (Vineland-II; Sparrow & Cicchetti, 2005; n = 169), a semi-structured parent interview. The Vineland-II was administered by clinicians and research coordinators who were trained and reliable (i.e., >90% exact agreement across all items with senior clinicians) in the administration and scoring of the assessment. Based on evaluation date, reflecting changing clinic practices over time, parents alternately completed the Vineland Adaptive Behavior Scales, Interview Edition, Expanded Form (Sparrow, Balla, & Cicchetti, 1984; n = 54), or Vineland Adaptive Behavior Scales, Second Edition, Parent-Caregiver Rating Form (Sparrow & Cicchetti, 2005; n = 7).

Inclusion and exclusion criteria. Two experienced clinicians independently assigned an overall clinical diagnosis or classification for each participant upon review of all available data. To qualify for inclusion in the ASD group, children met the following three criteria: (1) they met criteria for ASD on the ADOS-2 (Catherine Lord et al., 2012); (2) they met criteria for ASD on the ADI-R (Rutter, LeCouter, et al., 2003); and (3) they received a consensus clinical diagnosis of ASD by two experienced clinicians upon independent review of all available clinical data, including standardized testing and video of the diagnostic examination. In the case of disagreement between ADOS-2 and ADI-R criteria, the consensus clinician-assigned, best-estimate diagnosis took priority as a widely accepted gold standard (Lord & Corsello, 2005). The rate of concordance between experienced clinicians is high (Klin, Lang, Cicchetti, & Volkmar, 2000), but in the rare case of disagreement in diagnostic assignment between clinicians, a case conference was called to attain a better understanding of the case and reach clinical consensus among at least 2 of the clinicians involved. Diagnostic guidelines followed either DSM-IV-TR or DSM-5 criteria based on the evaluation date (American Psychiatric Association, 2004,

2013). All children who were diagnosed under DSM-IV-TR criteria would also meet criteria for ASD per current, DSM-5 criteria.

For inclusion in the typically-developing group, children met the following inclusionary criteria: (1) they exhibited no cognitive impairments either currently or by history (i.e., all IQ scores > 70); (2) they exhibited no social disability either currently or by history, assessed based on a score below the cut-off that did not meet criteria for ASD on the SCQ (Rutter, Bailey, et al., 2003); (3) they had no family history of ASD or major neurological or psychiatric illness in first- or second-degree relatives based on parent report; and (4) they received a consensus clinical classification as typically-developing.

All children had normal or corrected-to-normal vision and no history of major visual or auditory impairments. Children were not eligible to participate if they had significant auditory impairments or significant visual acuity deficits that were not corrected with corrective lenses. Participants were also excluded in case of abnormal functional eye movements ascertained through tests of visuomotor function (Klin et al., 2002; Leigh & Zee, 2006).

Eye-tracking procedures

Stimuli. Children's visual scanning and fixations were measured using eye-tracking technology while they viewed developmentally-appropriate, naturalistic social scenes selected to reflect typical childhood experiences in everyday settings (see Figure 1 for example stills). The film scenes, presented as full-screen audiovisual stimuli, formed 2 self-contained 5- to 7-minute scenarios of nuanced social interaction. The first clip offered a narrative of a girl trying to fit in and make friends with classmates at school. The clip was selectively edited from the 1995 film *Welcome to the Dollhouse* to contain

only themes, language, and content appropriate for children. The second clip, edited from the 1993 film *The Sandlot*, showed a group of boys playing baseball together.

The film scenes were shown as full-screen audiovisual stimuli on a 20-inch (50.8-cm) computer monitor with a refresh rate of 60 Hz noninterlaced. Video frames were 8-bit color images 640 x 480 pixels in resolution. Video frame rate of presentation was 30 frames per second. The audio track was a single (mono) channel sampled at 44.1 kHz.

Experimental Procedure. Eye-tracking data were concurrently collected during the experimental protocol using a video-based, dark pupil-corneal reflection system with hardware and software created by ISCAN, Inc. (Woburn, MA). The system was mounted unobtrusively on the bill of a baseball hat and used a target-tracking method that enables highly accurate eye-tracking without needing to restrain the participant's head. The equipment is accurate to within ±0.3° across a ±20° horizontal and vertical field of view. Participants sat in a comfortable armchair placed 25 inches from a 20-inch computer screen mounted flush within a black wooden panel. Data were collected at a rate of 60 samples/second and recorded to video at the standard rate of 30 frames/second.

Children were individually calibrated using a 5-point system prior to presentation of experimental videos. To ensure accurate eye-tracking data, the calibration was checked regularly between videos through the duration of testing. If the recorded point of regard shifted more than 3°, data collection was paused and the 5-point calibration procedure repeated.

All aspects of the experimental protocol were performed by personnel blinded to the diagnostic status of the children. All diagnostic measures were administered by trained clinicians blinded to the results of experimental procedures.

Data processing. Most aspects of data acquisition and all aspects of coding and data processing were automated to ensure separation between clinical characterization and the experimental protocol. Analysis of eye movements and coding of fixation data were performed with in-house software. Non-fixation data, comprising saccades, blinks, and off-screen fixations, were automatically identified in the first phase of analysis. Saccades were identified based on eye movement velocity, using a threshold of 30° per second. Blinks were identified as in Shultz, Klin, & Jones, 2011. The blink detection algorithm was previously verified through visual analysis of video images and through simultaneous eye-tracking and EMG recording. Off-screen fixations, when a child looked away from the video screen, were identified by fixation coordinates beyond the possible screen bounds.

Across 752.3 seconds of total viewing data (22,570 video frames), measures of fixation time (as percentage of total time spent fixating on the stimuli presentation monitor) significantly differed across diagnostic group (typically-developing, 69.2% (7.5); ASD, 57.3% (13.5); data given as mean (standard deviation); $F_{1,255} = 38.860, p < 0.001$). However, neither the main effect of sex ($F_{1,255} = 0.03, p = 0.87$) nor the diagnosis X sex interaction were significant ($F_{1,255} = 0.07, p = 0.79$). Measures of fixation time were not significantly different across sexes within the typically-developing group (male, 69.3% (6.1); female, 69.1% (8.6); $t_{57} = 0.82, p = 0.42$) or the ASD group (male, 57.1% (12.7); female, 57.9% (15.1); $t_{198} = 0.70, p = 0.49$). Measures of non-fixation data were also not significantly different across sexes within each group for saccades (typically-developing: male, 20.5% (3.2); female, 19.3% (4.1); $t_{57} = 1.19, p = 0.24$; ASD: male, 21.7% (4.7); female, 20.6% (4.2); $t_{198} = 1.56, p = 0.12$), blinks (typically-developing:

male, 7.0% (5.2); female, 5.8% (5.3); $t_{57} = 0.82, p = 0.42$; ASD: male, 7.9% (5.9); female, 7.2% (6.9); $t_{198} = 0.70, p = 0.49$), and off-screen fixations (typically-developing: male, 3.3% (3.3); female, 5.8% (5.7); $t_{57} = -2.00, p = 0.05$; ASD: male, 13.4% (11.9); female, 14.3% (13.6); $t_{198} = -0.50, p = 0.62$).

Eye movements identified as fixations were coded relative to 4 regions-of-interest (ROIs) defined within all video stimuli: eyes, mouth, body (neck, shoulders, and contours around the eyes and mouth, including hair), and object (background setting and inanimate objects). ROIs were hand-traced for all video frames (22,570 frames) and were then stored as binary bitmaps via software written in MATLAB. Automated coding of fixation time to each ROI consisted of a numerical comparison of each child's coordinate fixation data against the bitmapped ROIs. Percentage of fixation time on each ROI was calculated relative to an individual's total fixation time.

Study 1: Categorical sex differences in social disability

We examined categorical, between-group differences across sexes in ASD based on multiple level of measurement of social disability. Parent report of social disability was assessed using total scores from the SRS-2. Clinician rating of social disability was assessed using the Social Affect domain score of the ADOS-2. Performance-based measurement of social disability was derived from eye-tracking analyses quantifying social visual engagement. Categorical differences across males and females with ASD were assessed by independent samples *t*-tests. For measures with data available for both the typically-developing and ASD groups, analyses were conducted by two-way factorial ANOVA with diagnosis and sex as fixed factors followed by planned contrasts within diagnostic group.

Quantifying social visual engagement. In addition to evaluating the distribution of visual fixations to each ROI across clips, we also calculated more temporally- and spatially-sensitive measures of social visual engagement. Analyses were conducted in MATLAB and utilized moment-by-moment coordinate fixation data to yield a continuous measure of normative time-varying visual scanning in typically-developing participants and deviation therefrom in participants with ASD. The computational method is described in greater detail in Jones & Klin (n.d.).

We first evaluated time-varying visual scanning patterns for typically-developing male and female participants separately to create a continuous measure of normative, sex-specific visual engagement. Calculations applied kernel density analysis to moment-by-moment coordinate fixation data throughout the duration of clips to quantify the visual salience of all areas of the onscreen image in each frame (Silverman, 1986). To test whether participants' visual scanning significantly converged on a common location more than expected by chance alone in a given frame, data were compared against 100,000 iterations of simulated kernel density estimates from randomized fixation data.

