



Universiteit  
Leiden  
The Netherlands

## **The Effect of the X Chromosome Imbalance on Sex Differences in Brain Variability and Autism Risk**

Rijn, Lotte van

### **Citation**

Rijn, L. van. (2023). *The Effect of the X Chromosome Imbalance on Sex Differences in Brain Variability and Autism Risk*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master thesis in the Leiden University Student Repository](#)

Downloaded from: <https://hdl.handle.net/1887/3563849>

**Note:** To cite this publication please use the final published version (if applicable).

# The Effect of the X Chromosome Imbalance on Sex Differences in Brain Variability and Autism Risk

Name: Lotte van Rijn

Thesis coördinator: Dr. Lara Wierenga

Date: 16-01-2023

Wordcount: 9.794

### **Abstract**

Autism Spectrum Disorder (ASD) is a highly heritable disorder that is more prevalent in males than females. Genetic mechanisms of the sexual differentiation could therefore contribute to the sex differences in ASD. An unexplored pathway lies in the potential contribution of the direct effect of the X chromosome imbalance. Due to the X chromosome's implications in brain structure, the imbalance in X chromosomes could contribute to recently discovered sex differences in brain structure variance. As males showed larger brain structure variance, they have generally larger brain structure deviations compared to females. If increasing brain structure deviations relate to increased ASD risk, this mechanism could explain why males show higher ASD prevalence compared to females. To assess an influence of the X chromosome imbalance on sex differences in brain structure variance, we assessed the sex difference in brain structure variance in brain regions with a high and low expression of the X chromosome. We then assessed whether larger brain structure deviations related to increased ASD symptoms. This was examined in a population of typically developing individuals aged 7 to 14. The X chromosome expression did not affect sex differences in brain structure variance. Therefore, our results did not support an effect of the X chromosome imbalance on sex differences in brain structure variance. Additionally, brain deviation showed no relationship with ASD symptoms. Our results do not support a contribution of sex differences in brain structure variance to sex differences in ASD. Our findings could suggest potential influences of other downstream effects of the chromosomal imbalance and the high heterogeneity of ASD.

**Layman's abstract**

Autism Spectrum Disorder (ASD) is highly determined by genetic processes and shows a higher prevalence in males than females. This indicates that something genetic that differs between the sexes is involved in determining ASD risk, causing higher ASD risk in males compared to females. The development of sex differences is initiated by a sex difference in the sex chromosomes: where males receive XY chromosomes, females receive XX chromosomes. The imbalance in X chromosomes can directly affect sex differences in X chromosome expression. As females receive two X chromosomes while males receive one, females will compensate for the double dosage by randomly silencing one of the X chromosomes in each cell. As the X chromosome has been implicated to play a role in brain structure, we would expect females to show each of their X chromosomes to be expressed in 50% of brain cells, while males will always show their one X chromosome expressed in all cells. This mechanism could explain why males have reported larger brain structure variance than females, as males will not be able to compensate for an X chromosome which codes for a small or large brain structure. This mechanism could contribute to sex differences in ASD prevalence by increasing ASD risk for individuals with more deviating brain structures, thereby increasing ASD risk in males. We tested this by assessing the effect of the X chromosome imbalance on sex differences in brain structure variance, and by assessing the relationship between brain structure deviations and ASD symptoms. Our results did not support the hypothesized relationships. Other mechanisms might be involved in determining sex differences in brain structure and sex differences in ASD prevalence, thereby explaining why no associations were found in the current study.

