CrossMark

ORIGINAL PAPER

# Parent-Endorsed Sex Differences in Toddlers with and Without ASD: Utilizing the M-CHAT

Roald A. Øien[1,2] · Logan Hart[1,2] · Synnve Schjølberg[3] · Carla A. Wall[2] ·
Elizabeth S. Kim[4] · Anders Nordahl-Hansen[5] · Martin R. Eisemann[1] ·
Katarzyna Chawarska[2] · Fred R. Volkmar[2] · Frederick Shic[2]

**Abstract** Sex differences in typical development can provide context for understanding ASD. Baron-Cohen (Trends Cogn Sci 6(6):248–254, 2002) suggested ASD could be considered an extreme expression of normal male, compared to female, phenotypic profiles. In this paper, sex-specific M-CHAT scores from $N = 53,728$ 18-month-old toddlers, including $n = 185$ (32 females) with ASD, were examined. Results suggest a nuanced view of the "extreme male brain theory of autism". At an item level, almost every male versus female disadvantage in the broader population was consistent with M-CHAT vulnerabilities in ASD. However, controlling for total M-CHAT failures, this male disadvantage was more equivocal and many classically ASD-associated features were found more common in non-ASD. Within ASD, females showed relative strengths in joint attention, but impairments in imitation.

✉ Roald A. Øien
roald.a.oien@uit.no; roald.oien@yale.edu

1 Department of Psychology, UiT - The Arctic University of Norway, PB 6050, N-9037, Tromsø, Norway

2 Child Study Center, Yale School of Medicine, Yale University, New Haven, CT, USA

3 Mental Health, Norwegian Institute of Public Health, Oslo, Norway

4 Center for Autism Research, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

5 Department of Special Needs Education, University of Oslo, Oslo, Norway

## Introduction

Studies examining gender differences in typically developing infants and toddlers show sex-specific patterns in behavior and development. Differences include higher activity level in males, while social orienting, reciprocity, eye-contact and language development tend to represent areas of strength for females (Bouchard et al. 2009; Connellan et al. 2000; Hittelman and Dickes 1979; Lutchmaya and Baron-Cohen 2002; Maccoby and Jacklin 1974; Reilly et al. 2009; Riddoch et al. 2007; Trouton et al. 2002; Zambrana et al. 2012). One study found that infants as young as 1-day of age showed sex-specific looking preferences, with males preferring mechanical objects while females showed a greater degree of interest in faces (Connellan et al. 2000). These findings are consistent with several studies suggesting that females in the general population outperform males in a variety of skills typically perceived as being deficits within ASD, e.g. sensitivity to facial expressions (McClure 2000; Montagne et al. 2005), performance on questionnaires measuring empathy (Davis 1994), age when reaching developmental milestones e.g. theory of mind (Happé 1995), and language development (Halpern 1997; Zahn-Waxler et al. 2006).

Sex differences in autism-related symptoms among children with ASD is an emerging but under-researched area. The most frequently reported sex difference in ASD is the disproportionally higher male to female prevalence, consistently reported since the seminal studies by Kanner (1943) and Asperger (1944). Fombonne (2003, 2005, 2007) reported across studies male to female prevalence ratios

ranging from 4.3:1, with 5.5:1 in groups within the normal IQ range. For moderate to severe intellectual disability male to female ratios of 1.33:1 (McCarthy et al. 1984) and 1.95:1 (Fombonne 2005, 2007) have been reported. While numerous theories have been forwarded to explain the causal mechanisms of this predominantly high male–female ratio in ASD, the topic remains widely debated in the current literature.

For example, the positive correlation between intellectual disability and severity of symptoms (Carter et al. 2007; Kopp and Gillberg 2011; Lai et al. 2012; Mayes and Calhoun 2011), combined with the fact that males are more prone to developmental delay have led some to hypothesize that the higher prevalence of autism in males stems from a greater risk of developmental disability (Boyle et al. 2011). The exact nature of this relationship is unclear, and studies have found evidence that sex differences in cognitive performance, adaptive abilities and repetitive behaviors do not appear to be ASD specific, but instead bear a closer resemblance to those found in typically developing children (Messinger et al. 2015; Zweigenbaum et al. 2012). However, other viewpoints stress specific biological factors related to autism, e.g. as illustrated by findings which note that higher genetic risk for autism may occur in females with idiopathic autism (Gilman et al. 2011; Levy et al. 2011; Robinson et al. 2013; Skuse 1997, 2000).