We next obtained a measure of each participant's degree of relative convergence with normative visual scanning patterns by leave-one-out resampling. Each typically-developing participants' coordinate fixation data in each frame was assessed relative to the visual salience of all areas of the onscreen image based on the rest of the typically-developing group's data. Each ASD participants' data was assessed relative to the full same-sex typically-developing group's data. We identified moments of significant deviation from normative visual scanning patterns by comparing group means in each frame, with permutation testing used to correct for multiple comparisons. To test for

between-group differences in the frequency of significant deviation from normative convergence, we calculated bootstrapped 95% confidence intervals across 1,000 permutations randomly resampled with replacement from the original comparison group. Although the calculations statistically account for differences in sample size, as an added control for between-group differences, we used equal group sizes for resampled permutation testing.

As an individual summary measure of relative divergence from normative visual scanning patterns for dimensional analyses within the ASD group, we took the median of all individual scores for each child during moments of significant between-group divergence in visual scanning; higher scores indicated greater deviation from normative patterns. As an additional measure of individual relative divergence, we also calculated the frequency of fixations away from the ROI of highest normative visual salience during moments of significant between-group divergence; higher scores indicated greater deviation from normative patterns.

Evaluating the structure of social disability across sexes. We completed within-sex exploratory factor analyses (EFA) to identify the underlying factor structure of social disability in males and females with ASD across parent-report, clinician-rated, and performance-based measures of social disability. Analyses were conducted in SPSS v24 and MATLAB. Due to relatively low correlations between measures, the EFA was carried out with principal axis factoring, which has sensitivity to potentially low factor loadings (de Winter & Dodou, 2012). Oblique rotation was chosen based on the assumption that dimensions may correlate with each other. To determine the number of factors to retain, we both visually inspected the scree plot and evaluated structure

outcomes using an eigenvalue cut-off of 1. We then calculated Tucker's congruence coefficient following Procrustes rotation to assess the equivalence of factor loadings across sexes (Lorenzo-Seva & ten Berge, 2006; McCrae, Zonderman, Costa, Bond, & Paunonen, 1996).

Study 2: Dimensional association with adaptive social ability across sexes

We examined the moderating effect of sex on the dimensional association between measures of social disability and functional social ability in order to identify measures associated with greater social-communicative competency in males and females with ASD. Parent report of social disability was assessed using total scores from the SRS-2. Clinician rating of social disability was assessed using the Social Affect domain score of the ADOS-2. Performance-based measurement of social disability was derived from eye-tracking analyses quantifying social visual engagement, as described previously. Functional social ability was measured using the Socialization domain of the Vineland-II. Associations were measured by Pearson correlation coefficient in each sex. To examine moderation, the measure of social disability was included along with sex and the social disability measure's interaction with sex as predictors in a stepwise linear regression model. The dependent variable was functional social ability. A significant interaction term would indicate a potential moderating effect of sex (Baron & Kenny, 1986).

Results

Study 1: Categorical sex differences in social disability

We first tested for between-group differences in cognitive functioning given prior studies identifying an association between level of cognitive functioning and level of social disability (e.g., Frazier et al., 2013; Havdahl et al., 2016). The ASD group and typically-developing group were matched on age but significantly differed in cognitive functioning; the ASD group performed significantly lower than typically-developing peers across measures of Full Scale IQ (typically-developing, 110.9 (16.6); ASD, 91.7 (23.5); $F_{1,255} = 31.33, p < 0.001$), Verbal IQ (typically-developing, 113.2 (17.1); ASD, 90.9 (25.5); $F_{1,255} = 36.94, p < 0.001$), and Nonverbal IQ (typically-developing, 107.2 (15.5); ASD, 93.8 (22.0); $F_{1,255} = 17.58, p < 0.001$). There was no main effect of sex (Full Scale IQ, $F_{1,255} = 8.07, p = 0.34$; Verbal IQ, $F_{1,255} = 1.02, p = 0.31$; Nonverbal IQ, $F_{1,255} = 0.58, p = 0.45$) or diagnosis-by-sex interaction across cognitive functioning measures (Full Scale IQ, $F_{1,255} = 0.39, p = 0.53$; Verbal IQ, $F_{1,255} = 0.58, p = 0.45$; Nonverbal IQ, $F_{1,255} = 0.44, p = 0.51$). Within both the typically-developing group and ASD group, males and females did not significantly across measures of cognitive functioning (Table 1). Within the ASD group, the proportion of participants with cognitive functioning in the intellectual disability range (Full Scale IQ < 70) did not differ across sexes (Male, 16.9%; Female, 17.2%; $X^2 = 0.002, p = 0.96$). We also tested for sex differences in the discrepancy between verbal and nonverbal IQ given past studies identifying an effect of cognitive profile on measures of social disability (Rice, Moriuchi, Jones, & Klin, 2012). Males and females with ASD did not differ in the proportion of participants with a discrepancy (>10 points) between Verbal IQ and Nonverbal IQ (Male, Verbal IQ advantage, 16.2%, Nonverbal IQ advantage, 30.9%; Female, Verbal IQ advantage, 25.0%, Nonverbal IQ advantage, 31.3%; $X^2_2 = 2.52, p = 0.28$). Across all comparisons

between males and females, there were no significant differences in cognitive functioning or cognitive profile, suggesting that any identified sex differences in social disability could not be attributed to cognitive functioning differences alone.

Parent-report and clinician-rated measures of social disability. We proceeded to assess sex differences across measures of social disability in ASD. On parent-report measures of social disability from the SRS-2, males and females with ASD did not differ in raw total score (Table 1; Cohen's $d = 0.05$), but females were rated as significantly more affected based on standardized scores using same-sex norms (Cohen's $d = 0.33$). On clinician ratings of social disability from the ADOS-2, males and females with ASD did not differ in social affect symptomatology (Table 1; Cohen's $d = 0.18$). There were also no differences across sexes in the Restricted and Repetitive Behaviors domain (Cohen's $d = 0.15$) or in overall Calibrated Severity Score (Cohen's $d = 0.18$).

Performance-based measures of social disability

Distribution of visual fixations. We compared the distribution of visual fixations on all ROIs (eyes, mouth, body, and object regions) across sexes and diagnostic groups as an initial performance-based measure of social disability (Figure 2). There was no sex-by-diagnosis interaction effect on fixations to eyes ($F_{1,255} = 2.37, p < 0.13, \eta_p^2 = 0.01$), mouth ($F_{1,255} = 0.60, p = 0.44, \eta_p^2 = 0.002$), or body regions ($F_{1,255} = 2.13, p = 0.15, \eta_p^2 = 0.01$). The sex-by-diagnosis interaction for fixations to object regions was marginally significant ($F_{1,255} = 3.95, p = 0.05, \eta_p^2 = 0.02$) and remained significant when including Full Scale IQ as a covariate based on the identified differences in cognitive functioning across diagnostic groups ($F_{1,254} = 4.11, p = 0.04, \eta_p^2 = 0.04$).

We then conducted planned contrasts to assess sex differences in visual fixations within diagnostic groups and across diagnostic groups, with Bonferroni correction for multiple comparisons ($\alpha = 0.0125$). Typically-developing males and females did not differ in distribution of visual fixations across ROIs (eyes, $t_{57} = 1.38$, $p = 0.17$, Cohen's $d = 0.36$; mouth, $t_{57} = 0.55$, $p = 0.58$, Cohen's $d = 0.15$; body, $t_{57} = -2.11$, $p = 0.04$, Cohen's $d = 0.56$; object, $t_{57} = -0.71$, $p = 0.48$, Cohen's $d = 0.18$). Males and females with ASD similarly did not differ in fixations to eyes ($t_{198} = -0.59$, $p = 0.56$, Cohen's $d = 0.09$), mouth ($t_{198} = -0.38$, $p = 0.71$, Cohen's $d = 0.06$), or body regions ($t_{198} = -0.003$, $p = 0.99$, Cohen's $d < 0.01$); however, females with ASD looked significantly less at object regions compared to males with ASD ($t_{198} = 2.79$, $p = 0.006$, Cohen's $d = 0.45$).

Across diagnostic groups, males with ASD focused significantly less on eyes ($t_{161} = 4.72$, $p < 0.001$, Cohen's $d = 1.03$) and significantly more on body ($t_{161} = -5.31$, $p < 0.001$, Cohen's $d = 1.34$) and object regions ($t_{161} = -3.99$, $p < 0.001$, Cohen's $d = 0.99$) compared to typically-developing males. The difference in mouth fixations did not reach significance based on the Bonferroni-corrected α ($t_{161} = 2.28$, $p = 0.02$, Cohen's $d = 0.42$). The largest difference was in attention to faces overall, representing eyes and mouth regions combined (typically-developing, 64.7% (7.9); ASD, 47.4% (14.7); $t_{161} = 5.93$, $p < 0.001$, Cohen's $d = 1.47$). In contrast, females with ASD did not significantly differ from typically-developing females in distribution of visual fixations across ROIs after correcting for multiple comparisons (eyes, $t_{94} = 2.26$, $p = 0.03$, Cohen's $d = 0.48$; mouth, $t_{94} = 0.93$, $p = 0.35$, Cohen's $d = 0.19$; body, $t_{94} = -2.40$, $p = 0.02$, Cohen's $d = 0.53$; object, $t_{94} = -1.57$, $p = 0.12$, Cohen's $d = 0.32$), though it is important to note that effect sizes were in the medium range for eyes and body fixations, with trends toward

differences across diagnostic group in the same direction as in males. When eyes and mouth regions were combined, females with ASD looked significantly less at faces compared to typically-developing females (typically-developing, 57.9% (14.2); ASD, 49.1% (13.7); $t_{94} = 2.94$, $p = 0.004$, Cohen's $d = 0.63$).