### **Introduction**

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that is four times more likely to occur in males compared to females (Fombonne, 2005). As ASD has shown high heritability rates (Tick et al., 2015), genetic mechanisms involved in the sexual differentiation are proposed to increase the risk for ASD in males (Ferri, Abel & Brodtkin, 2018; see also Werling & Geschwind, 2013). All mechanisms that differ between the sexes however originate from one essential sex difference: an imbalance in sex chromosomes (Arnold, 2012). Where males receive XY chromosomes, females receive XX chromosomes. The imbalance in Y chromosomes causes the male phenotype to develop differently than the female phenotype by initiating downstream effects that are male-specific. The imbalance in X chromosomes however initiates mechanisms to compensate for the double dosage of X chromosomes in females (Arnold, 2012). The imbalance in X chromosomes thereby directly affects the expression of X chromosomes differently in males and females. As the X chromosome has been implicated in determining brain structure, having a different expression of the X chromosome in males and females could cause sex differences in brain structure (Mallard et al., 2021; Nguyen & Disteché, 2006; Raznahan & Disteché, 2021). Animal studies have shown that a direct effect of the chromosomal balance can influence phenotypical sex differences in brain structure, independent of other mechanisms (Cabrera Zapata et al., 2022). This provides a potential pathway for the direct effect of the X chromosome imbalance to influence sex differences in brain structure that could relate to a sex difference in ASD risk. Recent studies have found consistently larger brain structure variance in males compared to females, posing a potential sex difference that could relate to sex differences in ASD risk (Ritchie et al., 2018; Wierenga et al., 2018, 2019, 2020). However, this direct effect of the X chromosome imbalance on sex differences in the human brain and their potential link to sex differences in ASD risk remains an underexplored path. Therefore, the current study will examine the direct effect of the X chromosome imbalance on sex differences in brain structure variance and their potential link to sex differences in ASD risk.

#### **The direct effect of the X chromosome imbalance**

The chromosomal imbalance holds several important roles in the sexual differentiation (Arnold, 2012). One essential role relies on sex specific effects initiated by sex specific genes. The Y chromosome contains the SRY gene, which initiates the formation of the testes and the consequent secretion of testosterone. As only males receive the Y chromosome, and thus the SRY gene, this effect is sex specific for males. Early research on sex differences focused on the role of the gonads and their subsequent sex hormones that cause the male phenotype to differ from the female phenotype (Herting et al. 2014; Schulz & Sisk, 2016; Wierenga et al., 2018). However, theories based on sex hormones do not seem to provide a comprehensive explanation of sex differences in ASD. The Extreme Male Brain (EMB) hypothesis suggests that males are at a higher risk for developing ASD due to a prenatal peak in testosterone that causes masculinization of brain and behavior (Baron-

Cohen, 2011). However, this theory does not hold in populations of male transgenders, who have developed testes and subsequent testosterone secretion, but do not show a masculinization of behavior and increased ASD prevalence ((Pasterski, Gilligan & Curtis, 2013; see also Kurth et al., 2022 for preliminary evidence on brain masculinization). This suggests that downstream effects of the sexual differentiation initiated by the Y chromosome imbalance cannot fully explain the sex differences in ASD. We might be missing an essential influence here: The direct effect of the X chromosome imbalance. The imbalance in X chromosomes between males and females causes an unequal expression of X chromosomes and thus of X linked genes. Females therefore compensate for their double dosage of X chromosomes through X-inactivation, a mechanism that turns off one of the X chromosomes in each cell, so both females and males will have one X chromosome expressed in their tissue (Arnold, 2012). As each cell turns off one of the X chromosomes, females will have each of their X chromosomes expressed in 50% of cells, while males will have their one X chromosome expressed in 100% of cells. Genes on the X chromosome can thus be partly turned off in females, while they will always be expressed in males. This mechanism could put males at risk for developing disorders if X linked genes code for a certain disorder. For example, one gene on the X chromosome (FMR1) can code for a disorder called Fragile X syndrome (FXS; Marco, 2006). The expression of FXS can therefore be mediated in females through X-inactivation. If more cells silence the X chromosome that contains the genetic code for FXS than the other X chromosome that does not contain the code, then females will show less signs of the disease. As males do not initiate X-inactivation, they cannot compensate for the expression of FXS. Males therefore more often show FXS, and additionally show more extreme expressions of the disorder. Although the genetic code for ASD is not limited to a single gene or the X chromosome alone, the X chromosome could contain genes that affect the risk for ASD. As a result of X-inactivation, males might not be able to compensate for risk factors of ASD that are coded for on the X chromosome.