While the specific behavioral influences of sex differences in ASD presents as unclear within current literature. Some have posited that these behavioral differences could, in part, influence the observed asymmetry in sex prevalence if they contribute to the risk that subtle cases of ASD in females may go unrecognized (Dworzynski et al. 2012). This research suggests that this may be particularly true for females falling into the average range of IQ and who, as a group, typically tend to display fewer disruptive behavioral outbursts than their male peers (Dworzynski et al. 2012). Less disruptive behavior and outbursts might be related to the fact that females score higher on internalizing behavior and lower on externalizing behaviors compared to males (Bölte et al. 2011; Mandy et al. 2012; Solomon et al. 2012; Szatmari et al. 2012), an area of work which has achieved a somewhat greater degree of consensus than the existing literature on more specific sex difference in ASD. On the other hand Baron-Cohen and colleagues, have turned to the general population and suggest that he "systemizing cognitive profile" typically found in males within the general population is reflected in gender differences in autism (Auyeung et al. 2013; Baron-Cohen 2002; Baron-Cohen and Benenson 2003; Baron-Cohen et al. 2005; Bölte et al. 2011; Hattier et al. 2011; Mandy et al. 2012; Szatmari et al. 2012).

In line with these studies, which widen the context by which behavioral manifestations of autism are considered by considering population-based phenomena, the present study further extends continuum-based perspectives of ASD-related behaviors in a large population based sample of children between 17 and 30 months. This perspective is consistent with the Research Domain Criteria (RDoC: Insel et al. 2010). The behavior rated on the Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al. 2001) are examined regarding sex differences. The overall aim for the present study is to examine sex differences in ASD-relevant behaviors as endorsed by parents in a cohort of children between 17 and 30 months of age. Specific aims include the following:

(a) To examine differences in overall endorsement of autistic symptoms associated with sex and diagnosis; (b) to examine individual behavioral symptoms associated with a diagnosis of ASD versus non-ASD; (c) to examine if non-ASD children differ by sex in symptoms endorsed at an M-CHAT item level; and (d) to examine if ASD children differ by sex in symptoms endorsed at an M-CHAT item level. We hypothesize that sex differences observed in ASD would follow similar patterns as those seen in males not receiving an ASD diagnosis (Baron-Cohen 2002, 2009).

## Methods

### Participants

The study sample is derived from the Norwegian Mother and Child Cohort Study (MoBa) (Magnus et al. 2006) and one of its sub-studies, the Autism Birth Cohort Study (ABC) (Stoltenberg et al. 2010). MoBa is a prospective population-based pregnancy cohort study established by the Norwegian Institute of Public Health. Participants included pregnant mothers recruited during the years 1999–2010 at ultrasound examinations at approximately week 18 of pregnancy. In total, 40.6 % of invited mothers consented to participate. The cohort comprises 114,500 children and 95,200 mothers. The first data were collected during pregnancy, and each mother received several questionnaires containing items from a number of age-appropriate scales for their participating child with follow-ups at 6, 18 and 36 months as well as at 5, 7, 8 and 13 years of age. Diagnoses of ASD were obtained from the Autism Birth Cohort´s (ABC) clinical records, in our sample, assessed at 3.5 years (Stoltenberg et al. 2010). A child's diagnose were rendered completely blinded for ratings in MoBa-questionnaires and with no knowledge about any previous diagnosis by specialized services. In addition, from the clinical records registered in the Norwegian Patient Registry (NPR), ASD-diagnosis registered at any-time from the first 1 year of age were used. Exact age at first diagnosis and level of functioning is not available for patients

retrieved from the NPR, as it only lists the diagnostic status of a given child for the specific year(s) seen by specialized services. The possibility of person specific identification of diagnosis in the NPR registry started in 2008. The non-ASD group consist of mostly typically developing children, though some children might have other diagnoses. Children in the MoBa-sample is currently ranging from 7 to 17 years of age. Both the MoBa and the ABC study obtained written informed consent from participating mothers and were approved by the Norwegian Data Inspectorate, as well as the Regional Committee for Medical and Health Research Ethics South-East Norway (REK). The present study uses the MoBa data release version 9, more about the sample selection displayed in Fig. 1.

## Measures

The M-CHAT was designed to screen for ASD early in development, i.e. around 16–30 months of age (Robins et al. 2001). It includes 23 yes-or-no questions to be completed by parents and followed-up by an interview with parents of children yielding a positive M-CHAT screen score. The M-CHAT was designed to be completed quickly in the waiting room of a primary care provider and has become one of the most frequently used screening instruments for ASD (Ibanez et al. 2014). The later M-CHAT-R (Robins et al. 2014) has been recommended in the United States for use with toddlers between 18 and 24 months of age (American Academy of Pediatrics 2006). In the present study, the M-CHAT checklist is used as an ASD-specific behavior measure in a large cohort, making it possible to examine

early sex differences in children with or without ASD. No follow-up interviews were conducted.

Approximately 73 % of MoBa participants completed the 18-month questionnaire, which included the 23 items of the M-CHAT. The ASD sample in the present study does not reflect the true national prevalence rate of ASD, as the present study due to the fact that not all children in Norway participated, and that all children with missing responses on the M-CHAT were excluded. At the same time, new cases of ASD will be diagnosed with increasing age and subsequently listed in NPR (Súren et al. 2012). Children in the current sample were born between 2003 and 2009 and at the linkage to NPR, autumn 2014, the youngest children in the sample were still only 5 years of age.