Deviation from normative visual scanning patterns. To calculate more temporally- and spatially-sensitive performance-based measures of social disability, we quantified the degree of deviation from normative visual scanning patterns. The visual scanning of typically-developing males was significantly convergent during 86.8% of total viewing time (~12.5 minutes). Similarly, the visual scanning of typically-developing females was significantly convergent during 89.4% of total viewing time. In total, the visual scanning patterns of typically-developing males and females significantly diverged in only 2.7% of frames (Figure 3a), comprising about 20 seconds of the total 12.5 minutes of viewing time. In over 97% of viewing time, typically-developing males and females were looking at approximately the same location at the same time.

Relative to the normative visual scanning patterns of typically-developing same-sex peers, the visual scanning of males with ASD was significantly divergent in 35.2% of total viewing time, whereas the visual scanning of females with ASD was significantly divergent in 7.1% of total viewing time. The difference in frequency of divergence across sexes was significant ($X^2 = 34.46$, $p < 0.001$). To further assess between-group differences and add an additional control for sample size, we compared bootstrapped 95% confidence intervals for the frequency of divergence from same-sex normative visual scanning patterns. The 95% confidence intervals for males and females with ASD did not overlap (Figure 3b; male, 11.7-19.3%; female, 3.6-9.0%); on average, the visual

scanning of males with ASD diverged from normative same-sex patterns over 2.5 times more frequently compared to the visual scanning of females with ASD.

Factor structure of social disability across modes of measurement. Table 2 shows correlations across multiple modes of social disability measurement, including parent report (SRS-2 Total score), clinician rating (ADOS-2 Social affect score), and performance-based assessment of divergence from normative visual scanning patterns (frequency of fixations away from the ROI of highest normative visual salience during moments of significant between-group divergence), for both males and females with ASD. The only significant correlation was between parent report and clinician rating in females with ASD (Table 2; $r = 0.29, p = 0.02$). All measures were entered into an exploratory factor analysis (EFA) separately for males and females with ASD. Although the Kaiser-Meyer-Olkin measure indicated adequate sampling adequacy in both groups (male, 0.51; female, 0.52; H. Kaiser, 1974), both the male and females groups failed Bartlett's test of sphericity due to low correlations across measures (male, $X^2_3 = 2.03, p = 0.57$; female, $X^2_3 = 5.95, p = 0.11$). Consequently, results of the EFA should be interpreted with caution.

The EFA indicated that a single-factor model should be should be retained for both males and females with ASD (see Figure 4 for factor loadings). Based on Tucker's congruence coefficient, factor equivalence across males and females with ASD was low and below the threshold for either "fair" or "good" equivalence ($\phi = 0.09$; Lorenzo-Seva & ten Berge, 2006), indicating that the latent structure across social disability measures differed between males and females with ASD. Analyses were repeated using alternate

performance-based measures of social disability, including fixation time on ROIs and relative divergence from normative visual salience, with no material change in results.

Study 2: Dimensional association with adaptive social ability across sexes

Males and females with ASD did not differ in adaptive social ability (Table 1), but both groups exhibited significant deficits in adaptive social functioning; their adaptive social functioning was approximately 5 years below their chronological age. Sex did not moderate the relationship between parent report of social disability ($F_{1,157} = 3.09$, $p = 0.08$). Parent report of social disability was significantly associated with adaptive social ability in both males (Figure 5a; $r = -0.43, p < 0.001$) and females with ASD (Figure 5d; $r = -0.55, p < 0.001$). In contrast, the relationship between clinician rating of social disability and adaptive social ability was significantly moderated by sex ($F_{1,189} = 5.09, p = 0.03$). The moderation effect remained significant when controlling for age ($F_{1,189} = 4.65, p = 0.03$), given the use of age equivalents as a measure of adaptive social ability. Clinician rating of social disability was more strongly related to adaptive social ability in females with ASD (Figure 5b; $r = -0.55, p < 0.001$) than in males (Figure 5e; $r = -0.20, p = 0.02$).

First using distribution of visual fixation on ROIs as a performance-based measures of social disability, sex did not moderate the relationship between fixation time on any ROI and adaptive social ability (eyes, $F_{1,193} = 0.38, p = 0.54$; mouth, $F_{1,189} = 1.86$, $p = 0.17$; body, $F_{1,189} = 0.94, p = 0.33$; object, $F_{1,189} = 2.44, p = 0.12$). There was a marginally significant moderating effect of sex on the association between visual fixation time on faces (i.e., combined eyes and mouth regions) and adaptive social ability ($F_{1,189} = 3.75, p = 0.05$) that remained trending when controlling for age ($F_{1,189} = 3.46, p = 0.06$).

Increased attention to faces was more strongly associated with adaptive social functioning in females with ASD (Figure 5c; $r = 0.39, p = 0.002$) than in males (Figure 5f; $r = 0.20, p = 0.02$).

There was no moderating effect of sex when using time-varying measures of divergence from normative visual scanning as a performance-based measure of social disability ($F_{1,193} = 0.40, p = 0.53$). Divergence from normative visual scanning was not associated with adaptive social ability in either males (Figure 5c; $r = -0.12, p = 0.16$) or females with ASD (Figure 5f; $r = -0.03, p = 0.81$).

Discussion

We examined the manifestation of social disability in male and female school-age children with ASD both within and across multiple methods of measurement: parent report of daily social behavior, clinician rating of observed social behavior in a semi-structured context, and performance-based measures of social behavior during an unstructured task. By including temporally- and spatially-sensitive measures of social visual engagement, we aimed to have greater sensitivity in assessing potential sex-based differences in social disability. In addition, by leveraging multiple levels of behavioral metrics, we assessed how both core and sex-specific features of the clinical phenotype of ASD manifest on a moment-by-moment basis and may be influenced by method of ascertainment. Clarifying similarities and differences in clinical phenotype across sexes is critical to determining sex-specific risk or protective factors and to advancing clinical care in both diagnosis and treatment of ASD.

Assessing categorical differences across modes of measurement, we found that males and females with ASD did not differ on parent report or clinician rating measures

of social behavior. However, on performance-based measures of social behavior, the visual scanning patterns of females with ASD were significantly less divergent from normative same-sex social visual engagement compared to the visual scanning patterns of males with ASD. In other words, females with ASD viewed social scenes in a way that was more similar to typically-developing peers than males with ASD. There were diminished differences in females across diagnosis on both summary visual fixation measures and time-varying social visual engagement measures. Despite their reduced divergence from normative social behavior on performance-based measures, females with ASD did not have a concurrent advantage in parent rating of social behavior, clinician rating of social behavior, or in adaptive social functioning. Predictors of adaptive social functioning also had limited difference across sexes

These results in parent-report and clinician-rating measures of social disability were generally consistent with past studies assessing sex differences. In school-age children with ASD matched on cognitive functioning, males and females with ASD demonstrate limited differences in ASD symptomatology within the social communication and interaction domain (Holtmann et al., 2007; Van Wijngaarden-Cremers et al., 2014; Volkmar et al., 1993). On the SRS-2, the only current social disability measure with sex-specific norms (Constantino, 2012), we found that raw total scores were similar across sexes but standardized scores based on sex-specific norms did significantly differ across sexes, with females with ASD more impaired. Recent studies in large samples of children with ASD have found a similar discrepancy across raw and standardized scores (Frazier, Ratliff, et al., 2013b; Hus et al., 2013). The sex-specific norms were developed based on general population differences in scores, with females

across the general population rated as exhibiting less social impairment than males; similar differences in raw totals across sex are generally not observed in children with ASD, resulting in higher standardized scores for females with ASD (Frazier, Ratliff, et al., 2013b).

We did not replicate a dissociation between social disability and adaptive social ability in either males or females with ASD. The discrepancy may be due to the heterogeneous sample included in the current study, as participants represented a broad range of cognitive functioning. Past reports of a dissociation between social disability and social ability have generally been in samples with more homogeneous cognitive profiles, including both higher-functioning (Mcdonald, Thomeer, Donnelly, Tang, & Rodgers, 2015; Saulnier & Klin, 2007) and lower-functioning children with ASD (Liss et al., 2001). Instead, in the current study, we found that both parent-report and clinician-rating measures of social disability were significantly associated with adaptive social functioning in males and females with ASD, though there did seem to be an effect of information source. Parent report of social disability was more strongly related to adaptive social functioning than clinician rating of social disability, consistent with past studies and likely reflecting the fact that adaptive functioning assessment is based on parent report as well (Chang, Lung, Yen, & Yang, 2013; Kanne et al., 2011).

