Testing the Extreme Male Brain Hypothesis: Is Autism Spectrum Disorder Associated with a More Male-Typical Brain?

Testing the Extreme Male Brain Hypothesis

Liza van Eijk[a,b,c,d], Brendan P. Zietsch[c]

[a] Department of Psychology, College of Healthcare Sciences, Division of Tropical Health and Medicine, James Cook University, Douglas, Australia

[b] The Australian e-Health Research Centre, CSIRO, Herston, Australia

[c] Centre for Psychology and Evolution, School of Psychology, University of Queensland, St Lucia, Australia

[d] Queensland Brain Institute, University of Queensland, St Lucia, Australia

Corresponding author:

Liza van Eijk, Department of Psychology, College of Healthcare Sciences, Division of Tropical Health and Medicine, James Cook University, Douglas, 4811, +61 7 4781 5823, liza.vaneijk@jcu.edu.au

Number of text pages: 22

Number of tables: 1

Number of Figures: 2

This study is supported by the following grants: NIH; EB020403, 1U54MH091657. NHMRC; 1009064, 496682. NIMH; K23MH087770, R03MH096321, 5R21MH107045. ZonMw; 849200011.

Acknowledgments

QTIM was funded by the National Institutes of Health (NIH) (project HD050735; Award U54 EB020403, subaward no. 56929223) and the National Health and Medical Research Council (1009064, 496682). We thank the twins and siblings for their participation, Marlene Grace and Ann Eldridge for twin recruitment, the radiographers for scanning, and Kerrie McAloney and Daniel Park for research support. HCP data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. For ABIDE-I, primary support for the work by Adriana Di Martino was provided by the National Institute of Mental Health (NIMH) (K23MH087770) and the Leon Levy Foundation. Primary support for the work by Michael P. Milham and the INDI team was provided by gifts from Joseph P. Healy and the Stavros Niarchos Foundation to the Child Mind Institute, as well as by an NIMH award to MPM (R03MH096321). For ABIDE-II, primary support for the work by Adriana Di Martino and her team was provided by the National Institute of Mental Health (NIMH 5R21MH107045). Primary support for the work by Michael P. Milham and his team provided by the National Institute of Mental Health (NIMH 5R21MH107045); Nathan S. Kline Institute of Psychiatric Research). Additional Support was provided by gifts from Joseph P. Healey, Phyllis Green and Randolph Cowen to the Child Mind Institute.

L. van Eijk was supported in part by the Australian Government Research Training Program Scholarship and the Imaging Genomics Laboratory, Queensland Brain Institute, as well as by the Hendrik Muller Vaderlandsch Fund, the Vrijvrouwe van Renswoude Fund, and the Prince Bernhard Culture Fund. B.P. Zietsch is also supported in part by ZonMw grant

849200011 from The Netherlands Organisation for Health Research and Development. We thank Dr Alex Pagnozzi for careful reading of the manuscript.

Lay abstract

A popular theory proposes that individuals with Autistic Spectrum Disorder (ASD) have an "extreme male brain", but this has not been subject to rigorous, direct tests. We developed a measure of individual differences along a male-female dimension and then derived this measure for 1,060 individuals with ASD and 1,166 neurotypical controls. Individuals with ASD had slightly more male-type brains. However, this difference is accounted for by males and individuals with ASD having relatively larger brains than females and controls, respectively.

Abstract

Autism Spectrum Disorder (ASD) is more common in males than females and has been linked to male-typical behavior. Accordingly, the 'Extreme Male Brain' hypothesis suggests that ASD is associated with an exaggeratedly male-typical brain. To test this hypothesis, we derived a data-driven measure of individual differences along a male-female dimension based on sex differences in subcortical brain shape (i.e. brain maleness) by training our algorithm on two population samples (Queensland Twin Imaging study and Human Connectome Project; combined N=2,153). We then applied this algorithm to two clinical datasets (Autism Brain Imaging Data Exchange I and II; ASD N=1,060; neurotypical controls N=1,166) to obtain a brain maleness score for each individual, representing maleness of their brain on a male-female continuum. Consistent with the Extreme Male Brain hypothesis, we found a higher mean brain maleness score in the ASD group than in controls (d=0.20 (0.12-0.29)), parallel to higher scores for control males than control females (d=1.17 (1.05-1.29)). Further, brain maleness was positively associated with autistic symptoms. We tested the possibility this finding was driven by the ASD group's larger brains than controls (d= 0.17 (0.08-0.25)), given that males had larger brains than females (d= 0.96 (0.84-1.07)). Indeed, after adjusting for differences in brain size, the brain maleness difference between the ASD group and controls disappeared, and no association with autistic symptoms remained (after controlling for multiple comparisons), suggesting greater maleness of the autistic brain is driven by brain size. Brain maleness may be influenced by the same factors that influence brain size.

Keywords: Masculinity, Autism Spectrum Disorder, Magnetic Resonance Imaging, Brain Diseases, Neuroimaging, Sex Characteristics

Testing the Extreme Male Brain Hypothesis: Is Autism Spectrum Disorder Associated with a More Male-Typical Brain?

Autism Spectrum Disorder (ASD) is four times more common in males than females (Lord et al., 2020), and some autistic traits have been associated with male-typical behavior (Baron-Cohen, Wheelwright, et al., 2005; Baron-Cohen, 2009). For example, men score on average lower on empathy tasks than women, while adults with ASD, irrespective of their sex, score the lowest (Baron-Cohen, Wheelwright, et al., 2005; Baron-Cohen, 2009). Vice versa, men score on average higher than women on an attention-to-detail task, while individuals with ASD score the highest (Baron-Cohen, Wheelwright, et al., 2005; Baron-Cohen, 2009). These behavioral findings suggest that individuals with ASD are more shifted towards the exaggerated male-type (Baron-Cohen, Wheelwright, et al., 2005; Baron-Cohen, 2009), raising the question whether ASD is associated with a more male-typical brain.

A recent large behavioral study (N= 671,606; aged 16 to 89 yrs; 61% female) (Greenberg, Warrier, Allison, & Baron-Cohen, 2018) claimed to have found evidence for the 'Extreme Male Brain' hypothesis. They used a classification system proposed by Baron-Cohen (2005), which classifies individuals based on their behavior on two dimensions. The first dimension, empathizing, includes understanding thoughts and feelings and responding with an appropriate emotion (Baron-Cohen, Knickmeyer, et al., 2005; Baron-Cohen, Wheelwright, et al., 2005), while the second, systemizing, is the drive to identify rules within a system, and to analyze and predict accordingly (Baron-Cohen, Knickmeyer, et al., 2005; Baron-Cohen, Wheelwright, et al., 2005). Using these two dimensions, this recent study (Greenberg et al., 2018) showed that more females than males were classified as type E (better at empathizing than systemizing), and more males than females were classified as type

S (better at systemizing than empathizing), while the majority of males and females with autistic symptoms was classified as type S. Although these findings provide some evidence for an exaggerated male type in ASD, findings were based on behavioral questionnaires and did not include brain data.

In line with the idea of ASD being associated with an exaggerated male-typical brain type, brain structure differences between individuals with ASD and controls have been found (e.g. van Rooij et al., 2018) and it has been suggested that these brain differences may be associated with sex differences in the brain. For example, some studies (Freitag et al., 2009; Sparks et al., 2002; Stanfield et al., 2008) have reported larger brain volumes for individuals with ASD versus controls, in particular in younger children, which parallels the finding that male brains are about 10-15% larger than female brains (Goldstein et al., 2001; Lenroot et al., 2007; Rabinowicz, Dean, Petetot, & de Courten-Myers, 1999; Ruigrok et al., 2014). Also, individuals with ASD were found to have less white matter in the corpus callosum compared to controls (Radua, Via, Catani, & Mataix-Cols, 2011), which parallels the smaller corpus callosum volumes observed in males compared to females (Shiino et al., 2017), as well as the more interhemispheric connections in female brains compared to male brains (Ingalhalikar et al., 2014). Moreover, a recent large study (Postema et al., 2019) (N=3,583) reported brain asymmetry differences for ASD, with six out of nine cortical regions showing decreased leftward asymmetry while the other three showed increased left asymmetry or more prominent right (versus left) decreases in asymmetry. This is parallel to the sex differences reported for lateralization, including more leftward and less rightward asymmetry in cortical thickness of the medial temporal brain regions in males compared to females (Kong et al., 2018) (N=17,141). However, these findings are isolated observations, prone to reporting or publication biases; what is needed is a more systematic exploration linking whole-brain sex differences with brain structure differences in ASD versus neurotypical individuals.

Evidence from studies directly linking brain sex differences and brain differences in ASD is inconclusive. An earlier small study (Lai et al., 2013) (N=120) found overlap between brain regions associated with ASD diagnosis and brain areas showing sex differences (within controls) – although evidence for the Extreme Male Brain hypothesis was only found within females but not within males. Similarly, a recent study (Smith et al., 2019) (N=167) examining functional connectivity differences in resting-state functional Magnetic Resonance Imaging (MRI) also found partial evidence for the Extreme Male Brain hypothesis by showing that cortico-cerebellar functional connectivity in ASD females was more similar to typical males than typical females. However, functional connectivity for ASD males fell between control males and control females. The first and only study to test the Extreme Male Brain hypothesis more directly (Ecker et al., 2017) (N=196) showed no evidence for the hypothesis. Individuals within the same sex vary in their genetic predispositions as well as in exposure and sensitivity to gonadal hormones, so that some men will develop a more male-typical brain while other men develop a more female-typical brain (and similarly for women). Ecker et al. (2017) found no difference in ASD probability based on the degree of maleness of the brain – a measure derived by predicting sex based on cortical thickness. It is difficult to draw firm conclusions from these various findings, given the limited sample sizes (ranging from 120 to 196 participants) that risk false negatives (due to lack of power) as well as false positives (in combination with publication bias toward significant results, which tend to show larger effects in smaller samples).