Each item in M-CHAT was scored 0 = non failure, 1 = failure according to the manual (Robins et al. 2001) and a total score was calculated summarizing each child´s *number of failed items* to establish an overall measure for presence of autistic-like behavior. In addition, a list of 6 out of the 23 M-CHAT items constituting the most critical items in predicting an ASD diagnosis (Robins et al. 2001) was summarized. The mean of total failed items in M-CHAT and the six critical items are listed in Table 1.

## Statistical Analyses

In line with aim (a) to examine the overall endorsed autistic symptoms with respect to sex and diagnosis by examining M-CHAT total number of failed items, we first conducted a two-way ANOVA (sex*ASD diagnosis) with total number of failed M-CHAT items as outcome. As it is expected that
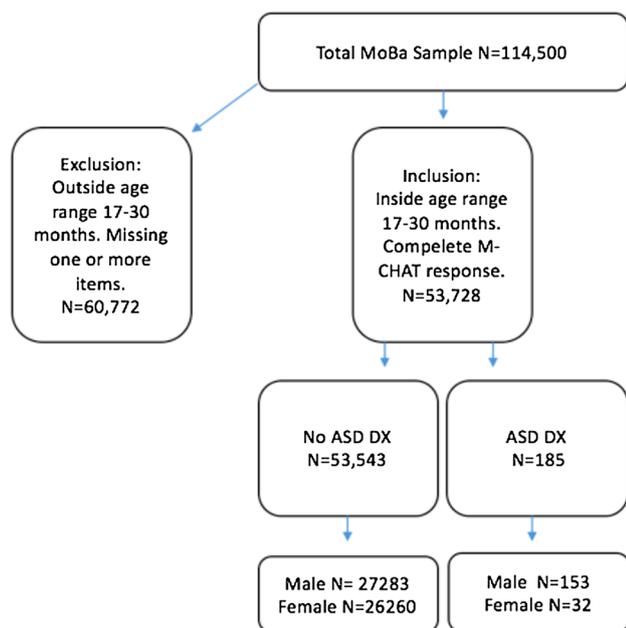


**Fig. 1** Sample selection

**Table 1** Demographics of patients performance on the M-CHAT

|  | Non ASD total | Non ASD male | Non ASD female | ASD total* | ASD male | ASD female |
|---|---|---|---|---|---|---|
| N | 53,543 | 27,283 | 26,260 | 185 | 153 | 32 |
| Mean age | 18.53 | 18.53 | 18.53 | 18.55 | 18.57 | 18.48 |
| Mean age (S.D.) | 0.58 | 0.58 | 0.58 | 0.66 | 0.68 | 0.49 |
| Mean of total failed items | 0.80 | 0.84 | 0.74 | 3.11 | 2.68 | 5.16 |
| Mean of total failed items (S.D.) | 1.17 | 1.22 | 1.11 | 4.00 | 3.54 | 5.34 |
| Mean critical items | 0.15 | 0.17 | 0.13 | 0.96 | 0.84 | 1.50 |
| Mean critical items (S.D.) | 0.44 | 0.47 | 0.40 | 1.38 | 1.26 | 1.76 |

*This does not reflect ASD prevalence in Norway. These are children with complete M-CHAT responses in the MoBa study. All participants with missing items are excluded

the total number of failed items would differ between ASD and non-ASD, the total failure rate (i.e. severity) should be controlled for in the subsequent analyses to better understand sex-specific phenomena taking into account symptom severity. Due to smaller sample size in the ASD female group, we did not include an interaction term between the severity score and diagnosis or sex (depending on analysis) in the subsequent logistic regression models. Including the interaction item introduced high collinearity between predictors for several item analyses leading to instability in parameter estimates. It was hence omitted from all the item level models for comparability.

In line with aim (b) conducting a logistic regression to explore the specificity of difficulties in ASD versus non-ASD groups by examining individual M-CHAT items. We did this analysis first without controlling for number of failed items showing the effect of diagnosis on each item. Next, to explore the difference in pattern of endorsed items comparing ASD with non-ASD children, we did the same analysis with diagnosis as predictor controlling for levels of failure (i.e. severity). To ease the interpretation of the beta we centered the total failure rate to the unweighted mean of failed items in the total sample based on mean failed items for non-ASD and mean failed items for ASD.

In line with aim (c), to examine if non-ASD children differ by sex in symptoms endorsed at an M-CHAT item level by performing a logistic regression for each M-CHAT item including sex as predictor controlling for levels of failure. The numbers of failures were expressed as a percentile score calculated separately for males and females and was included to control for overall failure. For ease of interpretation of coefficients, the total failure rate expressed as percentile-score was centered on the median in our statistical model. Male was used as reference group, leading to the interpreted of the beta with a focus on female advantage (low failing rate, negative beta) or disadvantage (high failing rate, positive beta). In addition, we performed a logistic regression without controlling for overall failed items (Supplementary Table 1). Bonferroni correction for multiple comparisons was applied and the α-level was set to 0.00217.