The effect of sex as a moderating factor in prediction of functional ability was minimal and only affected the relationship between clinician rating of social disability and adaptive social functioning, which were more strongly related in females compared to males with ASD. Similarly, differences across sexes in the latent structure of social disability were largely driven by a stronger relationship between parent report and

clinician rating of social disability in females with ASD compared to males. Taken together, parent and clinician judgments of a child's social behavior seemed more convergent in females than in males, though these observer-completed measures were not strongly related to females' social visual engagement.

It is important to contextualize these findings relative to the characteristics of our sample. Participants with ASD represented a wide range of cognitive profiles but were generally higher-functioning; the percentage of participants with IQ scores in the cognitively impaired range was about half the national estimate (Christensen et al., 2016). We did not have exclusion criteria based on cognitive functioning, so this was likely due to the demands of the eye-tracking procedure. Although the task was unstructured, required no or minimal verbal instruction, and did not constrain movement, the procedure did require use of a hat-mounted eye-tracking system.

An additional consideration is that the ASD group included children who met eligibility criteria for a formal diagnosis of ASD and had often received an initial diagnosis prior to study recruitment. Extant studies have suggested that under-identification of ASD in females is a factor contributing to the discrepancy in diagnostic prevalence across sexes (e.g., Dworzynski et al., 2012). Even when presenting with similar severity of ASD symptomatology, females required more comorbid cognitive or behavioral concerns to have a documented formal diagnosis (Giarelli et al., 2010; Hiller, Young, & Weber, 2015; Russell, Steer, & Golding, 2011). This is similar to diagnostic patterns in Attention-Deficit/Hyperactivity Disorder, another neurodevelopmental disorder with a male-biased diagnostic prevalence ratio (Gaub & Carlson, 1997). As a result, the females diagnosed with ASD in the current study may represent a biased

sample of individuals with a clearer diagnostic presentation, perhaps reflected in the greater convergence across parent report and clinician rating of social behavior and adaptive ability in females with ASD.

Nevertheless, the identified differences between males and females with ASD in performance-based measures of social visual engagement provide insight into the clinical phenotype of ASD even when accounting for potential bias in diagnosis. Across performance-based measures, the visual attention and visual scanning of females with ASD was closer to that of same-sex typically-developing peers than in males with ASD. This diminished difference from same-sex typically-developing peers in social visual behavior may contribute to the difficulty in identifying females with ASD, as their observable visual behavior may appear less 'atypical.' Consistent with this, school-age females with ASD were less likely than males with ASD to exhibit atypical eye contact based on clinician rating (Hiller, Young, & Weber, 2015).

In addition, based on the theory that visual engagement patterns reflect an individual's learning within and acting upon the social environment (Klin, Jones, Schultz, & Volkmar, 2003), females with ASD and typically-developing females appeared more frequently similar in how they were responding to and learning from the social environment. Still, although they may have been looking at the same location at the same time, that visual engagement may have had very different adaptive social value for females with ASD and typically-developing peers, as females' diminished difference from normative social visual engagement occurred in the absence of any sex-specific differences across measures of cognition, ASD symptomatology, and adaptive functioning. Similar differences in the social adaptive value of visual engagement

patterns for children with ASD have been noted in previous studies (e.g., Klin, Lin, Gorrindo, Ramsay, & Jones, 2009; Rice et al., 2012). Follow-up studies will seek to assess this potential difference in the social adaptive value of visual engagement using quantitative measures of moment-by-moment affective salience (Shultz et al., 2011).

Of note, sex-specific differences in social visual engagement would not have been identified in this study without assessing degree of deviation from the normative social visual engagement of typically-developing same-sex peers. Differences within diagnostic group alone were limited in magnitude, including between males and females with ASD as well between typically-developing males and females, but the direction of differences was distinct for children with ASD and typically-developing peers, suggesting sex-specific modulation (Constantino, 2016).

The existing literature on sex differences in typical social development has found subtle yet robust differences in communication development beginning in infancy and continuing through the school-age years. Typically-developing females exhibit more context-sensitive social visual engagement than male peers (Ashear & Snortum, 1971; Levine & Sutton-Smith, 1973; Podrouzek & Furrow, 1988), particularly during reciprocal conversation (Levine & Sutton-Smith, 1973). Other identified normative sex differences in childhood have focused on processing of social cues. Female children are more accurate in identifying emotions based on nonverbal cues (Eisenberg & Lennon, 1983). For females but not for males, success in this affective decoding of nonverbal cues is associated with increased social competence in peer relationships (Custrini & Feldman, 1989). Because ratings of perceived social competence in peer relationships are fairly similar across sexes during this age range (Cairns, Leung, Buchanan, & Cairns, 1995),

the findings add to evidence that the normative processes underlying successful peer relationships differs across sexes in school-age children (e.g., Maccoby, 1990).

Whether sex differences in ASD follow a parallel pattern as in typical development or diverge, the distinction is meaningful in evaluating potential risk and protective factors. Based on results from the current study, although females with ASD seemed to superficially exhibit a similarly context-sensitive advantage in social visual engagement like typically-developing peers, it did not translate to increased social adaptation. As future studies continue to examine the etiology and developmental pathogenesis of ASD, it will be critical to consider sex-specific differences in typical social development in order to understand deviations therefrom in ASD.

References

American Psychiatric Association. (2004). *Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition-Text Revision*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition*. Washington, DC: American Psychiatric Association.

Ashear, V., & Snortum, J. R. (1971). Eye contact in children as a function of age, sex, social and intellective variables. *Developmental Psychology*, *4*, 479.

Asperger, H. (1944). Die "autistischen psychopathen" im kindesalter. *Archiv Fur Psychiatrie Und Nervenkrankheiten*, *117*, 76–136.

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), 188–93. http://doi.org/10.1007/s10803-008-0612-4

Baron, R. M., & Kenny, D. a. (1986). The Moderator-Mediator Variable Distinction in Social The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. http://doi.org/10.1037/0022-3514.51.6.1173

Begeer, S., Mandell, D., Wijnker-Holmes, B., Venderbosch, S., Rem, D., Stekelenburg, F., & Koot, H. M. (2013). Sex differences in the timing of identification among children and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(5), 1151–6. http://doi.org/10.1007/s10803-012-1656-z

Bölte, S., Duketis, E., Poustka, F., & Holtmann, M. (2011). Sex differences in cognitive domains and their clinical correlates in higher-functioning autism spectrum disorders. *Autism*, *15*(4), 497–511. http://doi.org/10.1177/1362361310391116

Bölte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*, *1*(6), 354–63. http://doi.org/10.1002/aur.49

Buescher, A. V. S., Cidav, Z., Knapp, M., & Mandell, D. S. (2014). Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatrics*, *19104*(8), 1–8. http://doi.org/10.1001/jamapediatrics.2014.210

Cairns, R. B., Leung, M. C., Buchanan, L., & Cairns, B. D. (1995). Friendships and social networks in childhood and adolescence: fluidity, reliability, and interrelations. *Child Development*, *66*(5), 1330–45.

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

Chang, C.-L., Lung, F.-W., Yen, C.-F., & Yang, P. (2013). Adaptive Behaviors in High-Functioning Taiwanese Children with Autism Spectrum Disorders : an Investigation of the Mediating Roles of Symptom Severity and Cognitive Ability. *Journal of Autism and Developmental Disorders*, *43*, 1347–1355. http://doi.org/10.1007/s10803-012-1684-8

Chawarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2009). A prospective study of toddlers with ASD: short-term diagnostic and cognitive outcomes. *Journal of*

Child Psychology and Psychiatry, and Allied Disciplines, *50*(10), 1235–45. http://doi.org/10.1111/j.1469-7610.2009.02101.x

Chawarska, K., Shic, F., Macari, S., Campbell, D. J., Brian, J., Landa, R., … Bryson, S. (2014). 18-month predictors of later outcomes in younger siblings of children with autism spectrum disorder: a baby siblings research consortium study. *Journal of the American Academy of Child & Adolescent Psychiatry*.

Christensen, D. L., Baio, J., Van Naarden Braun, K., Bilder, D., Charles, J., Constantino, J. N., … Yeargin-Allsopp, M. (2016). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveillance Summary*, *65*(3), 1–23.

Coffman, M. C., Anderson, L. C., Naples, a J., & McPartland, J. C. (2013). Sex Differences in Social Perception in Children with ASD. *Journal of Autism and Developmental Disorders*. http://doi.org/10.1007/s10803-013-2006-5

Constantino, J. (2012). *Social Responsiveness Scale, Second Edition*. Los Angeles, CA: Western Psychological Services.

Constantino, J., & Gruber, C. P. (2005). *Social Responsiveness Scale*. Los Angeles, CA: Western Psychological Services.

Constantino, J. N. (2011). The quantitative nature of autistic social impairment. *Pediatric Research*, *69*(5), 55R–62R. http://doi.org/10.1203/PDR.0b013e318212ec6e

Constantino, J. N. (2016). Data from the Baby Siblings Research Consortium confirm and specify the nature of the female protective effect in autism: A commentary on Messinger et al. *Molecular Autism*, *7*(1), 32. http://doi.org/10.1186/s13229-016-

0092-x

Croen, L. A., Zerbo, O., Qian, Y., Massolo, M. L., Rich, S., Sidney, S., & Kripke, C. (2015). The health status of adults on the autism spectrum. http://doi.org/10.1177/1362361315577517

Custrini, R. J., & Feldman, R. S. (1989). Children's social competence and nonverbal encoding and decoding of emotions. *Journal of Child Clinical Psychology*, *18*, 336–342.

