Origins of gender differences in the human brain

van Hemmen, J.

2017

document version
Publisher's PDF, also known as Version of record

Chapter 5

Sexual differentiation of the human brain is affected by a multifactorial mechanism. Evidence from the complete androgen insensitivity syndrome

Judy van Hemmen
Paola Valsasina
Paolo Misci
Annemieke S. Staphorsius
Dick J. Veltman
Massimo Filippi
Julie Bakker

To be submitted

ABSTRACT

Animal studies have revealed a role for sex chromosome genes in addition to sex hormone effects in brain sexual differentiation. To investigate if a similar mechanism applies to humans, brain structure and function was studied in women with complete androgen insensitivity syndrome (CAIS), who have a 46,XY karyotype but are resistant to androgens. A multivariate probabilistic pattern recognition method was employed to analyze structural (gray matter [GM] volume, fractional anisotropy [FA]), and functional (during resting-state, mental rotation, and affective picture task performance) magnetic resonance imaging (MRI) data. First, the predictive value of these neuroimaging modalities in discriminating 31 control men from 31 control women was assessed. Second, successful classification algorithms were applied to 21 women with CAIS. Classification of women with CAIS into either of the control groups, using classifiers with significantly accurate classification of control men and women (GM volume, FA, combined multimodal), was not above chance level. These findings suggest that sex differences in these modalities do not solely reflect differences in androgen exposure. Instead, the mechanism underlying sexual differentiation of these neural characteristics seems to be multifactorial, and most likely includes a combination of sex hormone, sex chromosome, and environmental effects, with effects varying across brain regions.

INTRODUCTION

Delineating neural sex differences and the mechanisms underlying sexual differentiation of the brain is of great importance. Research on this topic increases our understanding of the origins of typical brain development, as well as the development of neuropsychiatric disorders with a sex difference in prevalence (Zahn-Waxler et al. 2006; Bao and Swaab 2011). Ever since the classic theory of brain sexual differentiation was presented by Phoenix et al. (1959), who demonstrated a permanent effect of prenatal testosterone on adult reproductive behavior in rodents, most research has focused on the organizational, i.e., permanent, and activational, i.e., transient, influence of sex hormones during different life stages; prenatal, adolescence, adulthood. There is, however, a growing interest in the role of genes on the sex chromosomes in the sexual differentiation of the brain (McCarthy and Arnold 2011). Although mouse models have provided evidence for direct sex chromosome effects in certain sexually differentiated behaviors (for review see Cox et al. 2014) and brain structure (Corre et al. 2016), it is challenging to disentangle hormonal and genetic effects in humans because genes on the sex chromosomes are involved in gonadal development and, thus, influence the hormonal profile indirectly. Since these factors cannot be manipulated in human research, scientists primarily have to rely on disorders/differences of sex development (DSDs), which is an umbrella term for "congenital conditions in which development of chromosomal, gonadal, or anatomical sex is atypical" (as cited in Hughes et al. 2006).

A DSD of particular interest when studying the effects of sex hormones versus sex chromosome genes on the sexual differentiation of the human brain is the complete androgen insensitivity syndrome (CAIS). CAIS is a rare condition characterized by (a) mutation(s) in the X-linked androgen receptor (AR) gene, causing complete androgen resistance in the presence of a 46,XY karyotype. Therefore, CAIS results in a female phenotype, regardless of normal or elevated amounts of testosterone secreted by the abdominal, pelvic or inguinal testes while *in situ* (Melo et al. 2003; Doehnert et al. 2015).

The few neuroimaging studies that have thus far been performed in women with CAIS have generally found female-typical neural characteristics. More specifically, these studies showed that localized sex differences in brain function during exposure to sexual images (Hamann et al. 2014) and visuospatial task performance (van Hemmen et al. 2014), and in white matter (WM) microstructure throughout extended WM regions, are not substantially influenced by genes on the sex chromosomes, but rather reflect differences in sex hormone exposure and/or female-typical socialization. These sex hormone effects may include masculinizing androgen effects through the AR, feminizing estrogen effects, or both. For some aspects of WM microstructure, however, a more subtle contribution of genes on the sex chromosomes, or androgen effects not mediated through the AR, could not be ruled out (van Hemmen et al. 2016).

In the present study we applied a multivariate pattern recognition (MPR) method to structural and functional neuroimaging data of women with CAIS, control men (CM), and control women (CW) to further investigate the relative influence and importance of sex chromosome genes and sex hormones on the sexual differentiation of multiple aspects of the human brain. An advantage of MPR is that it can detect subtle, spatially distributed patterns in the data (Schrouff et al. 2013), which might reflect between-group differences that have not been detected with mass-univariate analyses employed in previous studies. Modalities included in the present study were selected based on previously reported sex differences (e.g., reviewed in Sacher et al. 2013): regional gray matter (GM) volume, WM microstructure reflected by fractional anisotropy (FA), functional magnetic resonance imaging (fMRI) during performance of a mental rotation task (MRT) and an affective picture task (APT), and during rest (rs-fMRI). First, we employed MPR to examine the predictive value of each modality in discriminating control men from control women. Second, modalities with significant discriminative power were combined into a multimodal classifier. Finally, successful classifiers were applied to women with CAIS.

MATERIALS AND METHODS

Subjects

From the initial 85 subjects that underwent MRI scanning, the neuroimaging data of 2 subjects (control women) were excluded from analyses in all modalities because of anatomical abnormalities. This resulted in a total of 21 women with CAIS, 31 control men (CM), and 31 control women (CW) with neuroimaging data suitable for analysis in at least one modality. Previous data from this sample have been published (van Hemmen et al. 2014, 2016). General exclusion criteria included MRI contraindications, a history of a serious medical, neurological or psychiatric disease, and, for CW, the use of hormonal contraceptives. All subjects had a right-hand writing preference and, with one exception (ambidexter), were right-handed according to the Dutch Handedness Inventory (Van Strien 1992). All subjects had a gender identity and gender role behavior in accordance with their assigned gender; male in CM, female in CW and women with CAIS, and were heterosexual, that is, women with CAIS and CW were androphilic, and CM were gynephilic (see van Hemmen et al. 2014 for further details regarding gender-related psychological functioning questionnaires).