In line with aim (d) to examine if examine ASD children differ by sex in symptoms endorsed at an M-CHAT item level. We performed a logistic regression for each M-CHAT item, including sex as predictor, controlling for overall total failure rate. For ease of interpretation, the total failure rate was centered to the unweighted mean of the mean value of total failed items for ASD males and ASD females. The mean was used in this model instead of the median to represent even class priors on males and females with ASD and to provide a less prevalence-biased interpretation of any observed phenomena. The analytical approach of controlling for severity (i.e. total number of failed items) maximized the power in the comparisons, allowing individuals in the smaller ASD group to be analyzed together without the need of small, stratified samples with less power. In addition, we performed a logistic regression without controlling for overall failed items (Supplementary Table 1). Statistical analyses were conducted using IBM SPSS 23.

## Results

In line with aim (a), there were main effects both of diagnosis [$F(1,53724) = 723.859, p < 0.001, partial \, \eta^2 = 0.013$] and sex [$F(1,53724) = 104.645, p < 0.001, partial \, \eta^2 = 0.002$], with greater numbers of failed items for individuals with ASD compared to non-ASD and non-ASD males compared to non-ASD females. However, there was also an interaction effect between diagnosis and sex, showing that the relation between sex and severity of mean M-CHAT score depended on the diagnostic status of the child [$F(1,53724) = 123.374, p < 0.001, partial \, \eta^2 = 0.002$]. To reiterate, the ANOVA indicated that toddlers with ASD had more autism-related symptoms than the non-ASD toddlers and that non-ASD males had more autism-related symptoms than non-ASD females. Furthermore, females with ASD had higher failure rates on M-CHAT than males with ASD.

### Non-ASD Versus ASD

In line with aim (b) a logistic regression analysis on individual items of the M-CHAT, controlling for number of failed items, was conducted to explore differences in parent endorsed ASD symptoms between children receiving an ASD-diagnosis and children who have not received an ASD diagnosis. A positive β coefficient indicates that children with an ASD diagnosis were more likely to fail an item compared to non-ASD children, while a negative β coefficient indicates the opposite. After adjusting for multiple comparisons ($p = 0.00217$), ASD children were found to be less likely to fail items 4 (*Enjoy peek-a-boo*), 10 (*Eye contact*), 11 (*Oversensitivity to noise*), 18 (*Unusual finger/hand movements*), and 20 (*suspected deafness*), 22 (*Stare at nothing*) and 23 (*Check parents reaction*). A logistic regression analysis without controlling for number of failed items revealed that children with ASD were more likely to fail all items except item 11 (*Oversensitivity to noise*) compared to non-ASD. These findings reveal that without controlling for the overall number of failed items, non-ASD children are less likely to fail almost all items. However, when controlling for this factor, multiple items are non-specific to ASD and observed more in non-ASD.

### Non-ASD: Males Versus Females

In line with aim (c), a logistic regression on individual M-CHAT items was conducted to explore sex differences

in parent-endorsed ASD symptoms in children who currently did not have an ASD diagnosis. Only one item had a significant and positive β coefficient for the effect of sex after adjusting for multiple comparisons, i.e. item 3 (*Enjoy climbing on things*). Furthermore, non-ASD females were overall less likely than males to fail items, i.e. negative and significant β coefficients, 5 (*Pretend play*), 9 (*Show objects to others*), 10 (*Eye contact*), 13 (*Imitation*), 15 (*Follow to point*), 20 (*Suspected deafness*), 21 (*Understand speech*) and 23 (*Check parent's reaction*). These findings reflect that females are generally less likely to fail items related to social motivation on the M-CHAT.

### ASD: Males Versus Females

In line with aim (d) a logistic regression analysis on item level was conducted to explore sex differences in children receiving an ASD diagnosis controlling for number of failed items, centered on the median. This analysis revealed that ASD females were more likely to fail item 13 (*Imitation*) compared to males. However, ASD females were less likely than males to fail item 15 (*Follow to point*), which may indicate a female strength concerning joint attention.

## Discussion

The present study found that females in the non-ASD sample failed significantly fewer items ($M = 0.74$, SD = 1.11) than males ($M = 0.84$, SD = 1.22), which might reflect that non-ASD males show more autism-like symptoms at 18 months than females. This might reflect that males show slower developmental gains early in key areas relevant for autism. In children with ASD the opposite relation emerges. Females failed significantly more items ($M = 5.16$, $SD = 5.34$) than males ($M = 2.68$, $SD = 3.54$). This might suggest that in this sample, females with ASD, diagnosed at any time, expressed a greater load of ASD symptoms compared to males, as rated at 18 months of age. This might indicate that less severe cases of females are not identified and consequently has not received an ASD diagnosis. This is in line with previous studies (Dworzinsky et al. 2012; Robinson et al. 2013), that suggest that females need more severe developmental, behavioral or intellectual delay/deviance to be diagnosed with ASD. Another hypothesis is that the difference in sex-ratio and symptom pattern might be related to the fact that males tend to show a higher level of repetitive behaviors than females (Szatmari et al. 2012), while females tend to express better and more complex language than males (Salomone et al. 2016). This could potentially have a masking effect on social communication in females, causing complexities in the diagnostic process and fewer high

functioning females are identified and diagnosed with ASD.