Here, we test the Extreme Male Brain hypothesis in two large imaging samples, combining the Autism Brain Imaging Data Exchange I and II datasets (N=2,226). We focus on subcortical regions to create a global brain maleness score, as cortical regions showed too much measurement error and we previously showed a similar prediction of sex for subcortical versus cortical data points (van Eijk et al., in press). We apply our recently

developed data-derived measure (van Eijk et al., in press) of individual differences on a male-female dimension based on sex differences in subcortical brain shape (i.e. brain maleness) to test whether the ASD group have a more exaggeratedly male-typical brain type (i.e. higher brain maleness scores than controls), while also examining the association between brain maleness and autistic symptoms.

Methods

Participants

This study used two independent population sample imaging datasets to train the prediction model in order to derive the brain maleness measure for the two clinical samples (van Eijk et al., in press) (**Fig. 1**). The first population dataset consists of 1,040 twins as part of the Queensland Twin Imaging (QTIM) study (ages 15 to 30 years, mean age of 22.42 ±3.33 years, 64.81% female). Individuals with developmental, neurological or psychiatric disorders, impaired intellectual functioning, or head trauma were excluded. Only right-handed twins were included in the study. All individuals gave written informed consent. Ethics approval was given by the Human Research Ethics Committees of the QIMR Berghofer Medical Research Institute, University of Queensland, and UnitingCare Health. The second normal population sample was provided as part of the Human Connectome Project (HCP) (Van Essen et al., 2012), and comprised 1,113 individuals (ages 22 to 37 years, mean=28.80 ±3.70 years, 54.40% female). Individuals with severe neurodevelopmental disorders, documented neuropsychiatric disorders, neurologic disorders, diabetes, high blood pressure, or those born premature were excluded. All individuals gave written informed consent. Ethics approval was given by the institutional review board.

The two clinical datasets were provided by the Autism Brain Imaging Data Exchange consortium (ABIDE-I (Di Martino et al., 2014) and ABIDE-II (Di Martino et al., 2017)). The

ABIDE-I dataset includes 1,112 individuals (539 individuals with ASD and 573 controls), ages 7 to 64 years (mean age= 17.04 ±8.03, 85.16% male). The ABIDE-II dataset consists of 1,114 individuals (521 individuals with ASD and 593 controls), ages 5 to 64 years (mean age=14.86 ±9.16, 76.84% male). All individuals gave informed consent in line with the human research boards at each the participating institution.

Image Acquisition

For the QTIM dataset, structural MRI scans were obtained at 4-Tesla (Siemens Bruker), acquiring a 3D structural T1-weighted image (T1/TR/TE = 700/1500/3.35 ms; flip angle = 8°, voxel size = 0.9375 × 0.9375 × 0.90 mm^3). 81% with a coronal acquisition, 19% with a sagittal acquisition. For the HCP dataset, structural MRI scans were obtained at 3-Tesla (Siemens Connectome Skyra), acquiring a 3D structural T1-weighted image (T1/TR/TE = 1000/2400/2.14 ms; flip angle = 8°, slice thickness = 0.70 mm, voxel size = 0.70 x 0.70 x 0.70 mm^3) (Van Essen et al., 2012).

For the ABIDE-I sample, datasets of 17 different sites were aggregated after collection, which resulted in different imaging acquisition protocols per site. We used the structural T1-weighted images, which were all acquired at 3-Tesla scanners with a 1 mm^3 isotropic resolution (Haar, Dinstein, Berman, & Behrmann, 2014). More details of the scanning acquisition for each site can be found at http://fcon_1000.projects.nitrc.org/indi/abide/. For the ABIDE-II sample, data was acquired across 16 sites (Di Martino et al., 2017) and aggregated after collection. All structural T1-weighted images were acquired on a 3-Tesla scanner, except for one site (at a 1.5 Tesla scanner). For more information of the scanning protocol at the different sites, see http://fcon_1000.projects.nitrc.org/indi/abide/. For the analyses, a cohort variable

(categorical) was regressed out to adjust for differences in acquisition parameters for each data collection, using multiple dummy variables.

Image Processing

All structural scans were preprocessed to remove signal inhomogeneity using the software program Statistical Parametric Mapping (SPM) (Frackowiak, 1997; Friston et al., 1995) version 12 software package in Matlab version R2018a. Note, images were not registered to common template space to avoid distortions in the shape of the brain structures.

Obtaining the Brain Maleness Measure

The Landmarks

Using SPM's function 'normalize', we placed our recently developed 467 subcortical landmarks per hemisphere (934 in total) in seven subcortical regions, including the amygdala, caudate nucleus, hippocampus, lateral ventricle, pallidum, putamen, and thalamus (**Fig. 1**). For more details, see van Eijk et al. (in press). Landmarks were only placed in regions large enough for multiple landmarks, as it was the aim to capture the shape of each structure with multiple landmarks. Of the total 2,226 scans, 166 were removed from analyses due to processing errors as the result of insufficient coverage of the brain and poor image quality. After placing the landmarks for each individual, we extracted the 3D coordinates for each of the 934 landmarks. Only occasionally, landmarks were not transformed to native space due to the non-linear nature of the transformation, resulting in missing data. Missing data (0.15% of the total data points; i.e. 934 landmarks multiplied by 2,060 participants) were imputed with R statistics package 'Geomorph' TPG option (Adams & Otarola-Castillo, 2013), as the next analyses required no missing data.

Then, using the R statistics package 'Shapes' (Dryden, 2016), we brought landmark coordinates from each individual into standard space by applying a Generalized Procrustes Analysis, which removes variation in size, position, orientation, and rotation of brain images (**Fig. 1**), while preserving brain shape. During this process, a Principal Component Analysis was also performed, rotating the data into uncorrelated components (for more details see van Eijk et al., in press). We performed this analysis twice: with or without scaling all brains to the same brain size during the Procrustes Analysis – equivalent to adjusting versus not adjusting for differences in brain size. This analysis also yielded a measure of brain size, i.e. centroid size, which reflects the square root of the sum of the squared distances of all the landmarks from their centroid (Klingenberg, 2016).

Model Prediction

Next, we included the first 50 principal components of the shape variables (those with an eigenvalue of one or larger, explaining 91.03% of the variance) in a model to predict the biological sex of the participants with a Linear Discriminant Analysis, using the package 'MASS' (Venables, 2002) in R statistics version 3.4.4 (for more details see van Eijk et al., in press) (**Fig. 1**). First, we trained our prediction model on two population samples (QTIM and HCP), to derive the linear combination of the shape variables which best discriminated males from females. Then, we applied our prediction model to the clinical ABIDE datasets, assigning each individual a brain maleness score which reflects the position of their brain shape along this male-female continuum. This method to obtain a measure of brain maleness derived from subcortical brain shape has shown excellent test-retest reliability (QTIM r=0.955; HCP r=1.000) and good validity (unadjusted for brain size: AUC=94.81-95.30%; adjusted for brain size (with the Procrustes size adjustment): AUC=85.69-87.01%) (van Eijk et al., in press).

To make sure that our measure of brain maleness measured sex differences in brain shape in the clinical samples, we tested its validity by calculating the area under the curve (AUC), the true positive rate against the false positive rate, and its 95% confidence interval (DeLong) using the 'pROC' package (Robin et al., 2011). The AUC is, unlike accuracy, insensitive to class imbalance (Fawcett, 2006). As well as examining the validity for the full sample, we also tested the robustness of the prediction by examining the AUC for the following subsamples: 1) a sex-balanced subsample (391 male, 391 female), 2) controls (831 male, 263 female), and 3) individuals with ASD (838 male, 128 female). As the total dataset included 391 females and 1,669 males, for the sex-balanced subsample we included all 391 females and randomly selected 391 of 1,669 males using the R Statistics function 'Sample' of the R Statistics 'Dplyr' package.

Statistical Analyses

We used t-tests and Cohen's *d* to compare brain size and brain maleness in males versus females (after regressing out age, diagnosis, and a cohort variable) and in the ASD versus control group (after regressing out sex, age, and a cohort variable). In addition, we also adjusted for an interaction effect of sex by age for brain size and brain maleness (unadjusted for brain size) after finding a sex by age interaction effect on brain size (Beta (SD)=0.014(0.006), t=2.48, p=0.013) (Supplementary Fig. 2B) and similarly on brain maleness (not adjusted for brain size) (Beta (SD)=0.014(0.005), t=2.83, p=0.005). This interaction was not found for brain maleness adjusted for brain size (with Procrustes size-adjustment or Procrustes size-adjustment plus regression for brain size). No evidence was found for interaction effect for age by diagnosis on brain size and brain maleness.

Further, within the ASD group, we tested for an association between brain maleness scores and autistic symptoms by calculating Pearson's correlation coefficient, using the

following behavioral measures: the Autism Diagnostic Interview-Revised (ADI), the Autism Diagnostic Observation Schedule Module (ADOS), and the Social Responsiveness Scale (SRS). Multiple-comparison correction was applied using the Bonferroni method, with an adjusted significance level of $p \leq 0.0028$ (0.05/18).

Results

Validation: Prediction of Sex

Predicting sex in the clinical dataset (81.02% male, N=2,060), after training the model on two normal population samples (QTIM + HCP; 59.45% females, N=2,153), resulted in an AUC of 84.06% when not adjusting for brain size in the Procrustes Analysis, and 78.78% with the Procrustes size adjustment (**Table 1**). Irrespective of adjusting for brain size, the distribution of the brain maleness scores showed overlapping normal distributions (**Fig. 2**), with a higher mean score in males than females (unadjusted for brain size but adjusted for age, diagnosis, and cohort: t=22.34, $p<0.001$, d=1.17 (95% CI: 1.05-1.29); with Procrustes size adjustment: t=16.53, $p<0.001$, d=0.92 (95% CI: 0.80-1.03)). To test the robustness of the prediction, we showed a similar prediction performance in the three subsamples (a sex-balanced subset, controls, and ASD group) compared to the prediction in the full sample (**Table 1**). To summarize, we validated the brain maleness measure by predicting sex accurately in the clinical datasets. For all further analyses, outliers in the brain maleness scores (z-scores ± 3.29) were winsorized within each sex.