For each imaging modality, additional subjects were excluded when there was a modality-specific reason for exclusion; excessive head movement, technical failure, or task-related problems (e.g., reports of bad visual perception of the stimuli). In order to prevent a large reduction in sample size and allow maximal data utilization, subjects' MRI data were only excluded for the analysis of the modality in which a modality-specific problem was present, resulting in small between-modality differences in sample size (see Table 1). In the initial

sample, groups were matched for age and level of education. As presented in Table 1, no differences in age and level of education were found between the three groups for any of the samples used in the present study.

Table 1 Demographics for the complete sample and classifier-specific samples

	Control men	Control women	CAIS	χ^2	P-value
Sample sizes					
Complete sample	31	31	21		
Classifier-specific samples (range)	29 - 30	29 - 31	19 - 21		
Mean (SD) - Complete sample					
Age	31.43 (9.46)	31.90 (9.56)	32.05 (11.77)	0.078	0.962
Level of education (E)	6.39 (1.86)	6.32 (1.70)	6.10 (1.61)	0.566	0.754
Level of education (C)	5.58 (1.78)	5.90 (1.49)	5.90 (1.64)	0.633	0.729
Mean (SD) range - Classifier-specific samples					
Age	31.33 - 32.11 (9.40 - 9.60)	31.20 - 32.06 (9.39 - 9.72)	30.50 - 32.05 (10.96 - 11.77)	0.009 - 0.507	0.776 - 0.996
Level of education (E)	6.28 - 6.37 (1.86 - 1.88)	6.27 - 6.45 (1.67 - 1.70)	6.00 - 6.32 (1.53 - 1.61)	0.093 - 0.690	0.708 - 0.955
Level of education (C)	5.53 -5.55 (1.80 - 1.82)	5.83 - 6.10 (1.46 - 1.50)	5.80 - 6.11 (1.59 - 1.64)	0.522 - 1.867	0.393 - 0.770

Note: Ranges represent minimum and maximum values of N, mean and SD for classifier-specific samples; MRT, APT-positive and -negative, rs-fMRI, GM, FA, and multimodal. Scores for the level of education range from 1 (primary school) to 8 (university degree). C: current level at the day of testing; E: expected level once the participant has finished the current educational trajectory (if applicable).

In women with CAIS the diagnosis was based on clinical features, and mutation analysis of the AR gene using genomic DNA showed that 13 women had a confirmed AR gene mutation, 7 women an unclassified variant, and in 1 woman the result of the analysis was inconclusive. Since all women with CAIS were gonadectomized and thus lacked gonadal hormone production, all except one woman used estrogen replacement therapy consisting of only estrogens (n = 16) or estrogens combined with progestins (n = 4).

Women with CAIS were recruited from the support group DSDNederland and from the databases of the VU University Medical Center and the Erasmus University Medical Center-Sophia's Children's Hospital Rotterdam. CM and CW were recruited using advertisements and flyers. Study approval was obtained from the Medical Ethics Committee of the VU

University Medical Center Amsterdam (application number NL32740.029.10). Participants gave their written informed consent according to the Declaration of Helsinki.

MRI data acquisition

MRI data were acquired at 3T (Signa HDxt, General Electric, Milwaukee, WI, USA) using an 8-channel head coil. fMRI data were acquired with gradient echo-planar imaging (EPI) sequences during the performance of the MRT and APT (TR 2100 ms, TE 30 ms, matrix size 64 × 64, 40 slices, 3.75 × 3.75 × 3.0 mm^3 voxels) and during the resting-state (repetition time [TR] 1800 ms, echo time [TE] 35 ms, matrix size 64 × 64, 34 slices, 3.3 mm^3 isotropic resolution). Preceding the rs-fMRI acquisition, participants were instructed to keep their eyes closed and not to fall asleep during scanning, which was verified immediately afterwards. T1-weighted anatomical images were acquired using a 3D fast spoiled gradient echo sequence (TR 7.8 ms, TE 3.0 ms, inversion time [TI] 450 ms, matrix size 256 × 256, 172 slices, 1 mm^3 isotropic resolution). For DTI, single-shot EPI was used to acquire 5 nondiffusion weighted (b0) volumes and 30 volumes with noncollinear diffusion gradients (b 1000 s/mm^2, TR 13 s, TE 86 ms, matrix size 128 × 128, 45 slices, 2 × 2 × 2.4 mm^3 voxels).

fMRI task paradigms

Participants received task instructions prior to the MRI session, which were repeated immediately preceding the fMRI experiments. E-prime (Psychology Software Tools, Pittsburgh, PA, USA) was used to project instructions and test stimuli on a screen visible through a mirror attached to the head coil. Participants used MRI-compatible button boxes to respond to the stimuli.

Mental rotation task

We refer to van Hemmen et al. (2014) for a detailed description of the fMRI MRT paradigm. In short; an alternating block design was used to present a) rotation stimuli consisting of 2 individually rotated identical or mirrored 3D objects (Shepard and Metzler 1971) and b) control stimuli consisting of one 3D object with an arrow underneath pointing left or right. Participants used button boxes to indicate a) whether the objects in the rotation stimuli were identical or mirrored versions of each other, or b) the direction of the arrow for control stimuli.

Affective pictures task

The APT consisted of 3 conditions, comprising 30 images each; positive affective valence, negative affective valence, and a control condition. Stimulus images were selected from the International Affective Picture System (IAPS) (Lang, P.J., Bradley, M.M., & Cuthbert 2008). For the control condition, stimuli were scrambled images with an arrow in the center. Stimuli were successively presented with a duration of 2 s each and an inter-stimulus interval vary-

ing between 2.5 and 3.5 s. To ensure attention to the stimuli, participants were asked to rate the stimuli in the positive and negative affective valence conditions as pleasant or unpleasant by pressing the right or left button box, respectively. For the control stimuli, participants were asked to press the button corresponding to the direction the arrow was pointing at.