Comparing non-ASD to ASD on symptom pattern, without controlling for number of failed items (i.e.as an index of severity), suggests that the pattern of failed items in the entire ASD group compared to non-ASD, aligns with the pattern of non-ASD males compared to non-ASD females. Isolated, this provide a nuanced support to the assertion that ASD represents an extreme version of male developmental strengths and weaknesses. However, when controlling for number of failed items (severity), a male disadvantage was more equivocal and many typically ASD-associated features were found more common in non-ASD children at 18 months. This finding might be influenced by characteristics of the non-ASD sample, which includes some individuals with other developmental delays, i.e. disorders such as ADHD/ADD, profound disability and other diagnoses that might share many of the developmental characteristics seen in ASD. However, in a study if this sample-size, they will be in minority to typically developing children. To reiterate, when comparing non-ASD to ASD, controlling for number of failed items, inconsistencies are present in terms of the theory that ASD are an extreme version of the typical male developmental profile, as the ASD sample also exhibit strengths that are in line with a typical female developmental profile.

The logistic regression comparison of non-ASD males and females revealed sex-specific strengths. Pretend play, imitation and follow to point among others emerged as especially strong for females (Table 2c). Strengths in these areas are in line with previous studies showing that infant females have advantages in social orientation (Maccoby and Jacklin 1974), imitation (Hittelman and Dickes 1979), and joint attention in early childhood (Mundy et al. 2007).

An analysis of sex differences in the ASD group, after controlling for failed items, revealed that ASD females do not possess the relative strength when compared to ASD males on imitation, contrary to the results in the non-ASD sample. However, they show relative strength in following a pointing gesture, as is also seen in the non-ASD group. Except for these two screened behaviors, strengths and weaknesses seem generally non-specific to sex and instead vary by the presence of an ASD diagnosis. This could indicate that females require greater impairment in imitation abilities before meeting the diagnostic criteria for ASD. It is also important to note that imitation might represent a complex construct encompassing important pillars of social cognition and communication, and the full wording of the question on the M-CHAT may convey an altogether narrower meaning in this instance. The M-CHAT offers a very specific example: "*Does your child imitate you? (e.g., you make a face, will your child imitate it*)." Parents´ interpretation of this item might be driven by the example of facial