#### **The direct effect of the X chromosome imbalance and sex differences in brain structure variability**

A pathway through which the X chromosome could influence ASD risk is by affecting the organization of the brain (Arnold, 2012; Nguyen & Disteche, 2006; Raznahan & Disteche, 2021). The X chromosome is linked to brain structure size as an increase in X chromosomes is associated with a decrease in brain volume, as females have generally smaller brains than males (Raznahan & Disteche, 2021). Specific genes on the sex chromosomes have also been linked to sex differences in brain volume (Liu et al., 2020). These associations between the X chromosome and brain structure could result from the relatively large influence the X chromosome has in the brain when compared to other autosomal chromosomes (Nguyen & Disteche, 2006). The X chromosome thus seems to play an important role in determining brain structure. Following from the X chromosome imbalance, we would therefore expect differences in brain structure between males and females. Wierenga et al.

(2018, 2020) showed that structural brain differences between the sexes follow a pattern expected from X-inactivation. As males will always have the same X chromosome expressed, brain structure determination within the male brain would be more similar within regions when compared to females, where two different X chromosomes are expressed. This pattern was confirmed by more homogeneity in interregional brain structure in males compared to females, reflected in a higher interregional correlation in males (Wierenga et al., 2018, 2020). These findings support the hypothesis that the X chromosome imbalance can directly affect sex differences in brain structure. Based on the pattern of X-inactivation in cells, the female brain structure would converge towards a mean brain structure coded for by the two X chromosomes. As males will express the brain structure coded for by their one X chromosome, they could show more varying expression of brain structure compared to females. The imbalance in X chromosomes could thereby cause a sex difference in brain structure variance, while both sexes could still show the same mean. Studies on brain structure have indeed shown a consistently higher brain structure variance in males compared to females, while mean differences between the sexes do not seem consistent (Ritchie et al., 2018; Wierenga et al., 2018, 2019, 2020). Males showed larger variance at both ends of the distribution, reporting smaller brains compared to females with small brains, and larger brains compared to females with large brains. A recent study by Mallard et al. (2021) showed that the expression of the X chromosome is associated with brain structure variance. The direct effect of the X chromosome imbalance might thus be observed in the brain by looking at sex differences in brain structure variance, providing a pathway through which sex differences in ASD risk could be influenced by the X chromosome imbalance.

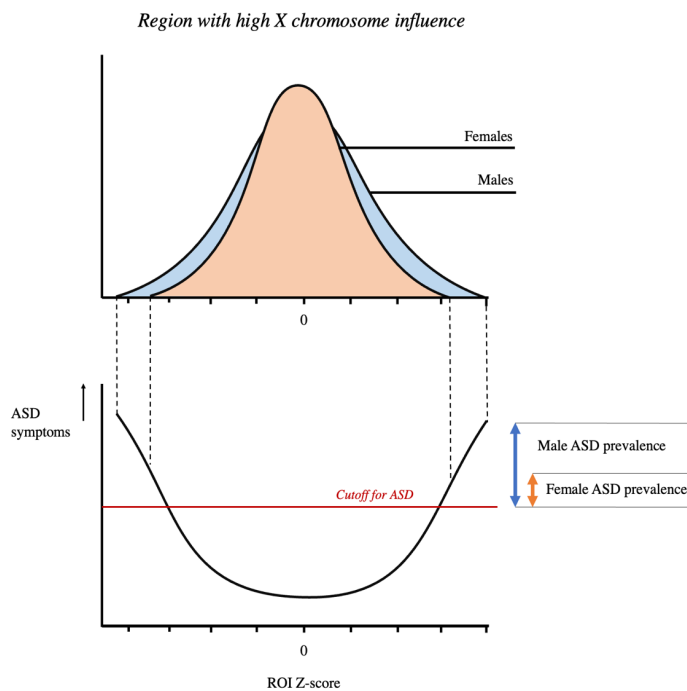
#### **Sex differences in brain structure variability and sex differences in ASD**