Brain Maleness Scores

First, we found that males had a larger brain size than females (controlled for age, diagnosis and cohort: t=18.50, $p<0.001$, d=0.96 (95% CI: 0.84-1.07); Supplementary Fig. 1). In the total sample (including both males and females), we also found that brain size was

slightly larger in individuals with ASD compared to controls (controlled for age, sex, cohort, and interaction effect sex by age: t=3.71, p<0.001, d=0.17 (95% CI: 0.08-0.25)) (Supplementary Fig. 2A). Looking within each sex, this difference was also found within males with, on average, ASD males having larger brains than control males (controlled for age and cohort: t=3.35, p<0.001, d=0.16 (95% CI: 0.07-0.26); Supplementary Fig. 1); this size difference was not found within females (controlled for age and cohort: t=0.06, p=0.953, d=0.01 (95% CI: -0.21-0.22); Supplementary Fig. 1). These different findings for each sex are line with the interaction effect found for covariates sex and age on brain size (Supplementary Fig. 2B). Within males, the ASD group has, on average, larger brains than controls across a wide age range (Supplementary Fig. 2B; right panel). Within females, however, the ASD group only shows, on average, larger brains than controls during early young adulthood, while this difference appears to reverse after the age of twenty, showing on average larger brains for female controls than the female ASD group (Supplementary Fig. 2B; left panel). This effect should be interpreted with caution though due to the small number of data points for females older than twenty.

In the full sample, the distribution of the brain maleness scores showed on average a slightly higher score for the ASD group compared to controls (unadjusted for brain size but controlled for age and cohort) (**Fig. 2 left panel)**. This finding was confirmed with a t-test (unadjusted for brain size but controlled for age, sex, cohort, and interaction sex by age): t=4.61, p<0.001, d=0.20 (95% CI: 0.12-0.29)), and likewise within each sex (controlled for age and cohort), though the effect was only significant within males (within females: t=0.31, p = 0.757, d=0.04 (95% CI: -0.18-0.25); within males: t=4.05, p<0.001, d=0.20 (95% CI=0.10-0.29)). These findings are in line with the interaction effect found for sex by age on brain maleness, which is similar to the interaction effect observed for brain size.

Brain maleness scores derived from brains scaled to the same size (with Procrustes size adjustment) no longer showed a difference between the ASD and control group (adjusted for sex, age, and cohort: t=-1.113, p=0.266, d=0.05 (95% CI: -0.04-0.14)), and likewise no group difference was observed within each sex (controlled for age and cohort, **Fig. 2 middle panel**; within females: t=-0.34, p = 0.735, d=-0.04 (95% CI: -0.25-0.17); within males: t=0.822, p =0.411, d=0.04 (95% CI=-0.06-0.14)). These scores (with Procrustes size adjustment) still showed an association with brain size (after adjusting for sex, age, and cohort) (Pearson's r=0.221 (95% CI: 0.180-0.262), p<0.001) – which has also been found previously in the two normal population samples (van Eijk et al., in press). This finding is related to allometric scaling (i.e. a structure's shape is not independent of its size), suggesting that the brain shape data used to derive the maleness scores may still have contained shape differences that are associated with the original size differences. To make sure the brain maleness scores adjusted for brain size during the Procrustes Analysis were not driven by differences in brain size, we regressed out brain size from the scores for further analyses that focused on brain maleness adjusted for brain size. These scores (with Procrustes size adjustment plus regression for brain size) showed similar findings as found for the scores with Procrustes size adjustment alone: we found no significant difference in brain maleness between the ASD and control group (after also controlling for the covariates age, sex and cohort; t=0.31, p= 0.757, d=0.01 (95% CI: -0.07-0.10)), and similarly there was no significant difference within each sex (**Fig. 2 right panel**, after controlling for age and cohort; within females: t = -0.37, p =0.715, d=-0.04 (95% CI:-0.25-0.17); within males: t =-0.30, p =0.765, d= -0.01 (95% CI:-0.11-0.08)).

Next, we examined the association between brain maleness and autistic symptoms within the ASD group (controlling for age, sex, cohort, and interaction sex by age, but not brain size). We found several associations between brain maleness and autistic symptoms

(Supplementary Table 1), with all but one in the same direction as hypothesized, although only one association survived Bonferroni correction ($p \leq 0.0028$). In the total sample, brain maleness was positively associated ($0.0028 \leq p \leq 0.05$) with abnormalities in reciprocal social interaction (r=0.096 (95% CI: 0.019-0.172)), abnormalities in verbal communication (r=0.092 (95% CI: 0.015-0.169)), restricted, repetitive, and stereotyped patterns of behavior (r=0.097 (95% CI: 0.020-0.173)), and abnormal development being evident at or before 36 months (r=0.101 (95% CI: 0.022-0.179), which are subscales of the ADI. These associations were found within males as well, but did not reach significance within females (ranging between r=0.192 and r=0.193, between p=0.061 and p=0.063; Supplementary Table 1). The difference in significance of the effect in males and in females should be seen in light of the much smaller sample size of the female (N=391) versus male (N=1,669) sample.

In addition, within males, brain maleness was positively associated with autistic symptoms measured with the ADOS ($p \leq 0.05$) – including the total score (r=0.080 (95% CI: 0.004-0.156)), and the subscale related to stereotyped behaviors and restricted interest (r=0.082 (95% CI 0.003-0.160)). Opposite to what we hypothesized, within males, brain maleness was negatively associated with the SRS total score, reflecting the severity of social deficits (r= -0.068 (95% CI: -0.131; -0.005)) (Supplementary Table 1). Within females, brain maleness was positively associated with autistic symptoms measured with the ADOS version 2 – including the total score (r=0.247 (95% CI:0.029-0.443)) and its subscales related to deficits in social affect (r=0.247 (95% CI: 0.029-0.443)), stereotyped behaviors and restricted interest (r=0.370 (95% CI: 0.163-0.545), p=0.001, surviving the Bonferroni correction), and severity of symptoms (r=0.257 (95% CI: 0.037-0.454) (Supplementary Table 1).

When scaling all brains to the same brain size using the Procrustes size adjustment, only one association remained in the total sample ($p \leq 0.05$) (Supplementary Table 2), between brain maleness and the ADOS subscale related to stereotyped behaviors and restricted

interest in the total sample (r=0.088 (95% CI: 0.002-0.172)), which was also found within females (r=0.293 (95% CI: 0.078-0.482)) but not within males. In addition, within females, brain maleness was positively associated with the ADI subscales related to abnormalities in reciprocal social interaction (r=0.205 (95% CI: 0.003-0.390)), abnormalities in verbal communication (r=0.204 (95% CI: 0.003-0.390)), restricted, repetitive, and stereotyped patterns of behavior (r=0.204 (95% CI: 0.003-0.390)), and abnormal development being evident at or before 36 months (r=0.207 (95% CI: 0.004-0.394)). However, none of these associations survived the Bonferroni correction (p≤0.0028), and no association was found in the total sample as well as within each sex.

These associations between brain maleness and autistic symptoms may still be driven by differences in brain size, even though brains are scaled to the same size when deriving the brain maleness score, as brain maleness scores are positively associated with brain size despite Procrustes size adjustment. Indeed, brain maleness scores adjusted for brain size with Procrustes size adjustment plus regression for brain size showed no associations with autistic symptoms measured on the ADI or the SRS (Supplementary Table 3). A positive association was found (p≤0.05) for the ADOS subscale related to stereotyped behaviors and restricted interest in the total sample (r=0.103 (95% CI: 0.018-0.187)), which was also found within each sex (Supplementary Table 3). However, none of these associations survived correction for multiple testing (Bonferroni corrected, p≤0.0028).

For comparison, associations between brain size and autistic symptoms were found within each sex (after adjusting for age and cohort) (p≤0.05) (Supplementary Table 4). Two associations were found within males: brain size was positively associated with stereotypic behavior of the ADOS scale (r=0.087 (95% CI: 0.008-0.164)), but negatively associated with the SRS total score (r= -0.075 (95% CI: -0.137;-0.012)). Within females, brain size was positively associated with the ADOS subscale related to repetitive behavior (r=0.322 (95%

CI: 0.110-0.506)). However, none of these associations survived the Bonferroni correction, and no association was found in the total sample as well as within each sex.

Discussion

This study tested the Extreme Male Brain hypothesis in two large clinical imaging datasets (ABIDE-I and II, combined N=2,226) by applying our recently developed data-driven measure of brain maleness (i.e. individual differences on a male-female dimension based on sex differences in subcortical brain shape) (van Eijk et al., in press). Our results are partly in line with the hypothesis: ASD diagnosis and symptoms were associated with greater brain maleness, but the effects appeared to be driven by differences in brain size.

More specifically, we found on average higher brain maleness scores in the ASD versus control group (in parallel with higher scores in control males versus females), and brain maleness showed positive associations with autistic symptoms. Note, though, that ASD females had less masculine brains than control males, indicating that it is brain maleness relative to biological sex that is associated with ASD, rather than absolute brain maleness. Next, we examined brain maleness adjusted for brain size, as we found slightly larger brains in the ASD group compared to controls (in parallel with larger male than female brains). After adjusting for brain size, we no longer found differences in brain maleness scores between the ASD and control group, and no association between brain maleness and autistic symptoms survived after controlling for multiple comparison. These findings suggest that brain size may be driving the associations found between brain maleness scores (unadjusted for brain size) and autistic symptoms.

An alternative hypothesis to the Extreme Male Brain hypothesis is the Gender-Incoherence hypothesis (Lai et al., 2017), which hypothesizes that ASD females are masculinized while ASD males are feminized. Before adjusting for brain size, our results are

more in line with the Extreme Male Brain hypothesis than the Gender-Incoherence hypothesis due to our findings of brain masculinization in the ASD group within both sexes (**Fig. 2, left panel)**. After adjusting for brain size, however, we found no differences between the ASD group and controls, which does not provide evidence for either theory. It is possible that adjusting for brain size removes a lot of the variation of interest (i.e. brain maleness) due to the large sex differences in brain size (10-15% larger in males than females). However, subcortical shape adjusted for brain size still showed an accurate prediction of sex (**Table 1**), despite filtering out (most) of the brain size difference, suggesting that the brain maleness score is measuring additional sex brain differences than just differences in brain size. Without adjusting the measure of brain maleness for brain size, results could simply reflect effects of brain size and not brain maleness per se.