Table 2 Item level analysis M-CHAT: main effects of diagnosis and sex

| | (2a) Main effect diagnosis[a],* | | | (2b) Main effect Diagnosis[b],* | | | (2c) Main effect sex in non-ASD[c],* | | | (2d) Main effect sex in ASD[d],* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (β) | S.E | p[α] | (β) | S.E | p[α] | (β) | S.E | p[α] | (β) | S.E | p |
| 1. Enjoy being swung/bounced | 1.6 | 0.4 | <0.001* | −0.9 | 0.5 | 0.105 | −0.2 | 0.1 | 0.025 | −0.6 | 1.2 | 0.600 |
| **2. Interest in other children** | 2.6 | 0.3 | <0.001* | 0.09 | 0.5 | 0.846 | −0.4 | 0.2 | 0.008 | −0.8 | 1.0 | 0.469 |
| 3. Enjoy climbing on things | 2.6 | 0.3 | <0.001* | −0.4 | 0.5 | 0.409 | 0.4 | 0.1 | <0.001* | 0.6 | 0.7 | 0.424 |
| 4. Enjoy peek-a-boo | 1.7 | 0.4 | <0.001* | −2.0 | 0.6 | 0.001* | −0.3 | 0.1 | 0.032 | 0.4 | 1.2 | 0.735 |
| 5. Pretend play | 2.5 | 0.2 | <0.001* | 0.6 | 0.3 | 0.044 | −0.9 | 0.1 | <0.001* | 0.5 | 0.6 | 0.412 |
| 6. Point to ask for something | 2.1 | 0.2 | <0.001* | 0.2 | 0.3 | 0.509 | −0.1 | 0.1 | 0.307 | −0.7 | 0.7 | 0.298 |
| **7. Point to indicate interest** | 2.2 | 0.2 | <0.001* | 0.2 | 0.3 | 0.468 | −0.01 | 0.1 | 0.802 | −1.5 | 0.8 | 0.062 |
| 8. Functional play with objects | 2.5 | 0.2 | <0.001* | −0.6 | 0.5 | 0.195 | 0.1 | 0.01 | 0.136 | 0.9 | 0.7 | 0.202 |
| **9. Show objects to others** | 2.7 | 0.2 | <0.001* | 0.5 | 0.4 | 0.187 | −0.3 | 0.01 | <0.001* | 0.6 | 0.7 | 0.384 |
| 10. Eye contact | 1.1 | 0.3 | <0.001* | −2.7 | 0.5 | <0.001* | −0.4 | 0.01 | <0.001* | −0.6 | 0.9 | 0.526 |
| 11. Oversensitive to noise | 0.0 | 0.2 | 0.980 | −3.1 | 0.4 | <0.001* | −0.05 | 0.02 | 0.050 | −0.0 | 0.6 | 0.953 |
| 12. Response to smile | 2.2 | 0.6 | <0.001* | −1.6 | 0.8 | 0.034 | 0.08 | 0.2 | 0.691 | −0.8 | 1.5 | 0.584 |
| **13. Imitation** | 1.3 | 0.2 | <0.001* | −0.7 | 0.307 | 0.018 | −0.5 | 0.04 | <0.001* | 1.1 | 0.5 | 0.036* |
| **14. Response to name** | 2.5 | 2.5 | <0.001* | −0.7 | 0.5 | 0.194 | −0.2 | 0.1 | 0.028 | −0.2 | 0.9 | 0.791 |
| **15. Follow to point** | 2.1 | 0.3 | <0.001* | 0.3 | 0.3 | 0.331 | −0.4 | 0.1 | <0.001* | −2.4 | 0.9 | 0.011* |
| 16. Walk unaided | 2.3 | 0.2 | <0.001* | 0.7 | 0.3 | 0.010 | 0.01 | 0.1 | 0.766 | 0.5 | 0.5 | 0.362 |
| 17. Look at things you are looking at | 2.0 | 0.2 | <0.001* | −1.3 | 0.5 | 0.006 | −0.2 | 0.1 | 0.003 | 0.3 | 0.7 | 0.701 |
| 18. Unusual finger/hand movements | 0.6 | 0.2 | 0.002* | −2.0 | 0.4 | <0.001* | 0.04 | 0.03 | 0.202 | −0.2 | 0.6 | 0.779 |
| 19. Attract attention to play/activity | 1.7 | 0.2 | <0.001* | −1.0 | 0.4 | 0.014 | −0.1 | 0.1 | 0.052 | 0.2 | 0.7 | 0.798 |
| 20. Suspected deafness | 1.4 | 0.2 | <0.001* | −2.0 | 0.5 | <0.001* | −0.3 | 0.1 | <0.001* | 0.4 | 0.8 | 0.660 |
| 21. Understand speech | 2.4 | 0.3 | <0.001* | −0.7 | 0.5 | 0.112 | −0.4 | 0.1 | <0.001* | 0.6 | 0.8 | 0.459 |
| 22. Stare at nothing | 0.8 | 0.2 | <0.001* | −0.9 | 0.3 | <0.001* | 0.1 | 0.02 | 0.080 | 0.0 | 0.5 | 0.956 |
| 23. Check parents reaction | 0.5 | 0.2 | 0.004 | −1.9 | 0.3 | <0.001* | −0.1 | 0.03 | 0.001* | −0.3 | 0.6 | 0.659 |

β values: positive beta values indicate a higher likelihood of failure for ASD-females. Negative values express lower likelihood of failure

*Significant, [α]$p < 0.00217$, bold item refer to Robins et al. (2001) six critical items

[a]Main effect of diagnosis without controlling for #item failed on M-CHAT

[b]Main effect of diagnosis shown after controlling for #item failed on M-CHAT

[c]Main effect of sex in non-ASD shown after controlling for #item failed on M-CHAT

[d]Main effect of sex in ASD shown after controlling for #item failed on M-CHAT

expressions to the possible exclusion of other forms of early imitation. Facial imitation represents a basic form of imitation emerging in early infancy (Meltzoff and Moore 1983), whereas later forms of imitation are more socially nuanced and complex.

Depending on how the imitation question is perceived and evaluated by parents, *follow to point*, a strength for ASD females, may in some instances encompass a relatively more complex social ability (Woodward and Guajardo 2002), when compared to very early imitation of facial expressions. This complexity arises from the need to understand the cue or dyadic bid from another person, and follow their direction to focus on a point of joint attention. Our findings concerning *follow to point* might suggest that females possess a strength over males on this socially oriented parent endorsed behavior even though the child received an ASD diagnosis at some point. In contrast to Mundy et al. (2007), Harrop et al. (2015) did not find significant differences in joint attention between males and females with an ASD diagnosis.

Another issue might be that caregivers interpret the various items differently for males and females. For example, excessive correct use of a toy car in males might not be screened as a failure of functional play with objects, though it could be a circumscribed behavior or a stereotyped replication of a movement. For females, parents might endorse the presence of imitative play until a point where severity in autism symptoms make severe impairment more salient. Another possibility is that for females, imitative play and pretend play situations may contain a higher level of complexity as development progresses. Thus, the difficulties in development that comes with an ASD diagnosis may be more impairing in female play than in male play situations. Another item that could be misinterpreted is item 16 (*Walking unaided*). Parents of children with bolting issues, often seen in children with ASD, could interpret the need of supervision to avoid bolting, as a failure. Such interpretation may not be limited to assessment of motor skills. It is important to keep in mind that the M-CHAT, without the follow-up interview, does not present the parents with exemplifying situations.