This sex difference in brain structure variance following from the X chromosome imbalance could explain why males show a higher prevalence of ASD compared to females. The larger brain structure variance in males presents two distinct profiles that could relate to increased ASD risk: a positive and a negative deviation from the mean (figure 1). Similar profiles have been observed for ASD individuals when compared to typically developing individuals (Zabihi et al., 2019). ASD individuals showed relatively larger deviations from the mean, thus showing both smaller and larger brain structures when compared to typically developing individuals. This aligns with the hypothesis that a larger deviation from the mean brain structure could relate to increased risk for ASD (Wierenga et al., 2018, 2019, 2020). Genetic evidence also provides a common ground for brain structure and ASD risk, as genes that code for extreme brain structure expressions are also involved in ASD risk (Li et al., 2017). Genes for microcephaly (small brains, thus negative deviations) and macrocephaly (large brains, thus positive deviations) both associate with ASD risk, confirming that the two distinct neuroanatomical profiles can be related to increased ASD risk. These profiles align with the findings of increased brain structure variance in males compared to females. As males are more widely distributed around the mean, more males will show larger deviations when compared to females.

Therefore, males could be more at risk for ASD due to their increased brain structure variance compared to females, following the effect of the X chromosome imbalance.

**Figure 1**

*Sex differences in ASD risk according to sex differences in brain structure variance*



*Note.* Sex differences in ASD risk according to the hypothesis by Wierenga et al. (2018, 2019, 2020) based on findings of higher brain structure variance in males compared to females. A risk for developing ASD could be represented by an extreme brain structure, thus a high deviation from the mean brain structure. As males would have more variance in brain structure, they show larger deviations from the mean brain structure compared to females. These larger deviations could pose a risk for developing ASD, thereby explaining the higher prevalence of ASD among males. We expect sex differences in brain structure variance to occur in regions with a high expression of the X chromosome, as the X chromosome imbalance would cause a different expression of the X chromosome in males compared to females.



### Research questions and hypotheses

We hypothesize that the direct effect of the X chromosome imbalance can cause males to show larger brain structure variance compared to females. Additionally, we hypothesize that larger brain structure deviations relate to increased ASD risk. This would increase ASD risk in males as males have shown larger deviations from the mean brain structure compared to females (Wierenga et al., 2018, 2019, 2020). The direct effect of the X chromosome imbalance could thereby contribute to the higher prevalence of ASD among males. We will examine our hypotheses by assessing whether: 1) X chromosome effects on the brain are related to increased brain structure variance in males compared to females, and 2) a larger deviance of brain structure relates to increased reporting of ASD symptoms. To examine hypothesis 1, we will assess variance ratios (male: female) in structural brain regions where a high expression of the X chromosome has been found (enrichment; Mallard et al., 2021). We will compare these variance ratios to variance ratios for regions with a low X chromosome expression (depletion) to assess the influence of X chromosome effects on sex differences in brain structure variance. To examine hypothesis 2, we will assess the relationship between brain structure deviation (Z-score) and ASD symptoms. Brain structure will be measured as Surface Area (SA), as X chromosome enriched and depleted regions were identified within SA (Mallard et al., 2021). ASD symptoms will be measured using the Social Responsiveness Scale (SRS) and a sensorimotor synchronization task (SMS; Constantino & Gruber, 2011; Schultz & Vugt, 2016; Vugt, 2020). The SRS measures parent reported social communicative difficulties often found for individuals with ASD (APA, 2013). The SMS provides a measure of sensorimotor synchronization that has been found to differ for ASD individuals (McNaughton & Redcay, 2020), and was therefore proposed as an objective measurement of ASD (Peper et al., 2016). We will examine the first hypothesis in a sample of typically developing same-sex twins who have T1 weighted imaging scans at the age of 7-9. The second hypothesis will be examined in a subset of this group who have corresponding measurements of ASD symptoms. Due to the broad autism phenotype, subclinical symptoms are also prevalent within the typically developing population, allowing us to infer a potential relationship between brain structure deviance and ASD symptoms (Dawson et al., 2002; Losh et al., 2009; Sucksmith & Hoekstra, 2011; see also Hyseni et al., 2018). Additionally, subclinical symptoms are assumed to have emerged within our sample, as ASD diagnosis is usually diagnosed around the age of 5 (van 't Hof, 2021). If our hypotheses can be confirmed, the direct effect of the X chromosome imbalance would provide a phenotypical expression of ASD risk that has already be observed from the age of 3 (Wierenga et al., 2018). Measurements of brain structure in combination with normative modelling could function as early biomarkers for ASD. This would allow for earlier detection of ASD compared to the mean age of diagnosis now (60.48 months; 5.03 years), which cannot be achieved using screening methods based on clinical symptoms (Hyman et al., 2020; van 't Hof, 2021).