Our results somewhat contrast the findings of Ecker et al. (2017), who used a tenfold smaller sample and derived brain maleness from cortical thickness. They found that brain maleness was not associated with a higher risk for ASD, and female and male individuals with ASD displayed cortical thickness patterns similar to female and male controls respectively. In contrast to their findings, we showed some differences in brain maleness scores between the ASD and control group when not adjusting for brain size. However, this discrepancy might partly be explained by the measures used. That is, cortical thickness is less correlated with brain size than most other brain measures (Barnes et al., 2010). Thus their results may be more comparable to our findings adjusted for brain size, showing no difference in brain maleness scores between the ASD and control group, in line with Ecker et al. (2017).

Previous findings of extreme male-typical behavior observed in ASD versus controls (Baron-Cohen, Wheelwright, et al., 2005; Baron-Cohen, 2009) could partly be caused by an exaggerated male-typical brain type (although driven by differences in brain size). However,

other causal possibilities to consider are, for instance, maleness in sex hormone levels or gene expression (Auyeung et al., 2009; Seidlitz et al., 2020) influencing both brain and behavior. A recent large study (Liu, Seidlitz, Blumenthal, Clasen, & Raznahan, 2020) (N=2,096) showed sex differences in cortical and subcortical brain structures in young and older adults, after adjusting for age and total brain size, and showed that these brain sex differences were associated with expression of sex-chromosome genes. Sex differences may be the result of many factors including societal and cultural differences, but one of the biological factors that has been found to contribute is the masculinization process that occurs early during gestation. Males become more masculinized than females once the SRY gene on the Y-chromosome starts the process that will result in the secretion of testosterone during early gestation (Ferri, Abel, & Brodkin, 2018). In parallel, some studies have shown an association of higher fetal testosterone levels with ASD diagnosis and/or autistic symptoms (Ferri et al., 2018).

Larger brains in individuals with ASD compared to controls have been found previously (Riddle, Cascio, & Woodward, 2017; Stanfield et al., 2008; van Rooij et al., 2018). Although some have suggested that enlarged brains in children with ASD normalize after the age of four (see review Courchesne, Redcay, & Kennedy, 2004), overall, we found slightly larger brains in individuals with ASD compared to controls across a wide age range (5-40 years) (Supplementary Fig. 2A) – in line with Stanfield et al. (2008) as well as a recent longitudinal study (Lee et al., 2020) (2-13 years). This finding found in the total sample may be mostly driven by males though, as within females this size difference may reverse after young adulthood (Supplementary Fig. 2B). However, caution is needed as there is a smaller number of data points for females than males in the ABIDE datasets, and even less so for females older than twenty. The reason for an enlarged brain in ASD is as yet unclear, but it has been thought to be the result of various factors (Freitag et al., 2009), such as an overproduction of neurons, glia, or astrocytes, or decreased pruning. A larger brain in

individuals with ASD compared to controls has been thought to be associated with reduced brain efficiency, such as reduced structural connectivity in the brain. However, a previous review (Ecker, Bookheimer, & Murphy, 2015) found that although most studies have reported poorer long distance connectivity for ASD compared to controls, short distance connectivity was increased in ASD. So, the consequences of an enlarged brain in ASD appear to be complex and are not yet fully understood. Further, it remains unclear whether the same factor(s) influencing brain size differences between the ASD and control group contribute to brain size differences between males and females.

In this study the measure of brain maleness was derived from landmarks placed in subcortical brain regions, as placement showed too much error for cortical landmarks. Our other method to place cortical landmarks more accurately requires high-quality diffusion data (van Eijk et al., in press), which was not available for the ABIDE datasets. We previously used multiple approaches to derive the brain maleness measures in a normal population sample with high-quality diffusion data (The Human Connectome Project) (van Eijk et al., in press), using either subcortical landmarks, cortical landmarks or both type of landmarks (subcortical and cortical). It appeared that mostly subcortical shape was used for the prediction of sex when using both type of landmarks, as the measure derived from the subcortical landmarks alone showed a greater association with the measure derived from both type of landmarks (r=0.686, p<0.001), than the measure derived from cortical landmarks alone (r=0.278, p<0.001). Also, the performance of predicting sex based on shape data (adjusted for brain size) was similar when using subcortical shape data alone versus when using both cortical and subcortical shape (AUC for subcortical data: 85.65-91.93%; AUC for subcortical and cortical data: 85.88-91.83%), while the prediction performance was lower when predicting sex based on cortical data alone (AUC for cortical data: 69.54-79.74%) (van Eijk et al., in press). It remains unclear whether focusing on cortical measures would yield

different results, in particular as it has been found that different genetic factors influence cortical versus subcortical regions, as shown by separate genetic clusters for cortical and subcortical regions (Wen et al., 2016). Further, this research focused on global brain shape, but further research is needed to examine whether the same findings apply to other brain data (e.g. brain function) as well as regional brain differences.

This study has several limitations. First, the clinical ABIDE datasets include a sex-imbalance and a wide age-range. However, we mitigate this by examining effects within each sex, and by adjusting results for age, sex, and a sex by age interaction (if this interaction was significant). We found no significant interaction effects for age by diagnosis for brain maleness or brain size. Second, the clinical datasets include only high-functioning ASD individuals and do not represent all individuals with ASD. Third, in relation to the automatic placement of the landmarks, it is possible that despite our best efforts to check for and exclude processing errors, some landmarks may have been placed inaccurately. However, we have no reason to believe that placement errors affected the scans of the ASD group more than the control group (e.g. scans for both groups had the same image quality), and manual placement would have been extremely time-consuming for such large datasets. In addition, we have previously shown excellent test-retest reliability for brain maleness derived from subcortical shape in two normal population samples (QTIM r=0.955; HCP r=1.000) and good validity (unadjusted for brain size: AUC = 94.81-95.30%; adjusted for brain size with the Procrustes size adjustment: AUC = 85.69-87.01%) (van Eijk et al., in press). Last, it should be noted that our findings are based on mean group statistics, with large overlap observed between the ASD and control group, as well as between males and females, providing evidence for larger individual (within-group) than between-group differences.

Larger multimodal and longitudinal studies of individuals with ASD (or those at high risk for ASD) are needed to unravel ASD-related brain and behavioral changes and to

examine whether these changes are associated with sex differences, in particular during early development. This will enable the identification of biomarkers that can assist with earlier diagnosis to improve outcomes and enhance our understanding of the large individual variability within ASD.

References

Adams, D. C., & Otarola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution, 4*, 393-399. doi:10.1111/2041-210X.12035

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2009). Fetal testosterone and autistic traits. *British Journal of Psychology, 100*(1), 1-22. doi:10.1348/000712608X311731

Barnes, J., Ridgway, G. R., Bartlett, J., Henley, S. M. D., Lehmann, M., Hobbs, N., . . . Fox, N. C. (2010). Head size, age and gender adjustment in MRI studies: a necessary nuisance? *NeuroImage, 53*(4), 1244-1255. doi:10.1016/j.neuroimage.2010.06.025

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: implications for explaining autism. *Science, 310*(5749), 819-823. doi:10.1126/science.1115455

Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R., Ashwin, C., Billington, J., & Chakrabarti, B. (2005). Empathizing and systemizing in autism spectrum conditions. *Handbook of autism and pervasive developmental disorders, 1*, 628-639.

Baron-Cohen, S. (2009). Autism: the empathizing–systemizing (E-S) theory. *Annals of the New York Academy of Sciences, 1156*(1), 68-80. doi:10.1111/j.1749-6632.2009.04467.x

Courchesne, E., Redcay, E., & Kennedy, D. P. (2004). The autistic brain: birth through adulthood. *Current opinion in neurology, 17*(4), 489-496. doi:10.1097/01.wco.0000137542.14610.b4

Di Martino, A., O'Connor, D., Chen, B., Alaerts, K., Anderson, J. S., Assaf, M., . . . Milham, M. P. (2017). Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Scientific Data, 4*, 170010. doi:10.1038/sdata.2017.10

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., . . . Milham, M. P. (2014). The Autism Brain Imaging Data Exchange: Towards Large-Scale Evaluation of the Intrinsic Brain Architecture in Autism. *Molecular psychiatry, 19*(6), 659-667. doi:10.1038/mp.2013.78

Dryden, I. L. (2016). shapes: Statistical Shape Analysis. Retrieved from https://CRAN.R-project.org/package=shapes

Ecker, C., Andrews, D. S., Gudbrandsen, C. M., Marquand, A. F., Ginestet, C. E., Daly, E. M., . . . Consortium, f. t. M. R. C. A. I. M. S. (2017). Association between the probability of autism spectrum disorder and normative sex-related phenotypic diversity in brain structure. *JAMA Psychiatry, 74*(4), 329-338. doi:10.1001/jamapsychiatry.2016.3990

Ecker, C., Bookheimer, S. Y., & Murphy, D. G. M. (2015). Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. *The Lancet Neurology, 14*(11), 1121-1134. doi:10.1016/S1474-4422(15)00050-2

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861-874. doi:10.1016/j.patrec.2005.10.010

Ferri, S. L., Abel, T., & Brodkin, E. S. (2018). Sex Differences in Autism Spectrum Disorder: a Review. *Current Psychiatry Reports, 20*(2), 9. doi:10.1007/s11920-018-0874-2

Frackowiak, R. S. J. (1997). *Human brain function*. San Diego: Academic.