## Limitations

The ABC study (Stoltenberg et al. 2010) is a prospective study, which was terminated in 2009. Diagnoses are still being registered, and children may receive an ASD diagnosis later than 3, 5 or 7 years of age. Thus, it is likely that the high male to female ratio found in our study may also be due to the relatively young age of the cohort included in our study. There is a possibility that some of these participants receive a diagnosis later. Another limitation that the person

specific identifiable diagnosis registered in the NPR can only utilized if diagnosed after 2008. Children diagnosed earlier than 2008 and not being seen by the specialized services, might have an ASD diagnosis without our knowledge. This also provides an issue in determining age at diagnosis and level of functioning, and would not be reliable with the current data. This sample do not represent the prevalence of ASD in Norway, which is approximately 1% (Surén et al. 2012), but the current dataset only has 185 fully completed M-CHAT questionnaires of children that went on to receive an ASD diagnosis. Nevertheless, this remains a limitation of the present study. Furthermore, being a large population sample, certain characteristics have been found to be over- or under-represented due to self-selection (Nilsen et al. 2009). The low number of females diagnosed ($n = 32$) might affect power in multiple comparisons. Because of a lack of clear examples on the M-CHAT, there is room for individual interpretation, e.g. parents of males might interpret the context of items differently than parents of females.

## Conclusion

The present study suggests that sex differences in ASD symptoms are present between 17 and 30 months of age in children who have not received an ASD diagnosis, both in terms of differences on number of failed items and on M-CHAT items. Non-ASD females develop certain behaviors and skills earlier than non-ASD males. Results suggest a nuanced view of the "extreme male brain theory of autism". At an item level, almost every male versus female disadvantage in the broader population was consistent with M-CHAT vulnerabilities in ASD. However, controlling for total M-CHAT failures, this male disadvantage was more equivocal and many classically ASD-associated features were found more common in non-ASD. Within ASD, females showed relative strengths in joint attention, but impairments in imitation. Further research is needed to disentangle sex differences in ASD symptoms at 18 months, taking into account children's language level and intellectual impairment. It is important move forward to understand how the presence of key developmental milestones moderates the development of autistic-like behavior at 18 months in children who later develop ASD and in children with other developmental problems.

### Compliance with Ethical Standards

## References

American Academy of Pediatrics; Council on Children with Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children with Special Needs Project Advisory Committee. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. Pediatrics, 118(1), 405–420.

Asperger, H. (1944). Die ''Autistischen Psychopathen'' im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten, 117*(1), 76–136.

Auyeung, B., Lombardo, M., Knickmeyer, R., & Baron-Cohen, S. (2013). Extreme male brain (EMB) theory. *Encyclopedia of Autism Spectrum Disorders*, 1193–1202.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248–254.

Baron-Cohen, S. (2009). Autism: The empathizing–systemizing (E–S) theory. *Annals of the New York Academy of Sciences, 1156*(1), 68–80.

Baron-Cohen, S., & Benenson, J. F. (2003). Books and arts-essential difference: Men, women and the extreme male brain/essential difference: The truth about the male and female brain. *Nature, 424*(6945), 132–132.

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: Implications for explaining autism. *Science, 310*(5749), 819–823.

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*. doi:10.1177/1362361310391116.

Bouchard, C., Trudeau, N., Sutton, A., Boudreault, M. C., & Deneault, J. (2009). Gender differences in language development in French Canadian children between 8 and 30 months of age. *Applied Psycholinguistics, 30*(04), 685–707.

Boyle CA, Boulet S, Schieve L, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, Visser S, Kogan MD. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics, 127*, 1034–1042.

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.

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior and Development, 23*(1), 113–118.

Davis, M. H. (1994). *Empathy: A social psychological approach*. Westview Press.

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

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders, 33*(4), 365–382.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18*(4), 281–294.

Fombonne, E. (2007). Epidemiological surveys of pervasive developmental disorders. In F. R. Volkmar (Ed.), Autism and pervasive developmental disorders (2nd ed., pp. 33–68). New York, NY: Cambridge University Press.

Gilman, S. R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M., & Vitkup, D. (2011). Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron, 70*(5), 898–907.

Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist, 52*(10), 1091.

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*(3), 843–855.

Harrop, C., Shire, S., Gulsrud, A., Chang, Y. C., Ishijima, E., Lawton, K., & Kasari, C. (2015). Does gender influence core deficits in ASD? An investigation into social-communication and play of girls and boys with ASD. *Journal of Autism and Developmental Disorders, 45*(3), 766–777.

Hattier, M. A., Matson, J. L., Tureck, K., & Horovitz, M. (2011). The effects of gender and age on repetitive and/or restricted behaviors and interests in adults with autism spectrum disorders and intellectual disability. *Research in Developmental Disabilities, 32*(6), 2346–2351.