(f)MRI data preprocessing and modelling

Mental rotation task

Preprocessing and first-level within-subject modelling of the fMRI data obtained during MRT performance has been described in detail in van Hemmen et al. (2014). Individual contrast images created by subtracting the control condition from the rotation condition were entered into a multiple regression model with individual mean reaction time (RT) to rotation stimuli as a covariate of no interest using Statistical Parametric Mapping 12 (SPM; Wellcome Trust Center for Neuroimaging, Institute of Neurology at UCL, UK), implemented in MATLAB (version R2011a; The MathWorks, Inc., Natick, MA, USA). Residual contrast images calculated with this regression model were used as input for the MPR analysis.

Affective pictures task

Preprocessing of the fMRI data with SPM8 included manual reorientation to the AC-PC line, slice-time correction to the middle slice, realignment to the first image, co-registration of the anatomical T1 image to the mean EPI image, normalization to MNI space using the normalization parameters derived from the segmented anatomical image, and smoothing using an 8 mm full width at half maximum (FWHM) Gaussian kernel. At the within-subject level, the trial onsets were convolved with a canonical hemodynamic response function for each condition. 'Scan nulling' regressors for each inter-volume head movement exceeding 0.5 mm and 6 motion regressors obtained during realignment were added to the model to account for motion-related artifacts. Furthermore, a high-pass filter (128 s) was applied to account for low-frequency signal drift. Individual contrast images were created for positive affective valence > control (APT-Positive) and negative affective valence > control (APT-Negative). The validity of the task was first verified by analyzing the main effect of these contrasts across groups using a whole-brain $P < 0.05$ family-wise error (FWE) corrected threshold. In both the APT-Positive and APT-Negative main effect, prominent bilateral activation was observed in limbic regions, including the amygdala and hippocampus, the occipital cortex, anterior cingulate cortex, superior medial frontal gyrus, insula, inferior frontal gyrus, thalamus, and caudate nucleus. Since these brain regions have previously been associated with positive and negative emotional stimuli (e.g., meta-analysis by Stevens and Hamann 2012), both contrast images were used as input images for the MPR analysis.

rs-fMRI

SPM8 was used for preprocessing of the rs-fMRI data. After manual reorientation and discarding of the first 2 volumes, the images were realigned to the mean image, normalized to MNI space while subsampling to 3 x 3 x 4 mm^3 voxel resolution, and smoothed using a 6 mm FWHM Gaussian kernel. The REST toolbox (version 1.7; Song et al. 2011) was used for linear detrending and band-pass filtering between 0.01 and 0.08 Hz. Whenever inter-volume head movement exceeded 0.5 mm, the succeeding volume was discarded to remove motion-related effects.

The default mode network (DMN), which shows increased activity during rest and decreased activity during cognitive performance (Raichle et al. 2001; Greicius et al. 2003; Raichle and Snyder 2007), is one of the most frequently studied resting-state networks (RSN) and has been found to show sex differences in functional connectivity (e.g., reviewed in Mak et al. 2016). The DMN was therefore selected as our RSN of interest and was identified by group independent component analysis (ICA) using the GIFT toolbox (Calhoun et al. 2001). First, the preprocessed functional data were reduced using a 2-step principal component analysis (PCA). Second, a spatially constrained ICA algorithm, which is a semi-blind ICA algorithm incorporating prior spatial information into the estimation process (Lin et al. 2010), was applied to the reduced dataset to estimate independent group components. The 28 RSNs from Allen et al. (2011) were selected as spatial reference files. Third, individual subject components were computed through back reconstruction, and component spatial maps and time courses were converted to z-scores. The DMN was identified by both visual inspection and spatial correlation of the independent components resulting from ICA with the spatial reference file representing the DMN. DMN individual component images were used in a multiple regression model in SPM12 with covariate age. Calculated residual images, reflecting individual component images of the DMN corrected for age-related effects, were used as input for the MPR analysis.

T1-weighted MRI

Regional GM volume was calculated from the T1-weighted MRI data. First, the brain extraction tool (BET; Smith 2002, implemented in FSL v5.04 http://www.fmrib.ox.au.uk/fsl, Smith et al. 2004) was used to remove non-brain tissue from manually reoriented T1-weigted images using image bias and residual neck voxel reduction settings and a 0.1 fractional intensity threshold. A voxel-based morphometry (VBM) approach (Ashburner and Friston 2000) was employed for further preprocessing of these images with default parameters of the VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm8/) in SPM8. Images were segmented into GM, WM, and cerebrospinal fluid (CSF) and normalized to the IXI550 template using high-dimensional DARTEL normalization (Ashburner 2007). The nonlinear component of the Jacobian determinant, derived from spatial normalization, was used to modulate the normalized GM images, thereby correcting for individual brain size. These modulated normalized GM

images were smoothed with an 8 mm FWHM Gaussian kernel. To correct for age-related effects in the GM images, the smoothed modulated (non-linear only) normalized GM images were used in a multiple regression model in SPM12 with covariate age and a 0.2 absolute threshold. Residual GM images were used as input for the MPR analysis.

DTI

DTI data preprocessing has been described in detail in van Hemmen et al. (2016). All subjects' skeletonized fractional anisotropy (FA) images resulting from these preprocessing steps were used in a multiple regression model in SPM12 with covariate age to calculate residual FA images. These age-corrected skeletonized FA images were used as input for the MPR analysis.

MPR analysis

MPR analyses were performed using the Pattern Recognition for Neuroimaging Toolbox (PRoNTo v2.0; Schrouff et al. 2013). Individual images that resulted from the data preprocessing and modelling steps per modality were used as input for the MPR analyses. Whole-brain masks were applied to MPR analyses of all fMRI modalities, and a mean FA skeleton mask obtained from tract-based spatial statistics (TBSS) was applied to FA images. GM data were masked with a whole-brain GM mask in which two clusters representing non-GM voxels, most likely resulting from imperfect segmentation of dural venous sinuses and veins (e.g., Good et al. 2001), were removed to prevent an erroneous contribution of these voxels to the classifier function.