Freitag, C. M., Luders, E., Hulst, H. E., Narr, K. L., Thompson, P. M., Toga, A. W., . . . Konrad, C. (2009). Total Brain Volume and Corpus Callosum Size in Medication-Naïve Adolescents and Young Adults with Autism Spectrum Disorder. *Biological Psychiatry, 66*(4), 316-319. doi:10.1016/j.biopsych.2009.03.011

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., & Turner, R. (1995). Analysis of fMRI Time-Series Revisited. *NeuroImage, 2*(1), 45-53. doi:10.1006/nimg.1995.1007

Goldstein, J. M., Seidman, L. J., Horton, N. J., Makris, N., Kennedy, D. N., Caviness, V. S., . . . Tsuang, M. T. (2001). Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cerebral cortex, 11*(6), 490-497. doi:10.1093/cercor/11.6.490

Greenberg, D. M., Warrier, V., Allison, C., & Baron-Cohen, S. (2018). Testing the Empathizing–Systemizing theory of sex differences and the Extreme Male Brain theory of autism in half a million people. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1811032115

Haar, S., Dinstein, I., Berman, S., & Behrmann, M. (2014). Anatomical Abnormalities in Autism? *Cerebral Cortex, 26*(4), 1440-1452. doi:10.1093/cercor/bhu242 %J Cerebral Cortex

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., . . . Verma, R. (2014). Sex differences in the structural connectome of the human brain. *Proceedings of the National Academy of Sciences, 111*(2), 823-828. doi:10.1073/pnas.1316909110

Klingenberg, C. P. (2016). Size, shape, and form: concepts of allometry in geometric morphometrics. *Dev Genes Evol, 226*(3), 113-137. doi:10.1007/s00427-016-0539-2

Kong, X.-Z., Mathias, S. R., Guadalupe, T., Glahn, D. C., Franke, B., Crivello, F., . . . Francks, C. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *115*(22), E5154-E5163. doi:10.1073/pnas.1718418115 %J Proceedings of the National Academy of Sciences

Lai, M.-C., Lerch, J. P., Floris, D. L., Ruigrok, A. N. V., Pohl, A., Lombardo, M. V., & Baron-Cohen, S. (2017). Imaging sex/gender and autism in the brain: Etiological implications. *Journal of Neuroscience Research, 95*(1-2), 380-397. doi:10.1002/jnr.23948

Lai, M.-C., Lombardo, M. V., Suckling, J., Ruigrok, A. N. V., Chakrabarti, B., Ecker, C., . . . Baron-Cohen, S. (2013). Biological sex affects the neurobiology of autism. *Brain : a journal of neurology, 136*(Pt 9), 2799-2815. doi:10.1093/brain/awt216

Lee, J. K., Andrews, D. S., Ozonoff, S., Solomon, M., Rogers, S., Amaral, D. G., & Nordahl, C. W. (2020). Longitudinal Evaluation of Cerebral Growth Across Childhood in Boys and Girls With Autism Spectrum Disorder. *Biological Psychiatry*. doi:10.1016/j.biopsych.2020.10.014

Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., . . . Evans, A. C. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage, 36*(4), 1065-1073. doi:10.1016/j.neuroimage.2007.03.053

Liu, S., Seidlitz, J., Blumenthal, J. D., Clasen, L. S., & Raznahan, A. (2020). Integrative structural, functional, and transcriptomic analyses of sex-biased brain organization in humans. *Proceedings of the National Academy of Sciences, 117*(31), 18788-18798. doi:10.1073/pnas.1919091117

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., . . . Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers, 6*(1), 5. doi:10.1038/s41572-019-0138-4

Postema, M. C., van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., . . . Francks, C. (2019). Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets. *Nature Communications, 10*(1), 4958. doi:10.1038/s41467-019-13005-8

Rabinowicz, T., Dean, D. E., Petetot, J. M.-C., & de Courten-Myers, G. M. (1999). Gender differences in the human cerebral cortex: more neurons in males; more processes in females. *Journal of Child Neurology, 14*(2), 98-107. doi:10.1177/088307389901400207

Radua, J., Via, E., Catani, M., & Mataix-Cols, D. (2011). Voxel-based meta-analysis of regional white-matter volume differences in autism spectrum disorder versus healthy controls. *Psychological Medicine, 41*(7), 1539-1550. doi:10.1017/S0033291710002187

Riddle, K., Cascio, C. J., & Woodward, N. D. (2017). Brain structure in autism: a voxel-based morphometry analysis of the Autism Brain Imaging Database Exchange (ABIDE). *Brain imaging and behavior, 11*(2), 541-551. doi:10.1007/s11682-016-9534-5

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12*(1), 77. doi:10.1186/1471-2105-12-77

Ruigrok, A. N., Salimi-Khorshidi, G., Lai, M.-C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience & Biobehavioral Reviews, 39*, 34-50. doi:10.1016/j.neubiorev.2013.12.004

Seidlitz, J., Nadig, A., Liu, S., Bethlehem, R. A. I., Vértes, P. E., Morgan, S. E., . . . Raznahan, A. (2020). Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nat Commun, 11*(1), 3358. doi:10.1038/s41467-020-17051-5

Shiino, A., Chen, Y.-w., Tanigaki, K., Yamada, A., Vigers, P., Watanabe, T., . . . Akiguchi, I. (2017). Sex-related difference in human white matter volumes studied: Inspection of the corpus callosum and other white matter by VBM. *Scientific Reports, 7*, 39818. doi:10.1038/srep39818

Smith, R. E. W., Avery, J. A., Wallace, G. L., Kenworthy, L., Gotts, S. J., & Martin, A. (2019). Sex Differences in Resting-State Functional Connectivity of the Cerebellum in Autism Spectrum Disorder. *Frontiers in Human Neuroscience, 13*(104). doi:10.3389/fnhum.2019.00104

Sparks, B. F., Friedman, S. D., Shaw, D. W., Aylward, E. H., Echelard, D., Artru, A. A., . . . Dager, S. R. (2002). Brain structural abnormalities in young children with autism spectrum disorder. *59*(2), 184-192. doi:10.1212/WNL.59.2.184 %J Neurology

Stanfield, A. C., McIntosh, A. M., Spencer, M. D., Philip, R., Gaur, S., & Lawrie, S. M. (2008). Towards a neuroanatomy of autism: A systematic review and meta-analysis of structural magnetic resonance imaging studies. *European Psychiatry, 23*(4), 289-299. doi:10.1016/j.eurpsy.2007.05.006

van Eijk, L., Zhu, D., Couvy-Duchesne, B., Strike, L. T., Lee, A. J., Hansell, N. K., . . . Zietsch, B. P. (in press). Are sex differences in human brain structure associated with sex differences in behaviour? *Psychological Science*.

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., . . . Yacoub, E. (2012). The Human Connectome Project: A data acquisition perspective. *NeuroImage, 62*(4), 2222-2231. doi:10.1016/j.neuroimage.2012.02.018

van Rooij, D., Anagnostou, E., Arango, C., Auzias, G., Behrmann, M., Busatto, G. F., . . . Buitelaar, J. K. (2018). Cortical and Subcortical Brain Morphometry Differences Between Patients With Autism Spectrum Disorder and Healthy Individuals Across the Lifespan: Results From the ENIGMA ASD Working Group. *Am J Psychiatry, 175*(4), 359-369. doi:10.1176/appi.ajp.2017.17010100

Venables, W. N., and Ripley, B.D. (2002). *Modern Applied Statistics with S* (Fourth ed.). New York: Springer.

Wen, W., Thalamuthu, A., Mather, K. A., Zhu, W., Jiang, J., de Micheaux, P. L., . . . Sachdev, P. S. (2016). Distinct Genetic Influences on Cortical and Subcortical Brain Structures. *Scientific Reports, 6*(1), 32760. doi:10.1038/srep32760

Tables

Table 1. Prediction performance (displayed as the Area Under the Curve[1] with 95% confidence interval) in the total ABIDE sample and in three subsamples (sex-balanced subset, controls, and ASD group) – either unadjusted (upper row) or adjusted for brain size (lower row) during the Procrustes analysis.

	Full dataset (N=2,060)	**Sex-balanced (N=782)**	**Controls (N=1,094)**	**ASD (N=966)**
Not adjusted for brain size	84.06 (82.02-86.10)	83.64 (80.85-86.42)	84.52 (81.99-87.06)	82.54 (78.88-86.14)
Adjusted for brain size	78.78 (76.37-81.19)	79.40 (76.30-82.50)	79.00 (76.00-82.00)	77.82 (73.53-82.12)

[1]The Area Under the Curve is the true positive rate against the false positive rate. ABIDE=Autism Brain Imaging Data Exchange I and II. ASD=Autism Spectrum Disorder

Figure Legends

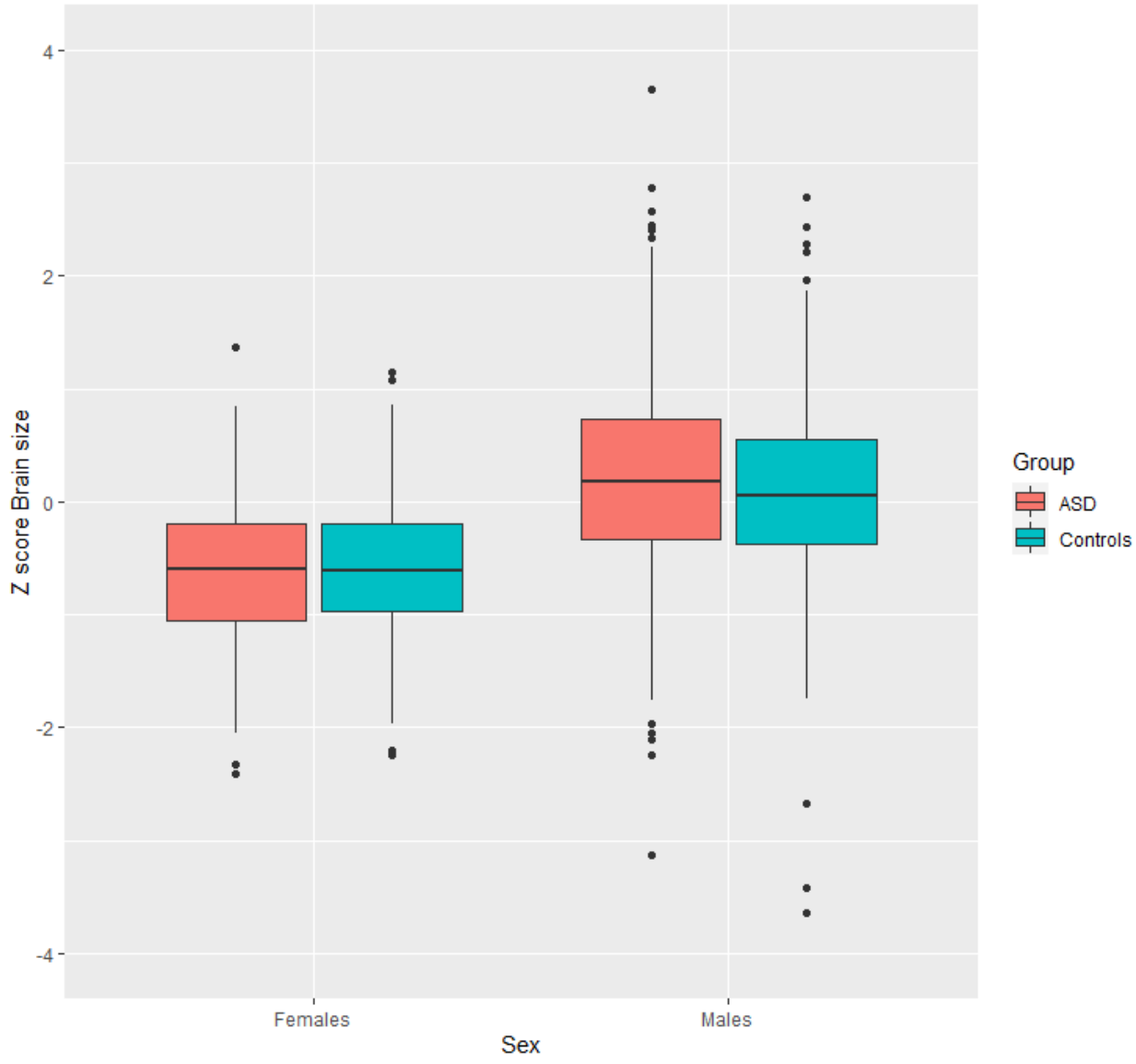
Fig. 1.