Dawson, G., & Bernier, R. (2013). A quarter century of progress on the early detection and treatment of autism spectrum disorder. *Development and Psychopathology*, *25*, 1455–1472. http://doi.org/10.1017/S0954579413000710

de Winter, J. C. F., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, *39*(4), 695–710. http://doi.org/10.1080/02664763.2011.610445

Duku, E., Szatmari, P., Vaillancourt, T., Georgiades, S., Thompson, A., Liu, X.-Q., … Bennett, T. (2013). Measurement equivalence of the autism symptom phenotype in children and youth. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. http://doi.org/10.1111/jcpp.12103

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), 788–97. http://doi.org/10.1016/j.jaac.2012.05.018

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities.

Psychological Bulletin, *94*(1), 100–131. http://doi.org/10.1037//0033-2909.94.1.100

Elliott, C. D. (2006). *Differential Ability Scales-Second Edition (DAS-II)*. San Antonio, TX: The Psychological Corporation.

Elsabbagh, M., & Johnson, M. H. (2016). Review Autism and the Social Brain : The First-Year Puzzle. *Biological Psychiatry*, 1–6. http://doi.org/10.1016/j.biopsych.2016.02.019

Farley, M. a, McMahon, W. M., Fombonne, E., Jenson, W. R., Miller, J., Gardner, M., … Coon, H. (2009). Twenty-year outcome for individuals with autism and average or near-average cognitive abilities. *Autism Research : Official Journal of the International Society for Autism Research*, *2*(2), 109–18. http://doi.org/10.1002/aur.69

Fein, D., Barton, M., Eigsti, I. M., Kelley, E., Naigles, L., Schultz, R. T., … Tyson, K. (2013). Optimal outcome in individuals with a history of autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(2), 195–205. http://doi.org/10.1111/jcpp.12037

Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Otr, L., Cook, E. H., … Volkmar, F. R. (2000). Practice parameter: Screening and diagnosis of autism: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology*, *55*, 468–479.

Folstein, S., & Rutter, M. (1977). Genetic influences and infantile autism. *Nature*, *265*, 726–728.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591–8. http://doi.org/10.1203/PDR.0b013e31819e7203

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and Cognitive Characteristics of Females and Males With Autism in the Simons Simplex Collection. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(3), 329–340.e3. http://doi.org/10.1016/j.jaac.2013.12.004

Frazier, T. W., & Hardan, A. Y. (2016). Equivalence of symptom dimensions in females and males with autism. *Autism*. http://doi.org/10.1177/1362361316660066

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2013a). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. http://doi.org/10.1177/1362361313500382

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. a, & Constantino, J. N. (2013b). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism : The International Journal of Research and Practice*. http://doi.org/10.1177/1362361313500382

Frazier, T. W., Youngstrom, E. A., Embacher, R., Hardan, A. Y., Constantino, J. N., Law, P., … Eng, C. (2013). Demographic and clinical correlates of autism symptom domains and autism spectrum diagnosis. http://doi.org/10.1177/1362361313481506

Frith, U. (1991). "Autistic psychopathy" in childhood. In U. Frith (Ed.), *Autism and Asperger Syndrome* (pp. 37–92). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511526770.002

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry*,

36, 1036–1045.

Gerhardt, P. F., Zawacki, J., & Satriale, G. (2008). Adaptive Behavior in Adolescents and Adults with Autism Spectrum Disorders: Ongoing Research and Future Implications.

Geschwind, D. H. (2011). Genetics of autism spectrum disorders. *Trends in Cognitive Sciences*, *15*(9), 409–416.

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), 107–16. http://doi.org/10.1016/j.dhjo.2009.07.001

Gotham, K., Brunwasser, S. M., & Lord, C. (2015). Depressive and anxiety symptom trajectories from school age through young adulthood in samples with autism spectrum disorder and developmental delay. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(5), 369–373.e3. http://doi.org/10.1016/j.jaac.2015.02.005

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(5), 693–705. http://doi.org/10.1007/s10803-008-0674-3

Hall, J. a. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*(4), 845–857. http://doi.org/10.1037//0033-2909.85.4.845

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., … Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry*, *68*(11), 1095–102.

http://doi.org/10.1001/archgenpsychiatry.2011.76

Harrison, P., & Oakland, T. (2015). *Adaptive Behavior Assessment System* (3rd Editio). Torrance: Western Psychological Services.

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), 1715–22. http://doi.org/10.1007/s10803-009-0810-8

Havdahl, K. A., Hus Bal, V., Huerta, M., Pickles, A., Øyen, A.-S., Stoltenberg, C., … Bishop, S. L. (2016). Multidimensional Influences on Autism Symptom Measures: Implications for Use in Etiological Research. *Journal of the American Academy of Child & Adolescent Psychiatry*. http://doi.org/10.1016/j.jaac.2016.09.490

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, 1381–1393. http://doi.org/10.1007/s10802-014-9881-x

Hiller, R. M., Young, R. L., & Weber, N. (2015). Sex differences in pre-diagnosis concerns for children later diagnosed with autism spectrum disorder. *Autism*, 1–10. http://doi.org/10.1177/1362361314568899

Hiller, R. M., Young, R. L., & Weber, N. (2015). Sex differences in pre-diagnosis concerns for children later diagnosed with autism spectrum disorder. http://doi.org/10.1177/1362361314568899

Holtmann, M., Bölte, S., & Poustka, F. (2007). Autism spectrum disorders: sex differences in autistic behaviour domains and coexisting psychopathology. *Developmental Medicine and Child Neurology*, *49*(5), 361–6.

http://doi.org/10.1111/j.1469-8749.2007.00361.x

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*. http://doi.org/10.1007/s10803-015-2501-y

Howlin, P., Moss, P., Savage, S., & Rutter, M. (2013). Social Outcomes in Mid- to Later Adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(6).

Hus, V., Bishop, S., Gotham, K., Huerta, M., & Lord, C. (2013). Factors influencing scores on the social responsiveness scale, *2*, 216–224. http://doi.org/10.1111/j.1469-7610.2012.02589.x

Hus, V., & Lord, C. (2014). The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *Journal of Autism and Developmental Disorders*. http://doi.org/10.1007/s10803-014-2080-3

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., … Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. http://doi.org/10.1038/nature13908

Johnson, C. P., Myers, S. M., & the Council on Children with Disabilities. (2007). Identification and Evaluation of Children With Autism Spectrum Disorders. *Pediatrics*, *120*(5), 1183–1215. http://doi.org/10.1542/peds.2007-2361

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, *65*(8), 946–54.

http://doi.org/10.1001/archpsyc.65.8.946

Jones, W., & Klin, A. (n.d.). Entrainment of adaptive action in typically-developing 2-year-olds and the description thereof of autism., Manuscript under review.

Kaiser, H. (1974). An index of factor simplicity. *Psychometrika*, *39*, 31–36.

Kaiser, M. D., Hudac, C. M., Shultz, S., Lee, S. M., Cheung, C., Berken, A. M., … Pelphrey, K. A. (2010). Neural signatures of autism. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(49), 21223–8. http://doi.org/10.1073/pnas.1010412107

Kanne, S. M., Gerber, A. J., Quirmbach, L. M., Sparrow, S. S., Cicchetti, D. V, & Saulnier, C. a. (2011). The role of adaptive behavior in autism spectrum disorders: implications for functional outcome. *Journal of Autism and Developmental Disorders*, *41*(8), 1007–18. http://doi.org/10.1007/s10803-010-1126-4

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

Kim, Y. S., Leventhal, B. L., Koh, Y.-J., Fombonne, E., Laska, E., Lim, E.-C., … Grinker, R. R. (2011). Prevalence of autism spectrum disorders in a total population sample. *The American Journal of Psychiatry*, *168*(9), 904–12. http://doi.org/10.1176/appi.ajp.2011.10101532

Klaiman, C., Fernandez-Carriba, S., Hall, C., & Saulnier, C. (2014). Assessment of Autism Across the Lifespan : A Way Forward. http://doi.org/10.1007/s40474-014-0031-5

Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1430), 345–60.

http://doi.org/10.1098/rstb.2002.1202

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*(9), 809–16.

Klin, A., Lang, J., Cicchetti, D. V, & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders*, *30*(2), 163–7.

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, *459*(7244), 257–61. http://doi.org/10.1038/nature07868

Klin, A., Saulnier, C. A., Sparrow, S. S., Cicchetti, D. V, Volkmar, F. R., & Lord, C. (2007). Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: the Vineland and the ADOS. *Journal of Autism and Developmental Disorders*, *37*(4), 748–59. http://doi.org/10.1007/s10803-006-0229-4

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). Nw York, NY: Kluwer Academic/Plenum Publishers.