In the first step of the MPR analysis, the accuracy of men-women classification in control groups was assessed for each imaging modality. A probabilistic pattern recognition algorithm, the binary Gaussian Process Classification (GPC) machine, was selected, because it produces probabilistic class membership predictions which can be used for further analyses. A leave-one-subject-out (LOSO) cross-validation strategy was used, which computes the accuracy of a classifier by iteratively excluding one subject for testing and using all other subjects for training until all subjects have been tested. Classifier performance for each modality is expressed in both balanced accuracy (BA), i.e., the percentage of correctly classified subjects while taking into account the number of subjects in each class, and area under the receiver operating characteristic (ROC) curve (AUC). Permutation testing using 1000 permutations was performed to determine statistical significance of the BA, that is, whether it exceeds chance level (50%). An FDR correction (Benjamini and Hochberg 1995) for the number of unimodal classifiers tested was applied to the resulting P-values (similar to e.g., Schmaal et al. 2015).

Second, a multimodal MPR analysis was performed to evaluate if a combination of unimodal classifiers would result in a classifier with improved performance compared with the performance of the individual unimodal classifiers. For this purpose, unimodal classifiers

with a significant BA ($P < 0.05$ FDR-corrected) were selected and combined into a multimodal classifier using a median rule (Kittler et al. 1998). An advantage of the median rule approach is that it allows within-subject missing data in one or more modalities, thereby enabling optimal usage of available data for training of the classifier function for each modality. Subjects with valid data in at least one of the selected unimodal classifiers were included in the multimodal analysis (see Table 1 for group sizes). The median rule was applied to each subjects' individual class membership probabilities extracted from the selected unimodal classifiers. Class probabilities are function values scaled between 0 and 1 that are computed during the MPR analysis and underlie the binary class predictions. These predictive values, therefore, provide exact information about the predictive uncertainty of binary classification. For the multimodal classifier, the medians calculated by applying the median rule to the unimodal class probabilities were similarly used to make individual binary class predictions. In this study, CM were selected as the reference class for all classifiers, therefore a > 0.5 median class probability resulted in a classification prediction for class 1 (CM) and a < 0.5 probability for class 2 (CW). Similar to the unimodal analyses, multimodal classifier performance regarding the classification of CM and CW is expressed in BA and AUC, and a P-value was calculated using permutation testing (1000 permutations). To assess differences in performance between the multimodal classifier and the significant unimodal classifiers, statistical comparisons of the AUCs were performed (Hanley and McNeil 1982, 1983) using MedCalc Statistical Software version 16.8 (MedCalc Software bvba, Ostend, Belgium; https://www.medcalc.org; 2016). A Bonferroni correction for multiple comparisons was applied.

In the third step of the MPR analysis, classifiers with a significant BA for the classification of CM and CW were selected to predict group membership in women with CAIS. For the unimodal classifiers, binary GPCs were applied with a custom cross-validation scheme; per iteration, one woman with CAIS was tested against the classifier trained using CM and CW, which was repeated until each woman with CAIS was tested. Individual class membership probabilities were extracted for each subject. For the multimodal classifier, binary class predictions were calculated for women with CAIS according to their median class probabilities based on the combined multimodal classifier using the median rule.

To assess the classification of women with CAIS relative to both control groups statistically, class membership probabilities were analyzed with non-parametric tests using IBM SPSS Statistics for Windows, Version 22 (IBM Corp., Armonk, NY, USA). Between-group comparisons of class probabilities were performed using Kruskal-Wallis and post hoc Mann-Whitney U tests. One-sample Wilcoxon signed rank tests were performed for each group independently to test if median class membership probabilities differed significantly from 0.5. A Bonferroni correction was applied for the number of post hoc between-group and Wilcoxon signed rank tests performed per classifier.

RESULTS

Gaussian process classification

Performance of unimodal and multimodal classifiers

Classifier performance is summarized in Table 2. Unimodal binary GPC analyses of the MRT, APT-Positive, GM, and FA data resulted in BAs significantly above chance level for the classification of CM and CW. The BAs of classifiers using APT-Negative and DMN information did not significantly exceed chance level. The multimodal classifier, in which information from the four unimodal classifiers with a significant BA was combined, was also able to discriminate between CM and CW with a significant BA. All subjects included in this multimodal analysis (see Table 1) had valid data in three (n = 7) or all four (n = 75) modalities. The ROC curves of the classifiers with significant BAs are displayed in Figure 1. Comparisons of the multimodal classifier AUC against the AUCs of the significant unimodal classifiers showed that the multimodal classifier AUC was significantly larger than the AUC of the MRT classifier (P = 0.035), while AUCs of the APT-Positive, GM, and FA classifier were not significantly different from the multimodal classifier AUC (P's 0.227, 0.615, 1.000, respectively).

Table 2 Classifier performance and classification of women with CAIS

	CM vs CW classification		CAIS classification	
	AUC	BA (P-value)	M	W
Unimodal classifiers				
MRT	0.66	66% (0.038)	62%	38%
APT-Positive	0.74	63% (0.042)	35%	65%
APT-Negative	0.62	59% (0.118)	—	—
rs-fMRI - DMN	0.60	57% (0.142)	—	—
GM	0.78	75% (0.003)	57%	43%
FA	0.90	78% (0.003)	40%	60%
Multimodal classifier				
MRT, APT-Positive, GM, FA	0.89	84% (0.001)	43%	57%

Note: P-values are based on 1000 permutations, unimodal P-values are FDR-corrected. AUC: area under the ROC curve; BA: balanced accuracy; CM: control men; CW: control women; CAIS: women with CAIS. CAIS classification was performed with classifiers with a significant BA (P < 0.05).