The different stages to obtain a measure of brain maleness for each individual in the ABIDE datasets, including placement of the landmarks, a Generalized Procrustes and Principal Component Analysis, and the prediction of sex (algorithm trained on population samples QTIM and HCP and predicted in ABIDE-I and II). QTIM=Queensland Twin IMaging; HCP=Human Connectome Project; ABIDE=Autism Brain Imaging Data Exchange.

Fig. 2.

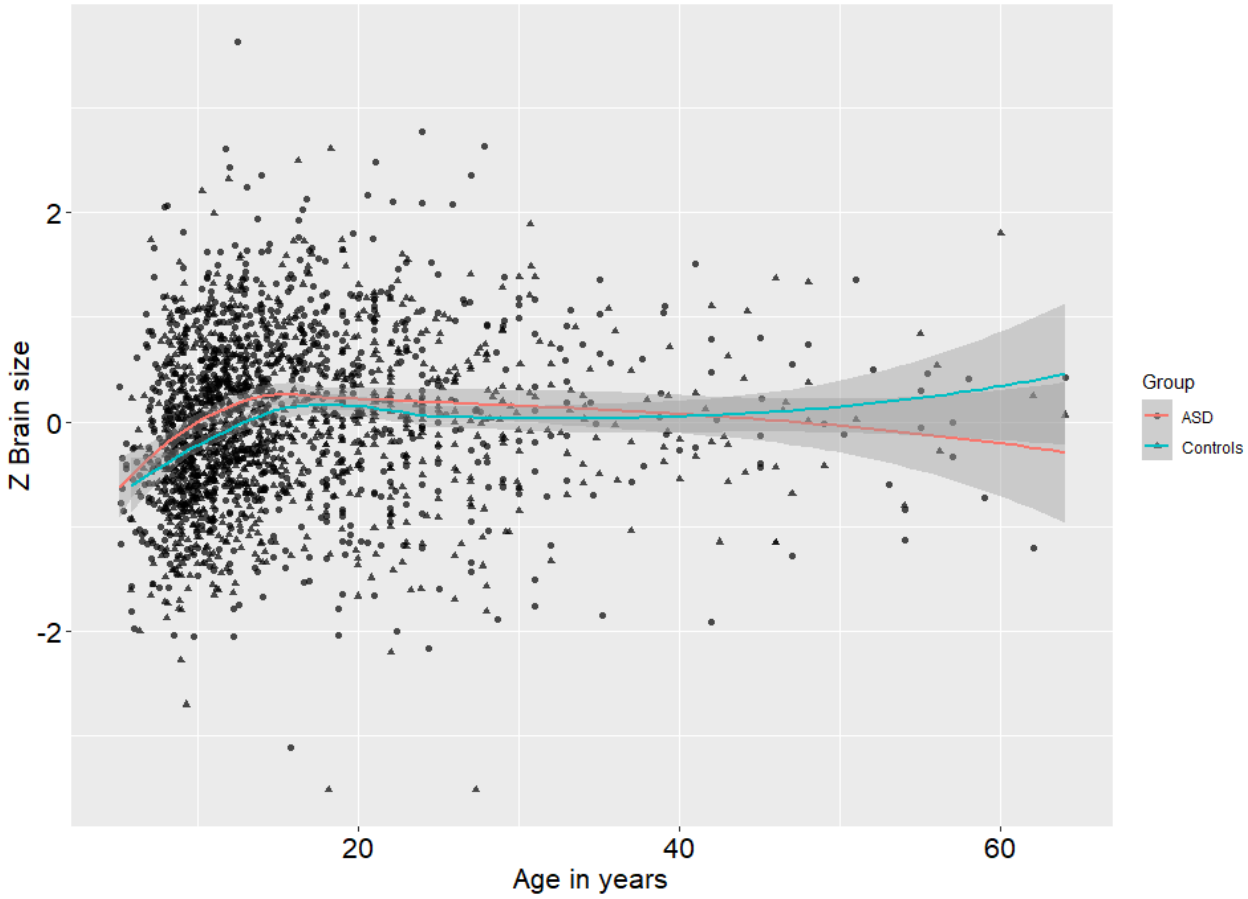
Distribution of brain maleness scores (adjusted for covariates age and cohort), with separate boxplots for the ASD group and controls, separately for males and females. Brain maleness scores without adjustment for brain size are displayed on the left, scores with Procrustes size adjustment are displayed in the middle figure, whereas the scores with Procrustes size adjustment plus regression for brain size are displayed on the right. The ASD group is displayed in red and controls in green. Note, 81.02% of sample is male (N ASD females =128, N ASD males =838, N control females =263, N control males =831).

Supplementary Figures



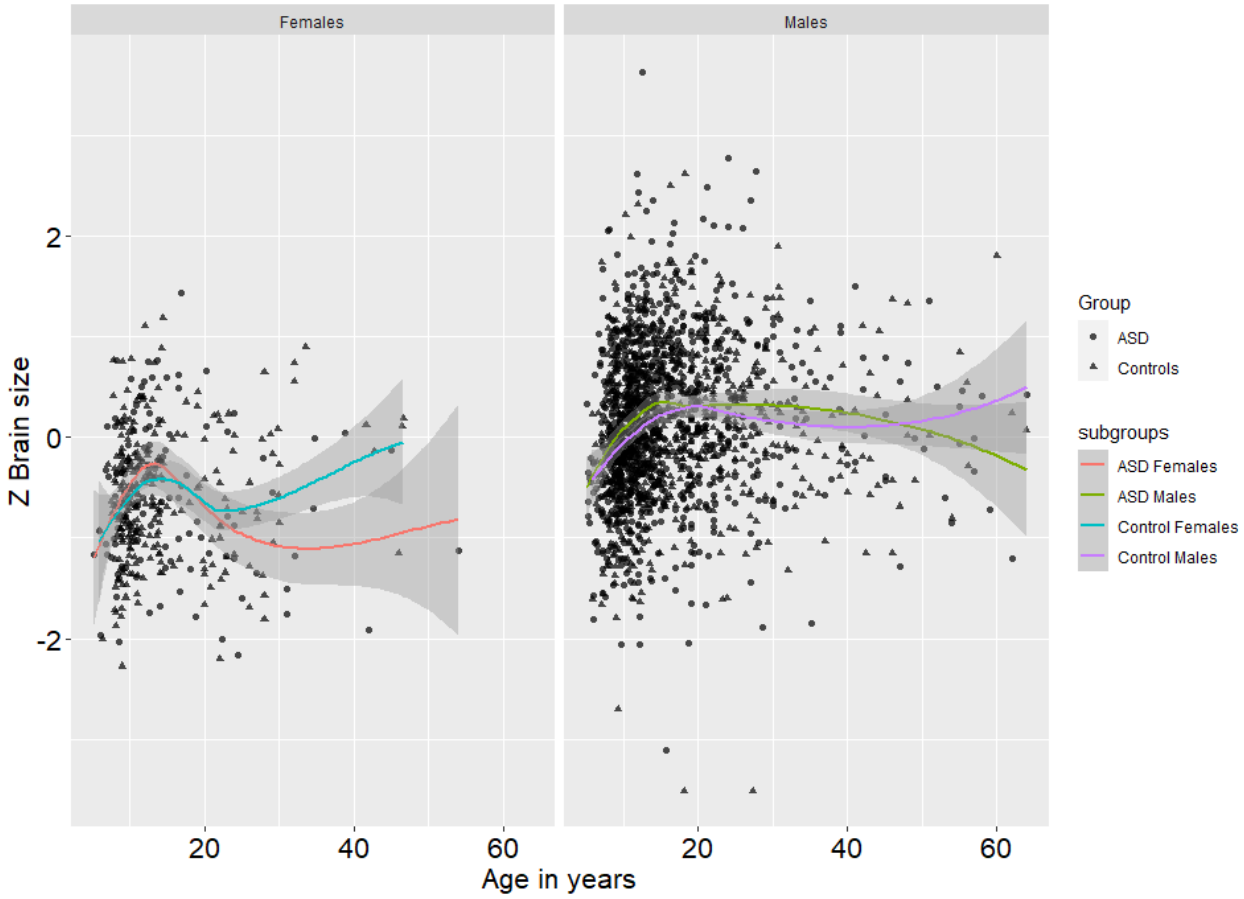
Supplementary Fig. 1.

Distribution of brain size (scaled), plotted by diagnosis and sex (adjusted for age and cohort). ASD group is displayed in red and controls in green. Note, sample is 81.02% male (N ASD females =128, N ASD males =838, N controls females =263, N controls males =831). ASD=Autism Spectrum Disorder.