Kopp, S., & Gillberg, C. (2011). The Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV): an instrument for better capturing the

autism phenotype in girls? A preliminary study involving 191 clinical cases and community controls. *Research in Developmental Disabilities*, *32*(6), 2875–88. http://doi.org/10.1016/j.ridd.2011.05.017

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. http://doi.org/10.1007/s10803-008-0610-6

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 & Adolescent Psychiatry*. http://doi.org/10.1016/j.jaac.2014.10.003

Lai, M.-C., Lombardo, M. V., Ruigrok, A. N. V., Chakrabarti, B., Auyeung, B., Consortium, M. A., … Baron-Cohen, S. (2016). Quantifying and exploring camouflaging in men and women with autism. *Autism*. http://doi.org/10.1177/1362361316671012

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

Lai, M.-C., Lombardo, M. V, Ruigrok, A. N. V, Chakrabarti, B., Wheelwright, S. J., Auyeung, B., … Baron-Cohen, S. (2012). Cognition in males and females with autism: similarities and differences. *PloS One*, *7*(10), e47198. http://doi.org/10.1371/journal.pone.0047198

Leigh, R. J., & Zee, D. S. (2006). *The Neurobiology of Eye Movements*. New York, NY: Oxford University Press.

Levine, M. H., & Sutton-Smith, B. (1973). Effects of age, sex, and task on visual behavior during dyadic interaction. *Developmental Psychology*, *9*(3), 400–405. http://doi.org/10.1037/h0034929

Levy, D., Ronemus, M., Yamrom, B., Lee, Y., Leotta, A., Kendall, J., … Wigler, M. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*, *70*(5), 886–97. http://doi.org/10.1016/j.neuron.2011.05.015

Liss, M., Harel, B., Fein, D., Allen, D., Dunn, M., Feinstein, C., … Rapin, I. (2001). Predictors and Correlates of Adaptive Functioning in Children with Developmental Disorders, *31*(2).

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *36*(8), 1365–82.

Lord, C., & Corsello, C. (2005). Diagnostic instruments in autistic spectrum disorders. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders* (3rd editio, pp. 730–771). New York, NY: Wiley.

Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism From 2 to 9 Years of Age. *Archives of General Psychiatry*, *63*, 694–701.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagostic Observation Schedule* (2nd Editio). Los Angeles, CA: 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), 185–93.

Lord, C., Schopler, E., & Revicki, D. (1982). Sex differences in autism. *Journal of Autism and Developmental Disorders*, *12*(4), 317–30.

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker' s Congruence Coefficient as a Meaningful Index of Factor Similarity. *Methodology*, *2*(2), 57–64. http://doi.org/10.1027/1614-1881.2.2.57

Lotter, V. (1966). Epidemiology of autistic conditions in young children. *Social Psychiatry and Psychiatric Epidemiology*, *1*, 124–135.

Maccoby, E. E. (1990). Gender and relationships. A developmental account. *The American Psychologist*, *45*(4), 513–20.

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 & Developmental Disorders*, *42*, 1304–1313. http://doi.org/10.1007/s10803-011-1356-0

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), 424–53.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory Factor Analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, *70*, 552–566.

Mcdonald, C. A., Thomeer, M. L., Donnelly, J. P., Tang, V., & Rodgers, J. D. (2015).

VABS-II Ratings and Predictors of Adaptive Behavior in Children with HFASD. *Journal of Developmental and Physical Disabilities*, *27*, 235–247. http://doi.org/10.1007/s10882-014-9411-3

McLennan, J. D., Lord, C., & Schopler, E. (1993). Sex differences in higher functioning people with autism. *Journal of Autism and Developmental Disorders*, *23*(2), 217–27.

McPartland, J. C. (2016). Considerations in biomarker development for neurodevelopmental disorders. *Current Opinion in Neurology*, *29*(2), 1. http://doi.org/10.1097/WCO.0000000000000300

Mullen, E. M. (1995). *Mullen Scales of Early Learning: AGS Edition*. Circle Pines, MN: American Guidance Service.

Neale, B. M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K. E., Sabo, A., … Makarov, V. (2012). Paterrns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*, *485*, 242–246.

Nicholas, J. S., Charles, J. M., Carpenter, L. A., King, L. B., Jenner, W., & Spratt, E. V. E. G. (2008). Prevalence and Characteristics of Children With Autism-Spectrum Disorders. http://doi.org/10.1016/j.annepidem.2007.10.013

Nordahl, C. W., Iosif, A., Young, G. S., Perry, L. M., Dougherty, R., Lee, A., … Amaral, D. G. (2015). Sex differences in the corpus callosum in preschool-aged children with autism spectrum disorder. *Molecular Autism*, *6*, 1–11. http://doi.org/10.1186/s13229-015-0005-4

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., … Stone, W. L. (2011). Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. *Pediatrics*, *128*(3), e488-95.

http://doi.org/10.1542/peds.2010-2825

Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., … Iosif, A.-M. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: a baby siblings research consortium study. *Journal of Child Psychology and Psychiatry*, *56*(9), 988–98. http://doi.org/10.1111/jcpp.12421

Pilowsky, T., Yirmiya, N., Shulman, C., & Dover, R. (1998). The Autism Diagnostic Interview-Revised and the Childhood Autism Rating Scale: differences between diagnostic systems and comparison between genders. *Journal of Autism and Developmental Disorders*, *28*(2), 143–51.

Podrouzek, W., & Furrow, D. (1988). Preschoolers' use of eye contact while speaking: The influence of sex, age, and conversational partner. *Journal of Psycholinguistic Research*, *17*(2), 89–98. http://doi.org/10.1007/BF01067066

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*. http://doi.org/10.1007/s10803-014-2223-6

Reszka, S. S., Boyd, B. a, McBee, M., Hume, K. a, & Odom, S. L. (2013). Brief Report: Concurrent Validity of Autism Symptom Severity Measures. *Journal of Autism and Developmental Disorders*. http://doi.org/10.1007/s10803-013-1879-7

Rice, K., Moriuchi, J. M., Jones, W., & Klin, A. (2012). Parsing heterogeneity in autism spectrum disorders: visual scanning of dynamic social scenes in school-aged children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(3), 238–48. http://doi.org/10.1016/j.jaac.2011.12.017

Robins, D., Fein, D., & Barton, M. (1999). *The Modified Checklist for Autism in Toddlers (M-CHAT)*. Storrs, CT: University of Connecticut.

Russell, G., Steer, C., & Golding, J. (2011). Social and demographic factors that influence the diagnosis of autistic spectrum disorders. *Social Psychiatry and Psychiatric Epidemiology*, *46*(12), 1283–93. http://doi.org/10.1007/s00127-010-0294-z

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.

Rutter, M., LeCouter, A., & Lord, C. (2003). *Autism Diagnostic Interview-Revised*. Los Angeles, CA: Western Psychological Services.

Sanders, S. J., Ercan-Sencicek, a G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., … State, M. W. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*, *70*(5), 863–85. http://doi.org/10.1016/j.neuron.2011.05.002

Saulnier, C. A., & Klin, A. (2007). Brief report: social and communication abilities and disabilities in higher functioning individuals with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *37*(4), 788–93. http://doi.org/10.1007/s10803-006-0288-6

Schaefer, G. B., & Mendelsohn, N. J. (2013). Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *15*(5), 399–407. http://doi.org/10.1038/gim.2013.32

Schopler, E., Van Bourgondien, M. E., Wellman, G. J., & Love, S. R. (2010). *Childhood*

Autism Rating Scale, Second Edition (CARS2). Torrance, CA: Western Psychological Services.

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., … Cuniff, C. (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 and Adolescent Psychiatry*, *48*(5), 474–83. http://doi.org/10.1097/CHI.0b013e31819b3848

Shultz, S., Klin, A., & Jones, W. (2011). Inhibition of eye blinking reveals subjective perceptions of stimulus salience. *Proceedings of the National Academy of Sciences*, *108*(52), 21270–21275. http://doi.org/10.1073/pnas.1109304108

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Sparrow, S., & Cicchetti, D. (2005). *Vineland Adaptive Behavior Scales-II (Interview Survey Form and Parent Rating Rating Form)*. Circle Pines, MN: American Guidance Service.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales, Interview Edition, Expanded Form*. Circle Pines, MN: American Guidance Service.

Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaire* (3rd Editio). Baltimore: Paul H Brookes Publishing Co.

Szatmari, P., Liu, X.-Q., Goldberg, J., Zwaigenbaum, L., Paterson, A. D., Woodbury-Smith, M., … Thompson, A. (2012). Sex differences in repetitive stereotyped behaviors in autism: implications for genetic liability. *American Journal of Medical*

Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of Psychiatric Genetics, *159B*(1), 5–12. http://doi.org/10.1002/ajmg.b.31238

Tassé, M. J., Schalock, R. L., Balboni, G., Bersani, H., Borthwick-Duffy, S. a, Spreat, S., … Zhang, D. (2012). The construct of adaptive behavior: its conceptualization, measurement, and use in the field of intellectual disability. *American Journal on Intellectual and Developmental Disabilities*, *117*(4), 291–303. http://doi.org/10.1352/1944-7558-117.4.291

Tsai, L. Y., & Beisler, J. M. (1983). The development of sex differences in infantile autism. *The British Journal of Psychiatry : The Journal of Mental Science*, *142*, 373–8.

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*, 627–635. http://doi.org/10.1007/s10803-013-1913-9

Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, *45*(1), 135–70.

Volkmar, F., Szatmari, P., & Sparrow, S. S. (1993). Sex differences in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *23*(4), 579–91.