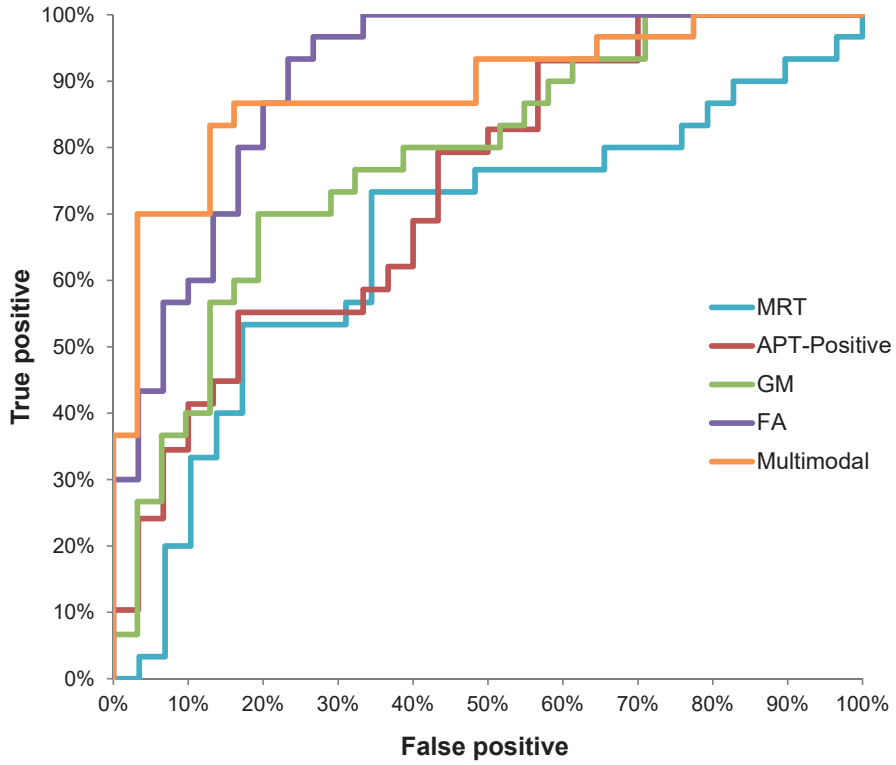

Figure 1 ROCs of classifiers with significant BAs.

Classification of women with CAIS

Class predictions in women with CAIS using the classifiers with a significant BA (MRT, APT-Positive, GM, FA, and multimodal) are presented in Table 2. Between-group comparisons of the individual class membership probabilities (using the class representing CM as the reference group) underlying the class predictions showed an effect of group for all modalities tested (Table 3). Results from post hoc pairwise comparisons of class probabilities are displayed in Figure 2. No significant between-group differences in class probabilities were found for the MRT classifier (CM versus CW P = 0.103; CM versus women with CAIS P = 1.000; CW versus women with CAIS P = 0.094). For the APT-Positive, GM, FA, and multimodal classifiers, class membership probabilities were significantly larger in CM than in CW (P's 0.005, < 0.001, < 0.001, < 0.001, respectively). Class probabilities in women with CAIS were not significantly different from CM (P = 0.132) and CW (P = 1.000) for the APT-Positive classifier. For the classifiers using GM and multimodal information, class membership probabilities in women with CAIS were significantly larger than in CW (P's 0.043 and 0.002, respectively) and were not significantly different from CM (P's 0.706 and 0.096 respectively). For the FA

classifier, class probabilities of women with CAIS were significantly smaller than in CM (P = 0.008) and not significantly different from CW (P = 0.055).

Table 3 Median class membership probability per group, difference from 0.5 and between-group statistics for classifiers with significant BAs

	Group	CM class probability	Difference from 0.5			Between-group	
		Median (IQR)	Z	P-value	Effect size	χ^2	P-value
Unimodal classifiers							
MRT	CM	0.62 (0.28)	1.717	0.258	0.66	6.240	**0.044**
	CW	0.47 (0.27)	-1.287	0.594	0.49		
	CAIS	0.53 (0.30)	1.651	0.297	0.77		
APT-Positive	CM	0.60 (0.25)	2.065	0.117	0.83	10.579	**0.005**
	CW	0.40 (0.33)	-2.314	0.063	0.93		
	CAIS	0.45 (0.29)	-1.045	0.888	0.48		
GM	CM	0.62 (0.31)	2.417	**0.048**	0.98	15.382	**< 0.001**
	CW	0.37 (0.24)	-2.959	**0.009**	1.25		
	CAIS	0.58 (0.26)	0.643	1.000	0.28		
FA	CM	0.58 (0.18)	3.486	**0.001**	1.65	30.867	**< 0.001**
	CW	0.38 (0.10)	-3.939	**< 0.001**	2.07		
	CAIS	0.48 (0.20)	-0.933	1.000	0.43		
Multimodal classifier							
MRT, APT-Positive, GM, FA	CM	0.58 (0.16)	3.281	**0.003**	1.50	30.323	**< 0.001**
	CW	0.37 (0.12)	-4.096	**< 0.001**	2.17		
	CAIS	0.49 (0.17)	0.330	1.000	0.14		

Note: CM class: median (IQR: interquartile range) class membership probability with the class representing control men as the reference class. Effect sizes are expressed in Cohen's *d*. Bold P-values represent a significant (P < 0.05) effect. CM: control men; CW: control women; CAIS: women with CAIS

Comparisons of the median class probability per group versus a probability value of 0.5 (Table 3 and Fig. 2), showed that medians were significantly higher than 0.5 in CM and lower in CW with large effect sizes for the GM, FA, and multimodal classifiers, but not for the MRT and APT-Positive classifiers. Median class membership probabilities in women with CAIS did not differ significantly from 0.5 for any of the classifiers tested.

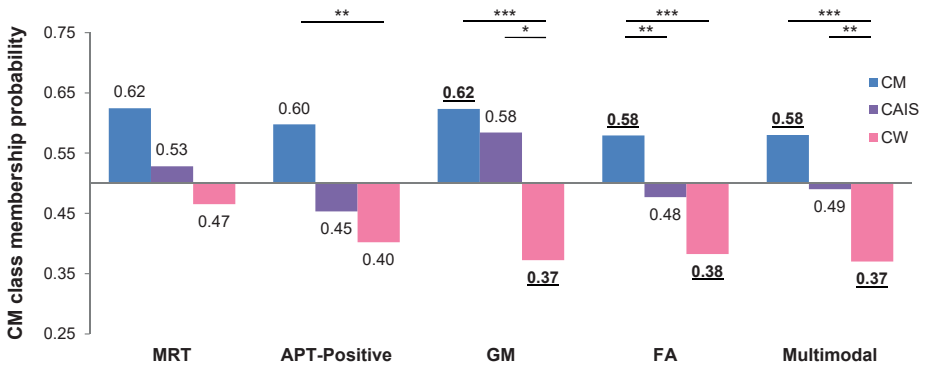

Figure 2 Median class membership probability for classifiers with significant BAs. Significant between-group differences are marked with asterisks: * = P < 0.05; ** = P < 0.01; P < 0.001. Medians with a significant difference from a 0.5 probability value are bold and underlined.