Hittelman, J. H., & Dickes, R. (1979). Sex differences in neonatal eye contact time. *Merrill-Palmer Quarterly of Behavior and Development, 25*, 171–184.

Ibanez, L. V., Stone, W. L. & Coonrod, E. E. (2014). Screening for autism in young children. In *Handbook of autism and pervasive developmental disorders* (4th ed., Vol. 2. pp. 585–608). Hoboken, NJ: Wiley.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., et al. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748–751.

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

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), 2875–2888.

Lai, D. C., Tseng, Y. C., Hou, Y. M., & Guo, H. R. (2012). Gender and geographic differences in the prevalence of autism spectrum disorders in children: Analysis of data from the national disability registry of Taiwan. *Research in Developmental Disabilities, 33*(3), 909–915.

Levy, D., Ronemus, M., Yamrom, B., Lee, Y. H., Leotta, A., Kendall, J., et al. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron, 70*(5), 886–897.

Lutchmaya, S., & Baron-Cohen, S. (2002). Human sex differences in social and non-social looking preferences, at 12 months of age. *Infant Behavior and Development, 25*(3), 319–325.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences* (Vol. 1). Stanford: Stanford University Press.

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjærven, R., Stoltenberg, C., & MoBa Study Group. (2006). Cohort profile: The Norwegian mother and child cohort study (MoBa). *International Journal of Epidemiology, 35*(5), 1146–1150.

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders, 42*(7), 1304–1313.

Mayes, S. D., & Calhoun, S. L. (2011). Impact of IQ, age, SES, gender, and race on autistic symptoms. *Research in Autism Spectrum Disorders, 5*(2), 749–757.

McCarthy, P., Fitzgerald, M., & Smith, M. A. (1984). Prevalence of childhood autism in Ireland. *Irish Medical Journal, 77*(5), 129–130.

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.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development, 54*, 702–709.

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

Montagne, B., Kessels, R. P., Frigerio, E., de Haan, E. H., & Perrett, D. I. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive Processing, 6*(2), 136–141.

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development, 78*(3), 938–954.

Nilsen, R. M., Vollset, S. E., Gjessing, H. K., Skjærven, R., Melve, K. K., Schreuder, P., et al. (2009). Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatric and Perinatal Epidemiology, 23*(6), 597–608.

Reilly, S., Bavin, E. L., Bretherton, L., Conway, L., Eadie, P., Cini, E., et al. (2009). The Early Language in Victoria Study (ELVS): A prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months. *International Journal of Speech and Language Pathology, 11*(5), 344–357.

Riddoch, C. J., Mattocks, C., Deere, K., Saunders, J., Kirkby, J., Tilling, K., et al. (2007). Objective measurement of levels and patterns of physical activity. *Archives of Disease in Childhood, 92*(11), 963–969.

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics, 133*(1), 37–45.

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131–144.

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

Salomone, E., Charman, T., McConachie, H., & Warreyn, P. (2016). Child's verbal ability and gender are associated with age at diagnosis in a sample of young children with ASD in Europe. *Child: Care, Health and Development, 42*(1), 141–145.

Skuse, D. H. (1997). Genetic factors in the etiology of child psychiatric disorders. *Current Opinion in Pediatrics, 9*(4), 360.

Skuse, D. H. (2000). Imprinting, the X-chromosome, and the male brain: Explaining sex differences in the liability to autism. *Pediatric Research, 47*(1), 9–9.

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

Stoltenberg, C., Schjølberg, S., Bresnahan, M., Hornig, M., Hirtz, D., Dahl, C., et al. (2010). The Autism Birth cohort: A paradigm for gene-environment-timing research. *Molecular Psychiatry, 15*, 676–680.

Surén, P., Bakken, I. J., Aase, H., Chin, R., Gunnes, N., Lie, K. K., et al. (2012). Autism spectrum disorder, ADHD, epilepsy, and cerebral palsy in Norwegian children. *Pediatrics, 130*(1), e152–e158.

Szatmari, P., Liu, X. Q., Goldberg, J., et al. (2012). Sex differences in repetitive stereotyped behaviors in autism: Implications for genetic liability. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics, 159B*(1), 5–12.

The Norwegian Mother and Child Cohort Study, End of enrolment—Protocol II, Revised October 2012 http://www.fhi.no/moba-en.

Trouton, A., Spinath, F. M., & Plomin, R. (2002). Twins early development study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Research, 5*(05), 444–448.

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development, 17*(1), 1061–1084.

Zahn-Waxler, C., Crick, N. R., Shirtcliff, E. A., & Woods, K. E. (2006). The origins and development of psychopathology in females and males. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Theory and method* (2nd ed., Vol. 1, pp. 76–138). Hoboken, NJ: Wiley.

Zambrana, I. M., Ystrom, E., & Pons, F. (2012). Impact of gender, maternal education, and birth order on the development of language comprehension: A longitudinal study from 18 to 36 months of age. *Journal of Developmental & Behavioral Pediatrics, 33*(2), 146–155.

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