Wang, H. K. C., Sung, C., Sa, J., Leahy, M. J., & Sa, J. (2015). Gender Differences in

Vocational Rehabilitation Service Predictors of Successful Competitive Employment for Transition-Aged Individuals with Autism. http://doi.org/10.1007/s10803-015-2480-z

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. Pearson.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence-Third Edition*. London: Pearson.

Wechsler, D. (2004). *The Wechsler Intelligence Scale for Children-Fourth Edition*. London: Pearson.

Wiggins, L. D., Baio, J., & Rice, C. (2006). Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. *Journal of Developmental and Behavioral Pediatrics*, *27*, S79–S87. http://doi.org/10.1097/00004703-200604002-00005

Wing, L. (1981). Sex Ratios in Early Childhood Related Conditions and related conditions. *Psychiatry Research*, *137*, 129–137.

Yuen, R. K. C., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., … Fernandez, B. a. (2015). Whole-genome sequencing of quartet families with autism spectrum disorder, (October 2014). http://doi.org/10.1038/nm.3792

Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., … Yirmiya, N. (2009). Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics*, *123*(5), 1383–91. http://doi.org/10.1542/peds.2008-1606

Figure 1. Stimuli and region-of-interest coding for performance-based measurement of social visual engagement.

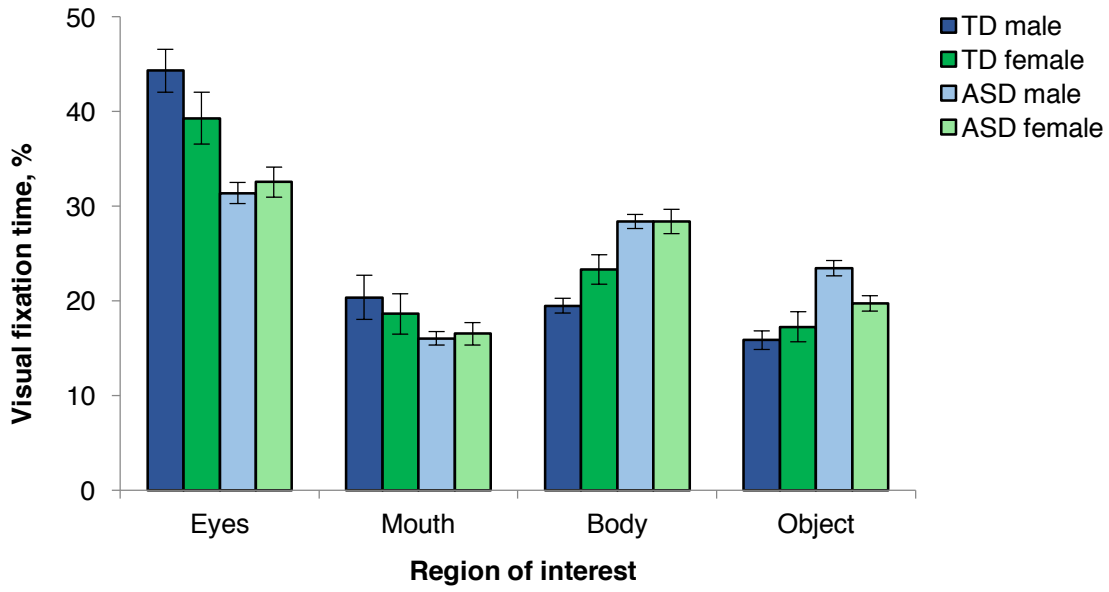

Figure 2. Distribution of visual fixations on regions of interest for males and females in the typically-developing (TD) and ASD groups. Error bars represent SEM.

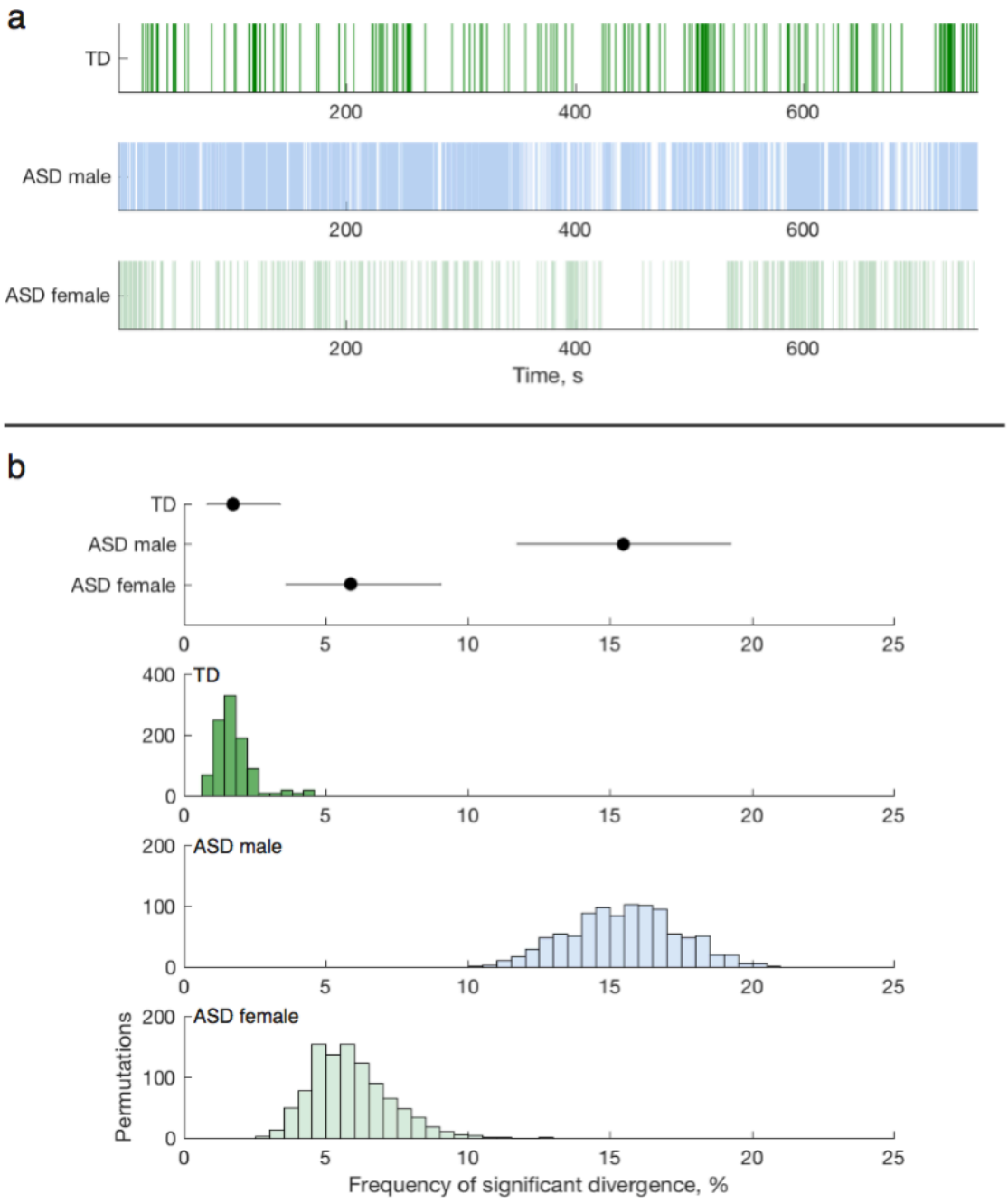

Figure 3. Sex-specific differences in divergence from normative social visual engagement in ASD. (a) Timelines of significant divergence between typically-developing males and females, between typically-developing males and males with ASD, and between typically-developing females and females with ASD. Each vertical line represents a frame in which groups significantly diverged. (b) Bootstrapped means and 95% confidence intervals derived from permutation testing with 1000 random resamplings indicated that the visual scanning of males with ASD significantly diverged from normative same-sex patterns of social visual engagement more frequently than that of females with ASD. Both males and females with ASD diverged from normative same-sex patterns of social visual engagement more frequently than typically-developing males and females diverged. TD = typically-developing.