DISCUSSION

To examine the factors involved in the sexual differentiation of the human brain, probabilistic MPR was employed to assess patterns of brain structure and function in women with CAIS relative to CM and CW. Consistent with the current view of neural sexual differentiation based on animal studies (McCarthy and Arnold 2011; Arnold 2017), the present findings provide evidence for a complex, multifactorial mechanism underlying the sexual differentiation of the human brain.

In the control groups, CM and CW could be accurately discriminated based on unimodal classifiers using spatially distributed patterns of mental rotation (MRT) and positive emotion-related (APT-positive) brain activation, GM volume, and WM microstructure (FA), as well as the multimodal classifier in which these successful unimodal classifiers were combined. This is in line with previous studies that have reported sex differences in these modalities when using mass-univariate analyses (e.g., reviewed in Sacher et al. 2013), and recent MPR studies that have achieved high classification accuracies (between 80% and 90%) using structural neuroimaging modalities (Wang et al. 2012; Feis et al. 2013).

Application of these successful classifier algorithms to women with CAIS showed that, overall, there was no compelling classification of women with CAIS into either the class representing CM or CW. The GM, FA, and multimodal classifiers were most informative, since their above chance discriminative power in control groups was also reflected in the class membership probabilities extracted from the models, thereby enabling further interpretation of the classification of women with CAIS. Predictive probabilities in women with CAIS were similar to CM, i.e. higher than in CW, for both the GM and multimodal classifier and were similar to CW, i.e. lower than in CM, for the FA classifier. These results alone should, however, not be interpreted in terms of male- or female-typical spatial patterns

of neural characteristics in women with CAIS, as this interpretation requires the predictive probabilities of this group to exceed chance level accordingly. Median class membership probabilities in women with CAIS, however, did not significantly differ from a 0.5 probability, i.e., chance level, for any of the classifiers. This shows that women with CAIS were equally likely to belong to the class representing CM as to the class representing CW based on their spatially distributed patterns of GM volume, WM microstructure, and multimodal neural characteristics. These multimodal characteristics reflect the combined information of GM volume, WM microstructure, and brain function related to mental rotation and positive emotions.

It is important to incorporate findings from previous neuroimaging studies in women with CAIS, including our own work in an overlapping sample, in the interpretation of the present results. These previous studies have found predominant female-typical regional brain activation in response to sexual stimuli and during mental rotation, as well as in WM microstructure (Hamann et al. 2014; van Hemmen et al. 2014, 2016). The apparent discrepancy between previous and current findings can be explained by the differences in methodological approaches. Mass-univariate methods analyze between-group differences at voxel-level, which introduces a significant bias towards the detection of spatially localized between-group differences (Davatzikos 2004). In contrast, the multivariate nature of MPR methods enables the detection of spatially complex and subtle patterns of differences (Schrouff et al. 2013), by simultaneous evaluation of all voxels. The present findings, therefore, do not contradict these previous findings, but instead increase our understanding of the sexual differentiation of the human brain and provide a more nuanced portrayal of the underlying mechanism. Thus, while previous mass-univariate studies have provided evidence for localized female-typical WM microstructure and functional activation in women with CAIS, the current MPR analyses indicate that these localized findings are part of spatially complex neural profiles which, overall, are neither male-, nor female-typical. Although the majority of previous neuroimaging findings in women with CAIS were female-typical, in line with the current findings, one of the DTI metrics that was used to further characterize FA in our previous DTI study was not convincingly female-typical (van Hemmen et al. 2016). Likewise, a behavioral study of spatial learning in women with CAIS did not show an overall female-typical pattern (Mueller et al. 2016).

The interpretation of the present findings with respect to the factors involved in the sexual differentiation of the brain, while incorporating previous findings, is twofold. First, previous studies showed that certain localized sex differences in WM microstructure and brain activation related to sexual stimuli and mental rotation predominantly reflect differences in sex hormone exposure and/or socialization effects, and not direct effects from genes on the sex chromosomes (Hamann et al. 2014; van Hemmen et al. 2014, 2016). These proposed sex hormone effects may include a) androgenic masculinization through activation of the AR, since these effects would be absent in women with CAIS, and/or b)

estrogenic feminization, as women with CAIS are exposed to androgen-derived estrogens produced by the testes while still *in situ*, or through estrogen replacement therapy following gonadectomy. Second, the findings from the present study show that these factors do not dominate the sexual differentiation throughout the entire brain. Instead, direct genetic effects related to differences in sex chromosome complement, or androgenic actions that are not mediated through the AR, are also likely to influence the development of sex differences in several neural characteristics. Based on the findings in the present study, these neural characteristics appear to include, but are not limited to, spatially distributed patterns of GM volume and WM microstructure. Similar effects were also reflected in the multimodal classifier, although interpretation of the exact neural characteristics related to this classifier is difficult due to the combination of four imaging modalities. Taken together, the present findings are in agreement with the current emphasis on a more complex model of brain sexual differentiation, which thus far has been primarily based on findings from animal studies that have for instance revealed an important role for sex chromosome genes in addition to steroid hormones in sexual behavior and brain structure (McCarthy and Arnold 2011; Arnold et al. 2016; Corre et al. 2016). This model integrates multiple factors that contribute to the development of neural sex differences, including sex hormones, environment, and direct effects of sex chromosome genes.