Supplementary Fig. 2A.

Distribution of brain size (scaled) plotted by age, separately for ASD and controls.
Note, sample is 81.02% male (N ASD females =128, N ASD males =838, N controls females =263, N controls males =831). ASD=Autism Spectrum Disorder.



Supplementary Fig. 2B.

Distribution of brain size (scaled) plotted by age, separately for each subgroup (diagnosis by sex), with females (ASD and controls) displayed in the left panel and males (ASD and controls) displayed in the right panel.

Note, sample is 81.02% male (N ASD females =128, N ASD males =838, N controls females =263, N controls males =831). ASD=Autism Spectrum Disorder.

Supplementary Tables

Supplementary Table 1.

Association between brain maleness (no adjustment for brain size) and behavioral measures (Autism Diagnostic Interview-Revised (ADI), the Autism Diagnostic Observation Schedule Module (ADOS), and the Social Responsiveness Scale Edition (SRS), both in the total sample (adjusted for age, sex, cohort, and interaction effect sex by age) and within each sex (adjusted for age and cohort).

	Total				***Males***				***Females***			
Questionnaire	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***
ADI_R_SOCIAL_TOTAL_A	0.096 (0.019-0.172)	2.44	639	0.015*	0.091 (0.007-0.174)	2.14	544	0.033*	0.193 (-0.009-0.380)	1.89	93	0.061
ADI_R_VERBAL_TOTAL_BV	0.092 (0.015-0.169)	2.34	638	0.019*	0.087 (0.003-0.169)	2.03	543	0.043*	0.193 (-0.009-0.379)	1.89	93	0.062
ADI_RRB_TOTAL_C	0.097 (0.020-0.173)	2.47	639	0.014*	0.093 (0.009-0.175)	2.17	544	0.031*	0.192 (-0.010-0.379)	1.89	93	0.062
ADI_R_ONSET_TOTAL_D	0.101 (0.022-0.179)	2.52	615	0.012*	0.098 (0.012-0.182)	2.24	522	0.026*	0.193 (-0.011-0.382)	1.88	91	0.063
ADOS_TOTAL	0.053 (-0.018-0.125)	1.46	747	0.144	0.080 (0.004-0.156)	2.07	658	0.039*	-0.057 (-0.263-0.153)	-0.54	87	0.593
ADOS_COMM	0.039 (-0.034-0.112)	1.06	722	0.291	0.066 (-0.012-0.143)	1.65	633	0.099	-0.059 (-0.264-0.152)	-0.55	87	0.586
ADOS_SOCIAL	0.049 (-0.024-0.121)	1.31	722	0.191	0.078 (0.000-0.154)	1.96	633	0.051	-0.058 (-0.263-0.152)	-0.54	87	0.589
ADOS_STEREO_BEHAV	0.044 (-0.030-0.117)	1.16	703	0.245	0.082 (0.003-0.160)	2.05	618	0.041*	-0.062 (-0.272-0.153)	-0.57	83	0.570
ADOS_GOTHAM_SOCAFFECT	-0.028 (-0.113-0.058)	-0.64	521	0.526	-0.057 (-0.149-0.037)	-1.20	441	0.233	0.247 (0.029-0.443)	2.26	78	0.027*
ADOS_GOTHAM_RRB	-0.016 (-0.102-0.070)	-0.37	521	0.710	-0.055 (-0.147-0.038)	-1.16	441	0.248	0.370 (0.163-0.545)	3.51	78	0.001**
ADOS_GOTHAM_TOTAL	-0.028 (-0.113-0.058)	-0.64	521	0.526	-0.057 (-0.149-0.037)	-1.20	441	0.233	0.247 (0.029-0.443)	2.26	78	0.027*
ADOS_GOTHAM_SEVERITY	-0.028 (-0.114-0.058)	-0.64	516	0.521	-0.060 (-0.153-0.034)	-1.26	438	0.210	0.257 (0.037-0.454)	2.32	76	0.023*
SRS_RAW_TOTAL	-0.043 (-0.099-0.014)	-1.48	1186	0.141	-0.068 (-0.131;-0.005)	-2.12	958	0.034*	0.105 (-0.025-0.232)	1.59	226	0.113
SRS_AWARENESS	0.027	0.77	819	0.444	0.007	0.19	644	0.852	0.060	0.79	173	0.433

	(-0.042-0.095)				(-0.070-0.084)				(-0.090-0.206)			
SRS_COGNITION	0.047 (-0.022-0.115)	1.33	819	0.183	0.018 (-0.059-0.095)	0.47	644	0.642	0.015 (-0.134-0.163)	0.19	173	0.849
SRS_COMMUNICATION	0.053 (-0.014-0.120)	1.55	847	0.121	0.025 (-0.051-0.100)	0.64	672	0.523	0.035 (-0.114-0.182)	0.46	173	0.645
SRS_MOTIVATION	0.046 (-0.022-0.113)	1.33	847	0.182	0.003 (-0.072-0.079)	0.09	672	0.929	0.040 (-0.109-0.187)	0.53	173	0.597
SRS_MANNERISMS	0.051 (-0.016-0.118)	1.49	847	0.138	0.028 (-0.047-0.104)	0.73	672	0.463	0.014 (-0.135-0.162)	0.18	173	0.857

* $p \leq 0.05$; ** $p \leq 0.0028$ (association survived the Bonferroni-correction); r=Pearson's correlation (95% confidence interval), df=degrees of freedom, t=t-statistic, p=p-value. ADI=Autism Diagnostic Interview-Revised, ADOS=Autism Diagnostic Observation Schedule Module, SRS=Social Responsiveness Scale Edition.

Supplementary Table 2.

Association between brain maleness (adjusted for brain size with the Procrustes size adjustment) and behavioral measures (Autism Diagnostic Interview-Revised (ADI), the Autism Diagnostic Observation Schedule Module (ADOS), and the Social Responsiveness Scale Edition (SRS), both in the total sample (adjusted for age, sex and cohort) and within each sex (adjusted for age and cohort).

	Total				***Males***				***Females***			
Questionnaire	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***
ADI_R_SOCIAL_TOTAL_A	0.065 (-0.012-0.142)	1.65	639	0.100	0.049 (-0.035-0.133)	1.15	544	0.249	0.205 (0.003-0.390)	2.02	93	0.047*
ADI_R_VERBAL_TOTAL_BV	0.057 (-0.021-0.134)	1.44	638	0.151	0.039 (-0.045-0.123)	0.92	543	0.361	0.204 (0.003-0.390)	2.01	93	0.047*
ADI_RRB_TOTAL_C	0.067 (-0.010-0.144)	1.70	639	0.090	0.052 (-0.032-0.135)	1.21	544	0.228	0.204 (0.003-0.390)	2.01	93	0.047*
ADI_R_ONSET_TOTAL_D	0.066 (-0.013-0.144)	1.64	615	0.101	0.051 (-0.035-0.136)	1.16	522	0.246	0.207 (0.004-0.394)	2.02	91	0.046*
ADOS_TOTAL	0.023 (-0.048-0.095)	0.64	747	0.523	0.043 (-0.034-0.118)	1.09	658	0.275	0.004 (-0.205-0.212)	0.04	87	0.971
ADOS_COMM	0.016 (-0.057-0.089)	0.43	722	0.671	0.035 (-0.043-0.112)	0.87	633	0.384	0.003 (-0.205-0.211)	0.03	87	0.977
ADOS_SOCIAL	0.018 (-0.055-0.091)	0.48	722	0.628	0.039 (-0.039-0.116)	0.97	633	0.331	0.003 (-0.205-0.211)	0.03	87	0.975
ADOS_STEREO_BEHAV	-0.002 (-0.076-0.072)	-0.06	703	0.953	0.024 (-0.055-0.103)	0.61	618	0.545	0.002 (-0.211-0.215)	0.02	83	0.985
ADOS_GOTHAM_SOCAFFECT	0.065 (-0.021-0.150)	1.49	521	0.138	0.056 (-0.037-0.149)	1.19	441	0.235	0.173 (-0.049-0.378)	1.55	78	0.125
ADOS_GOTHAM_RRB	0.088 (0.002-0.172)	2.01	521	0.045*	0.064 (-0.029-0.157)	1.36	441	0.176	0.293 (0.078-0.482)	2.71	78	0.008*
ADOS_GOTHAM_TOTAL	0.065 (-0.021-0.150)	1.49	521	0.138	0.056 (-0.037-0.149)	1.19	441	0.235	0.173 (-0.049-0.378)	1.55	78	0.125
ADOS_GOTHAM_SEVERITY	0.063 (-0.023-0.149)	1.44	516	0.151	0.055 (-0.039-0.148)	1.15	438	0.251	0.169 (-0.056-0.377)	1.50	76	0.139
SRS_RAW_TOTAL	0.007 (-0.050-0.064)	0.23	1186	0.816	-0.007 (-0.070-0.056)	-0.22	958	0.823	0.073 (-0.058-0.201)	1.10	226	0.274
SRS_AWARENESS	0.006 (-0.063-0.074)	0.16	819	0.871	-0.011 (-0.088-0.067)	-0.27	644	0.787	0.042 (-0.107-0.189)	0.55	173	0.582
SRS_COGNITION	0.004 (-0.064-0.072)	0.12	819	0.907	-0.008 (-0.085-0.069)	-0.21	644	0.838	-0.046 (-0.193-0.103)	-0.61	173	0.546

SRS_COMMUNICATION	0.006 (-0.061-0.074)	0.19	847	0.851	-0.014 (-0.090-0.061)	-0.37	672	0.714	-0.002 (-0.150-0.146)	-0.03	173	0.977
SRS_MOTIVATION	0.001 (-0.066-0.068)	0.03	847	0.979	-0.024 (-0.099-0.052)	-0.62	672	0.534	-0.034 (-0.182-0.115)	-0.45	173	0.654
SRS_MANNERISMS	-0.009 (-0.077-0.058)	-0.28	847	0.782	-0.030 (-0.106-0.045)	-0.79	672	0.430	-0.009 (-0.157-0.139)	-0.12	173	0.904

* $p \leq 0.05$; none of the associations survived the Bonferroni-correction ($p \leq 0.0028$); r=Pearson's correlation (95% confidence interval), df=degrees of freedom, t=t-statistic, *p*=*p*-value. ADI=Autism Diagnostic Interview-Revised, ADOS=Autism Diagnostic Observation Schedule Module, SRS=Social Responsiveness Scale Edition.