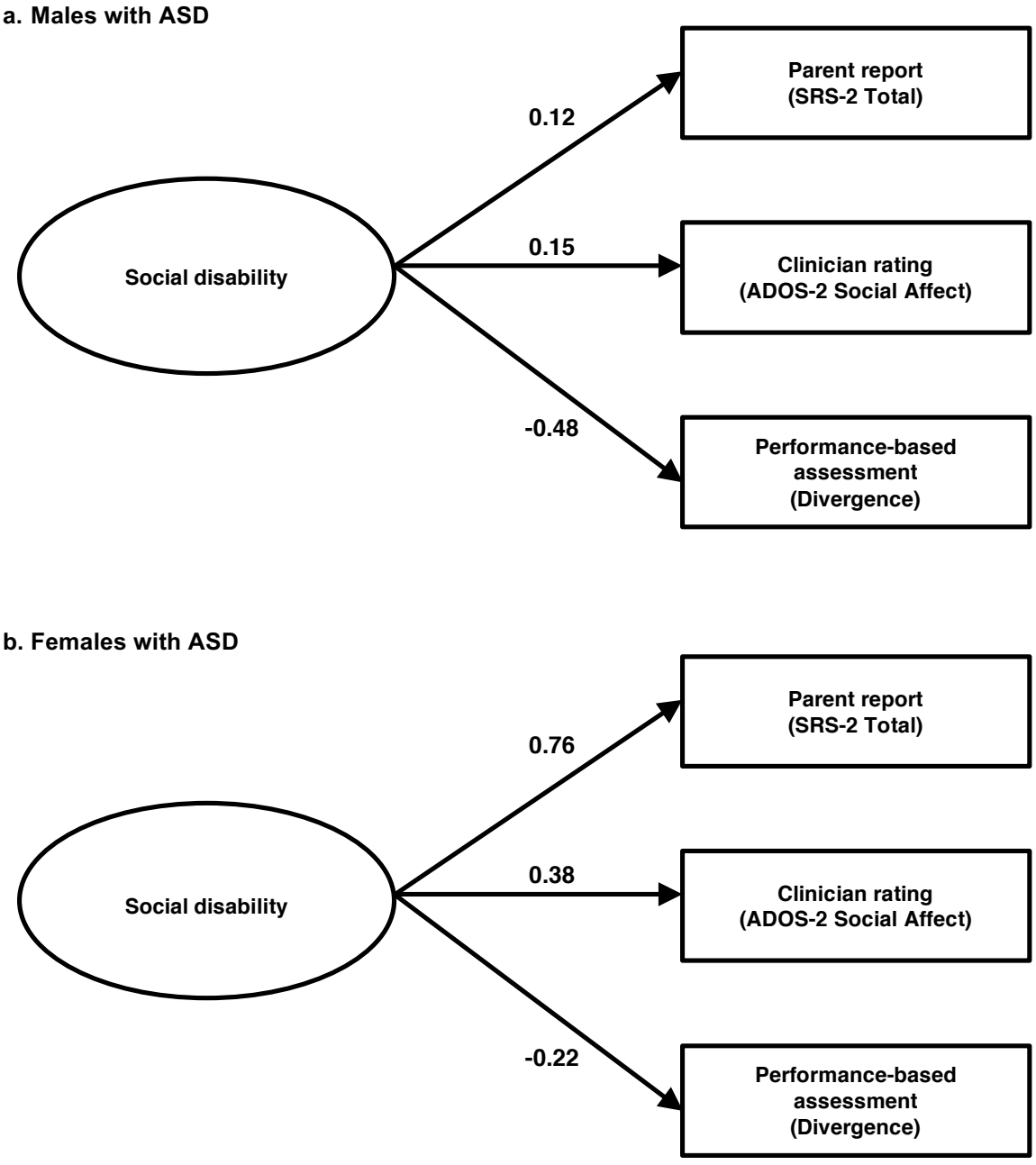

Figure 4. Single-factor model of social disability with factor loadings from exploratory factor analysis for (a) males and (b) females with ASD.

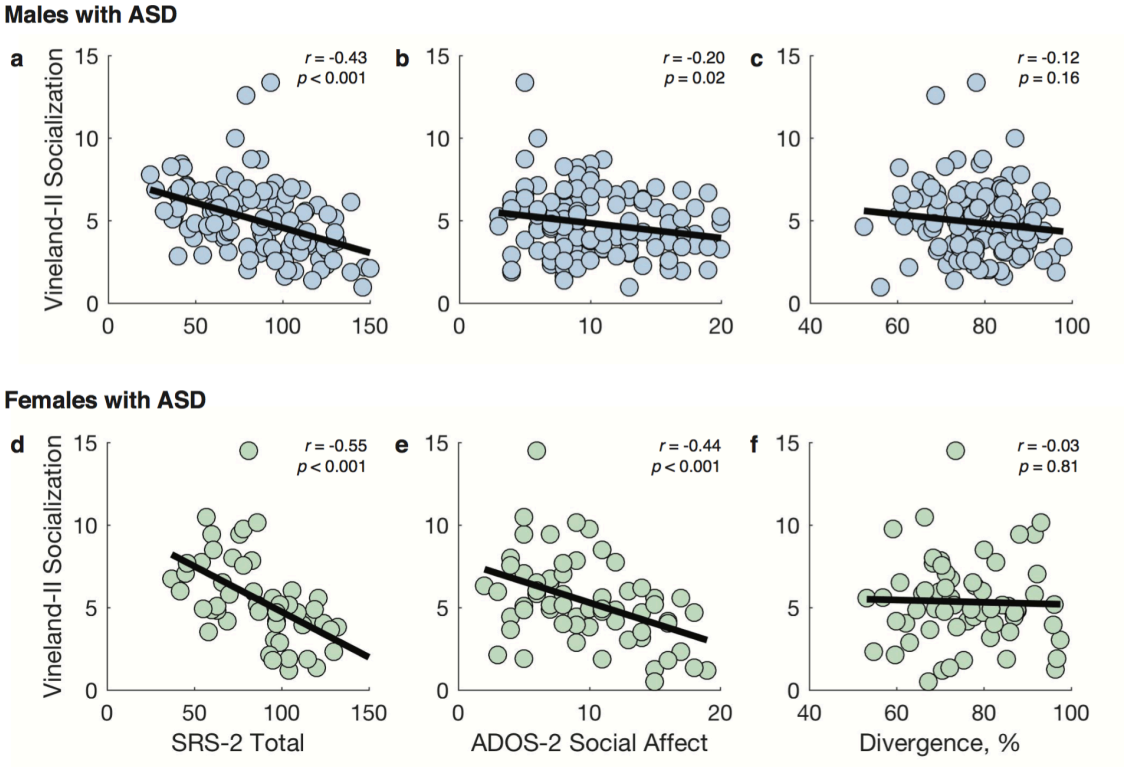

Figure 5. Dimensional associations between social disability and adaptive social functioning across modes of measurement. (a) In males with ASD, parent report of social disability on the SRS-2 was not significantly associated with adaptive social functioning, but (b) clinician rating of social disability was negatively correlated with adaptive social functioning. (c) There was no relationship with performance-based measurement of social disability in males with ASD. In females with ASD, both (d) parent report and (e) clinician rating of social disability were associated with adaptive social functioning. (f) There was no relationship with performance-based measurement of social disability in females with ASD. ADOS-2 = Autism Diagnostic Observation Schedule, Second Edition; SRS-2 = Social Responsiveness Scale, Second Edition; Vineland-II = Vineland Adaptive Behavior Scales, Second Edition.

Table 1. Participant characterization data

	TD Group		ASD Group		Test statistic	*p* value
	Male (n = 27)	**Female (n = 32)**	**Male (n = 136)**	**Female (n = 64)**		
Age, years	**9.9** (2.4)	**10.1** (3.5)	**10.2** (2.9)	**10.8** (3.0)	$F_{3,255} = 0.89$	0.45
Cognition						
Full Scale IQ	**108.0** (17.8)	**113.4** (15.4)	**91.3** (23.4)	**92.6** (24.0)	$F_{3,255} = 11.8$	<0.001
Verbal IQ	**109.8** (18.3)	**116.0** (15.7)	**90.6** (24.8)	**91.4** (27.0)	$F_{3,255} = 13.6$	<0.001
Nonverbal IQ	**104.8** (16.9)	**109.3** (14.2)	**93.7** (22.2)	**94.1** (21.7)	$F_{3,255} = 6.55$	<0.001
Social disability						
SRS-2 Total			**86.3** (30.5)	**87.8** (27.2)	$t_{161} = -0.30$	0.76
SRS-2 T-score			**70.9** (12.1)	**74.9** (11.6)	$t_{161} = -1.96$	0.05
ADOS-2 Social affect			**10.6** (4.4)	**9.8** (4.4)	$t_{194} = 1.24$	0.22
ADOS-2 RRB			**3.1** (1.9)	**3.4** (2.1)	$t_{194} = -1.03$	0.31
ADOS-2 CSS			**7.3** (2.1)	**6.9** (2.4)	$t_{194} = 1.19$	0.24
Adaptive ability						
Vineland-II Socialization AE			**4.8** (2.0)	**5.3** (2.6)	$t_{195} = -1.46$	0.15

Note: Data presented as mean (SD). ADOS-2 = Autism Diagnostic Observation Schedule, Second Edition; AE = Age Equivalent, years; ASD = Autism Spectrum Disorder; CSS = Calibrated Severity Score; RRB = Restricted and repetitive behaviors; SRS-2 = Social Responsiveness Scale, Second Edition; TD = Typically-developing; Vineland-II = Vineland Adaptive Behavior Scales, Second Edition.

Table 2. Correlations between measures of social disability

	Males with ASD			Females with ASD		
	SRS-2	**ADOS-2**	**Divergence**	**SRS-2**	**ADOS-2**	**Divergence**
SRS-2 Total	1.00			1.00		
ADOS-2 Social affect	0.02	1.00		0.29*	1.00	
Divergence from normative visual scanning	0.09	0.16	1.00	-0.18	0.01	1.00

Note: Data presented are Pearson correlation coefficients. Across all measures, higher scores indicate greater social disability. *$p < 0.05$; ADOS-2 = Autism Diagnostic Observation Schedule, Second Edition; ASD = Autism Spectrum Disorder; SRS-2 = Social Responsiveness Scale, Second Edition.