A limitation of the present study is that not all imaging modalities proved to be useful for the interpretation of MPR results in women with CAIS. First, regardless of previously reported sex differences in brain function related to negative emotions (see meta-analysis by Stevens and Hamann 2012) and the DMN (Bluhm et al. 2008; Allen et al. 2011; Tomasi and Volkow 2012; Filippi et al. 2013; Jung et al. 2015; Ypma et al. 2016; but see Weissman-Fogel et al. 2010), these modalities did not have sufficient predictive value to distinguish CM from CW in our sample. Second, even though the results from the MPR analyses revealed significant classification performance for the MRT and APT-positive classifiers, their discriminative power was insufficient to significantly exceed chance level during post hoc analyses of class membership probabilities. Unfortunately, we had to depend on these probabilistic values for the interpretation of the classification results in women with CAIS. The absence of sufficient discriminative power in these modalities might be explained by more subtle and complex spatial patterns compared with the other modalities. Larger samples available during training of the classifier might enable the detection of these patterns and the calculation of an algorithm with sufficient or increased predictive power to discriminate CM from CW. Regarding the classification results of women with CAIS when using the GM, FA, and multimodal classifiers, the lack of significant classification into either the class representing CM or CW is not likely to be explained by insufficient statistical power, as the effect sizes for the difference from 0.5 class probability are small to medium for women with CAIS, while these effect sizes in the control groups are large.

While superior classification performance in the control groups was expected when successful unimodal classifiers were combined into a multimodal classifier, the combined information from MRT, APT-positive, GM, and FA classifiers only outperformed the unimodal MRT classifier. This might be explained by insufficient predictive power of the two fMRI modalities, as became apparent from the post hoc analyses of class predictive values. In future studies, it would be interesting to investigate if other combinations of information, for instance also including behavioral data, would result in further improvement of classification performance. However, it is important to note that, while this would be valuable for the classification of CM and CW, it is questionable if it would provide further insights with respect to the aim of the present study. The classification results of women with CAIS using the multimodal classifier are difficult to interpret as such, because it is unclear to what extent each modality drives the classification and therefore which neural characteristics are responsible for these results.

In conclusion, the present study provides unique insight into the factors involved in the sexual differentiation of the human brain. Whole-brain spatial patterns of WM microstructure and GM volume in women with CAIS were neither predominantly female- nor male-typical. This demonstrates that the mechanism underlying the development of sex differences in these neural characteristics is multifactorial. Sexual differentiation of these spatial patterns are not only under the influence of masculinizing androgen effects through the AR, feminizing estrogen effects, or environmental influences, as has previously been shown for highly localized regions and activation clusters. Instead, at least with respect to the sexual differentiation of WM microstructure and GM volume, there is also a role for genes on the sex chromosomes, or androgen effects that are not mediated through the AR.

REFERENCES

Allen EA, Erhardt EB, Damaraju E, Gruner W, Segall JM, Silva RF, Havlicek M, Rachakonda S, Fries J, Kalyanam R, Michael AM, Caprihan A, Turner JA, Eichele T, Adelsheim S, Bryan AD, Bustillo J, Clark VP, Feldstein Ewing SW, Filbey F, Ford CC, Hutchison K, Jung RE, Kiehl K a, Kodituwakku P, Komesu YM, Mayer AR, Pearlson GD, Phillips JP, Sadek JR, Stevens M, Teuscher U, Thoma RJ, Calhoun VD. 2011. A baseline for the multivariate comparison of resting-state networks. Front Syst Neurosci. 5:Article 2.

Arnold AP. 2017. A general theory of sexual differentiation. J Neurosci Res. 95:291–300.

Arnold AP, Reue K, Eghbali M, Vilain E, Chen X, Ghahramani N, Itoh Y, Li J, Link JC, Ngun T, Williams-Burris SM. 2016. The importance of having two X chromosomes. Philos Trans R Soc Lond B Biol Sci. 371:20150113.

Ashburner J. 2007. A fast diffeomorphic image registration algorithm. Neuroimage. 38:95–113.

Ashburner J, Friston KJ. 2000. Voxel-based morphometry--the methods. Neuroimage. 11:805–821.

Bao A-M, Swaab DF. 2011. Sexual differentiation of the human brain: relation to gender identity, sexual orientation and neuropsychiatric disorders. Front Neuroendocrinol. 32:214–226.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 57:289–300.

Bluhm RL, Osuch EA, Lanius RA, Boksman K, Neufeld RWJ, Théberge J, Williamson P. 2008. Default mode network connectivity: effects of age, sex, and analytic approach. Neuroreport. 19:887–891.

Calhoun VD, Adali T, Pearlson GD, Pekar JJ. 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp. 14:140–151.

Corre C, Friedel M, Vousden DA, Metcalf A, Spring S, Qiu LR, Lerch JP, Palmert MR. 2016. Separate effects of sex hormones and sex chromosomes on brain structure and function revealed by high-resolution magnetic resonance imaging and spatial navigation assessment of the Four Core Genotype mouse model. Brain Struct Funct. 221:997–1016.

Cox KH, Bonthuis PJ, Rissman EF. 2014. Mouse model systems to study sex chromosome genes and behavior: relevance to humans. Front Neuroendocrinol. 35:405–419.

Davatzikos C. 2004. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. Neuroimage.

Doehnert U, Bertelloni S, Werner R, Dati E, Hiort O. 2015. Characteristic features of reproductive hormone profiles in late adolescent and adult females with complete androgen insensitivity syndrome. Sex Dev. 9:69–74.

Feis D-L, Brodersen KH, von Cramon DY, Luders E, Tittgemeyer M. 2013. Decoding gender dimorphism of the human brain using multimodal anatomical and diffusion MRI data. Neuroimage. 70:250–257.

Filippi M, Valsasina P, Misci P, Falini A, Comi G, Rocca MA. 2013. The organization of intrinsic brain activity differs between genders: A resting-state fMRI study in a large cohort of young healthy subjects. Hum Brain Mapp. 34:1330–1343.

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage. 14:21–36.

Greicius MD, Krasnow B, Reiss AL, Menon V. 2003. Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. Proc Natl Acad Sci. 100:253–258.