Supplementary Table 3.

Association between brain maleness (adjusted for brain size with the Procrustes size adjustment plus regression for brain size) and behavioral measures (Autism Diagnostic Interview-Revised (ADI), the Autism Diagnostic Observation Schedule Module (ADOS), and the Social Responsiveness Scale Edition (SRS), both in the total sample (adjusted for age, sex and cohort) and within each sex (adjusted for age and cohort).

	Total				***Males***				***Females***			
Questionnaire	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***
ADI_R_SOCIAL_TOTAL_A	0.054 (-0.023-0.131)	1.37	639	0.171	0.032 (-0.052-0.115)	0.74	544	0.459	0.170 (-0.032-0.360)	1.67	93	0.099
ADI_R_VERBAL_TOTAL_BV	0.046 (-0.031-0.123)	1.17	638	0.243	0.022 (-0.062-0.106)	0.52	543	0.603	0.170 (-0.033-0.359)	1.67	93	0.099
ADI_RRB_TOTAL_C	0.057 (-0.020-0.134)	1.45	639	0.147	0.036 (-0.048-0.120)	0.84	544	0.400	0.170 (-0.033-0.359)	1.66	93	0.100
ADI_R_ONSET_TOTAL_D	0.056 (-0.023-0.135)	1.40	615	0.162	0.035 (-0.051-0.120)	0.80	522	0.426	0.174 (-0.031-0.364)	1.68	91	0.096
ADOS_TOTAL	0.019 (-0.053-0.090)	0.51	747	0.608	0.026 (-0.051-0.102)	0.66	658	0.507	0.035 (-0.174-0.242)	0.33	87	0.743
ADOS_COMM	0.013 (-0.060-0.085)	0.34	722	0.733	0.020 (-0.058-0.098)	0.51	633	0.608	0.035 (-0.175-0.241)	0.33	87	0.746
ADOS_SOCIAL	0.013 (-0.060-0.086)	0.34	722	0.732	0.020 (-0.058-0.098)	0.51	633	0.608	0.035 (-0.175-0.241)	0.33	87	0.745
ADOS_STEREO_BEHAV	-0.014 (-0.088-0.060)	-0.38	703	0.707	-0.006 (-0.085-0.073)	-0.16	618	0.877	0.035 (-0.179-0.246)	0.32	83	0.751
ADOS_GOTHAM_SOCAFFECT	0.083 (-0.002-0.168)	1.91	521	0.057	0.092 (-0.001-0.184)	1.94	441	0.053	0.130 (-0.093-0.340)	1.16	78	0.251
ADOS_GOTHAM_RRB	0.103 (0.018-0.187)	2.37	521	0.018*	0.098 (0.005-0.189)	2.06	441	0.040*	0.221 (0.001-0.420)	2.00	78	0.049*
ADOS_GOTHAM_TOTAL	0.083 (-0.002-0.168)	1.91	521	0.057	0.092 (-0.001-0.184)	1.94	441	0.053	0.130 (-0.092-0.340)	1.16	78	0.251
ADOS_GOTHAM_SEVERITY	0.081 (-0.005-0.166)	1.85	516	0.064	0.091 (-0.003-0.183)	1.91	438	0.057	0.124 (-0.102-0.337)	1.09	76	0.281
SRS_RAW_TOTAL	0.017 (-0.039-0.074)	0.60	1186	0.548	0.020 (-0.044-0.083)	0.61	958	0.544	0.035 (-0.095-0.164)	0.53	226	0.598
SRS_AWARENESS	0.005 (-0.064-0.073)	0.14	819	0.890	-0.008 (-0.085-0.069)	-0.20	644	0.838	0.027 (-0.121-0.175)	0.36	173	0.719
SRS_COGNITION	-0.005 (-0.074-0.063)	-0.16	819	0.876	-0.016 (-0.093-0.061)	-0.41	644	0.679	-0.058 (-0.204-0.091)	-0.76	173	0.447

SRS_COMMUNICATION	-0.006 (-0.073-0.061)	-0.18	847	0.859	-0.029 (-0.104-0.047)	-0.75	672	0.454	-0.016 (-0.164-0.133)	-0.21	173	0.833
SRS_MOTIVATION	-0.013 (-0.080-0.054)	-0.39	847	0.700	-0.035 (-0.110-0.041)	-0.91	672	0.364	-0.059 (-0.206-0.090)	-0.78	173	0.435
SRS_MANNERISMS	-0.023 (-0.090-0.045)	-0.66	847	0.507	-0.049 (-0.124-0.026)	-1.28	672	0.202	-0.014 (-0.162-0.135)	-0.18	173	0.857

* $p \leq 0.05$, none of the associations survived the Bonferroni-correction ($p \leq 0.0028$); r=Pearson's correlation (95% confidence interval), df=degrees of freedom, t=t-statistic, *p*=*p*-value. ADI=Autism Diagnostic Interview-Revised, ADOS=Autism Diagnostic Observation Schedule Module, SRS=Social Responsiveness Scale Edition.

Supplementary Table 4.

Association between brain size and behavioral measures (Autism Diagnostic Interview-Revised (ADI), the Autism Diagnostic Observation Schedule Module (ADOS), and the Social Responsiveness Scale Edition (SRS), both in the total sample (adjusted for age, sex, cohort, and interaction effect sex by age) and within each sex (adjusted for age and cohort).

	Total				***Males***				***Females***			
Questionnaire	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***	***r***	**t**	**df**	***p***
ADI_R_SOCIAL_TOTAL_A	0.059 (-0.018-0.136)	1.50	639	0.134	0.056 (-0.028-0.139)	1.31	544	0.192	0.138 (-0.066-0.330)	1.34	93	0.183
ADI_R_VERBAL_TOTAL_BV	0.056 (-0.021-0.133)	1.42	638	0.156	0.052 (-0.032-0.136)	1.22	543	0.223	0.138 (-0.066-0.330)	1.34	93	0.183
ADI_RRB_TOTAL_C	0.055 (-0.023-0.132)	1.38	639	0.167	0.051 (-0.033-0.134)	1.18	544	0.238	0.137 (-0.066-0.330)	1.34	93	0.185
ADI_R_ONSET_TOTAL_D	0.056 (-0.023-0.134)	1.39	615	0.166	0.052 (-0.033-0.137)	1.20	522	0.231	0.140 (-0.065-0.335)	1.35	91	0.179
ADOS_TOTAL	0.023 (-0.048-0.095)	0.64	747	0.522	0.047 (-0.029-0.123)	1.22	658	0.224	-0.095 (-0.298-0.115)	-0.89	87	0.374
ADOS_COMM	0.017 (-0.056-0.089)	0.45	722	0.653	0.041 (-0.037-0.118)	1.02	633	0.307	-0.096 (-0.299-0.114)	-0.90	87	0.369
ADOS_SOCIAL	0.026 (-0.047-0.099)	0.70	722	0.485	0.052 (-0.026-0.129)	1.31	633	0.190	-0.096 (-0.298-0.115)	-0.90	87	0.371
ADOS_STEREO_BEHAV	0.052 (-0.022-0.125)	1.38	703	0.169	0.087 (0.008-0.164)	2.16	618	0.031*	-0.099 (-0.306-0.116)	-0.91	83	0.366
ADOS_GOTHAM_SOCAFFECT	-0.066 (-0.151-0.019)	-1.52	521	0.129	-0.090 (-0.182-0.003)	-1.91	441	0.057	0.191 (-0.030-0.394)	1.72	78	0.089
ADOS_GOTHAM_RRB	-0.051 (-0.136-0.035)	-1.17	521	0.242	-0.083 (-0.175-0.010)	-1.76	441	0.079	0.322 (0.110-0.506)	3.01	78	0.004*
ADOS_GOTHAM_TOTAL	-0.066 (-0.151-0.019)	-1.52	521	0.129	-0.090 (-0.182-0.003)	-1.91	441	0.057	0.191 (-0.030-0.394)	1.72	78	0.089
ADOS_GOTHAM_SEVERITY	-0.066 (-0.152-0.020)	-1.51	516	0.131	-0.093 (-0.185-0.001)	-1.95	438	0.052	0.206 (-0.017-0.410)	1.84	76	0.070
SRS_RAW_TOTAL	-0.045 (-0.102-0.012)	-1.56	1186	0.120	-0.075 (-0.137;-0.012)	-2.32	958	0.020*	0.129 (-0.001-0.255)	1.95	226	0.052
SRS_AWARENESS	0.008 (-0.060-0.077)	0.23	819	0.816	-0.008 (-0.085-0.069)	-0.21	644	0.837	0.056 (-0.094-0.202)	0.73	173	0.466
SRS_COGNITION	0.045 (-0.023-0.113)	1.30	819	0.194	0.021 (-0.056-0.098)	0.54	644	0.589	0.039 (-0.110-0.187)	0.52	173	0.605

SRS_COMMUNICATION	0.059 (-0.009-0.125)	1.71	847	0.087	0.038 (-0.038-0.113)	0.99	672	0.325	0.050 (-0.099-0.197)	0.66	173	0.511
SRS_MOTIVATION	0.063 (-0.004-0.130)	1.85	847	0.065	0.027 (-0.049-0.102)	0.70	672	0.484	0.089 (-0.060-0.234)	1.17	173	0.242
SRS_MANNERISMS	0.060 (-0.008-0.126)	1.74	847	0.083	0.047 (-0.029-0.122)	1.22	672	0.223	0.016 (-0.133-0.164)	0.21	173	0.835

* $p \leq 0.05$; none of the associations survived the Bonferroni-correction ($p \leq 0.0028$); r=Pearson's correlation (95% confidence interval), df=degrees of freedom, t=t-statistic, *p*=*p*-value. ADI=Autism Diagnostic Interview-Revised, ADOS=Autism Diagnostic Observation Schedule Module, SRS=Social Responsiveness Scale Edition.