Hamann S, Stevens J, Vick JH, Bryk K, Quigley CA, Berenbaum SA, Wallen K. 2014. Brain responses to sexual images in 46,XY women with complete androgen insensitivity syndrome are female-typical. Horm Behav. 66:724–730.

Hanley AJ, McNeil JB. 1982. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. Radiology. 143:29–36.

Hanley JA, McNeil BJ. 1983. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 148:839–843.

Hughes IA, Houk C, Ahmed SF, Lee PA. 2006. Consensus statement on management of intersex disorders. J Pediatr Urol. 2:148–162.

Jung M, Mody M, Saito DN, Tomoda A, Okazawa H, Wada Y, Kosaka H. 2015. Sex Differences in the Default Mode Network with Regard to Autism Spectrum Traits: A Resting State fMRI Study. PLoS One. 10.

Kittler J, Hatef M, Duin RPW, Matas J. 1998. On combining classifiers. IEEE Trans patterns Anal Mach Intell. 20:226–239.

Lang, P.J., Bradley, M.M., & Cuthbert BN. 2008. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. Univ Florida, Gainesville, FL.

Lin QH, Liu J, Zheng YR, Liang H, Calhoun VD. 2010. Semiblind spatial ICA of fMRI using spatial constraints. Hum Brain Mapp. 31:1076–1088.

Mak LE, Minuzzi L, MacQueen G, Hall G, Kennedy S, Milev R. 2016. The Default Mode Network in Healthy Individuals: A Systematic Review and Meta-Analysis. Brain Connect. brain.2016.0438.

McCarthy MM, Arnold AP. 2011. Reframing sexual differentiation of the brain. Nat Neurosci. 14:677–683.

Melo KFS, Mendonca BB, Billerbeck AEC, Costa EMF, Inácio M, Silva FAQ, Leal AMO, Latronico AC, Arnhold IJP. 2003. Clinical, hormonal, behavioral, and genetic characteristics of androgen insensitivity syndrome in a Brazilian cohort: five novel mutations in the androgen receptor gene. J Clin Endocrinol Metab. 88:3241–3250.

Mueller SC, Verwilst T, Van Branteghem A, T'Sjoen G, Cools M. 2016. The contribution of the androgen receptor (AR) in human spatial learning and memory: A study in women with complete androgen insensitivity syndrome (CAIS). Horm Behav. 78:121–126.

Phoenix CH, Goy RW, Gerall AA, Young WC. 1959. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. Endocrinology. 65:369–382.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. Proc Natl Acad Sci. 98:676–682.

Raichle ME, Snyder AZ. 2007. A default mode of brain function: A brief history of an evolving idea. Neuroimage. 37:1083–1090.

Sacher J, Neumann J, Okon-Singer H, Gotowiec S, Villringer A. 2013. Sexual dimorphism in the human brain: evidence from neuroimaging. Magn Reson Imaging. 31:366–375.

Schmaal L, Marquand AF, Rhebergen D, van Tol M-J, Ruhé HG, van der Wee NJA, Veltman DJ, Penninx BWJH. 2015. Predicting the Naturalistic Course of Major Depressive Disorder Using Clinical and Multimodal Neuroimaging Information: A Multivariate Pattern Recognition Study. Biol Psychiatry. 78:278–286.

Schrouff J, Rosa MJ, Rondina JM, Marquand AF, Chu C, Ashburner J, Phillips C, Richiardi J, Mourão-Miranda J. 2013. PRoNTo: pattern recognition for neuroimaging toolbox. Neuroinformatics. 11:319–337.

Shepard RN, Metzler J. 1971. Mental rotation of three-dimensional objects. Science (80-). 171:701–703.

Smith SM. 2002. Fast robust automated brain extraction. Hum Brain Mapp. 17:143–155.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage. 23 Suppl 1:S208-19.

Song X-W, Dong Z-Y, Long X-Y, Li S-F, Zuo X-N, Zhu C-Z, He Y, Yan C-G, Zang Y-F. 2011. REST: a toolkit for resting-state functional magnetic resonance imaging data processing. PLoS One. 6:e25031.

Stevens JS, Hamann S. 2012. Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. Neuropsychologia. 50:1578–1593.

Tomasi D, Volkow ND. 2012. Aging and functional brain networks. Mol Psychiatry. 17:549–558.

van Hemmen J, Saris IMJ, Cohen-Kettenis PT, Veltman DJ, Pouwels PJW, Bakker J. 2016. Sex Differences in White Matter Microstructure in the Human Brain Predominantly Reflect Differences in Sex Hormone Exposure. Cereb Cortex. bhw156.

van Hemmen J, Veltman DJ, Hoekzema E, Cohen-Kettenis PT, Dessens AB, Bakker J. 2014. Neural Activation During Mental Rotation in Complete Androgen Insensitivity Syndrome: the Influence of Sex Hormones and Sex Chromosomes. Cereb Cortex.

Van Strien JW. 1992. Classificatie van links- en rechtshandige proefpersonen. / Classification of left-handed and right-handed test subjects. Ned Tijdschr Psychol. 47:88–92.

Wang L, Shen H, Tang F, Zang Y, Hu D. 2012. Combined structural and resting-state functional MRI analysis of sexual dimorphism in the young adult human brain: an MVPA approach. Neuroimage. 61:931–940.

Weissman-Fogel I, Moayedi M, Taylor KS, Pope G, Davis KD. 2010. Cognitive and default-mode resting state networks: Do male and female brains "rest" differently? Hum Brain Mapp. 31:1713–1726.

Ypma RJF, Moseley RL, Holt RJ, Rughooputh N, Floris DL, Chura LR, Spencer MD, Baron-Cohen S, Suckling J, Bullmore ET, Rubinov M. 2016. Default Mode Hypoconnectivity Underlies a Sex-Related Autism Spectrum. Biol Psychiatry Cogn Neurosci Neuroimaging. 1:364–371.

Zahn-Waxler C, Crick NR, Shirtcliff E a, Woods KE. 2006. The origins and development of psychopathology in females and males. Dev Psychopathol Vol 1 Theory method.