## Methods

### Participants

Participants in this study took part in the Leiden Consortium on Individual Development (LCID), which is a sequential longitudinal twin study including children aged 3-13 (Crone et al., 2020). The study was approved by the Dutch Central committee on Research Involving Human Subjects (CCMO). Twins were recruited through municipality records in the Netherlands and were eligible if both parents were fluent in Dutch and both their parents and grandparents originated from Europa (Euser et al., 2016). Children were excluded from participation if they reported a congenital disability, psychological disorder, chronic illness, hereditary disease, hearing impairment, or visual impairment that could influence their behavioral performance. Additionally, if children were diagnosed with an intellectual disability ( $IQ < 70$ ) they were also excluded from participation. Exclusion criteria were checked with parents during a phone call, after which children and parents were invited for a home visit in the case of inclusion. Informed consent was given by both parents. Data collection was split in two sequential cohorts: an early childhood cohort (ECC) that initially included children aged 3-4 years old and followed them until they were 8-9 years old, and a middle child cohort (MCC) that initially included children aged 7-8 years old and followed them until the age of 11-14 years old. Both cohorts included six waves where home visits alternated with lab visits (see Crone et al., 2020 for an overview per cohort). Home visits mainly consisted of performing behavioral tasks, while lab visits included either EEG measurements or MRI measurements starting from the age of 7.

Due to the nature of the analyses (variance ratios and multilevel analyses to account for dependency) we were unable to perform a power analysis to calculate the required sample size. We therefore decided to include MRI data for an age group that was available for both cohorts (7-9; ECC wave 5 and MCC wave 1), allowing us to increase our sample size and consequently have a higher statistical power for our analyses. As the calculation of variance ratios does not allow for dependency induced by twins, variance ratios were calculated for two groups: the first containing one individual from each twin indicated with child number 1 ( $N = 337$ ), and the other containing the other individual from each twin indicated with child number 2 ( $N = 337$ ). We then calculated the mean ratios over both groups, resulting in an inclusion of 674 participants for the calculation of the variance ratios ( $M$  age = 7.99, age range = 7.02 – 9.68, 53.8% females, see table 1).

For the analyses of brain deviation and ASD symptoms (symptomatic analyses) we included different subsets of the sample for which either SRS or SMS data and corresponding MRI data for the age group (7-9) was available. As not all individuals performed both the SRS and SMS, using separate datasets for the SRS and SMS allowed us to include most available data for each measure. Additionally, including MRI data for the age group 7-9 allowed us to include behavioral data from both cohorts to increase the sample size and to infer potential relationships between the variance ratios and the symptomatic analyses. For the SRS analyses, behavioral data from the ECC wave 5 and the MCC wave 3 was included, yielding a total of 524 participants for the symptomatic analyses with the

SRS ( $M$  age = 9.13, age range = 7.02 – 11.67, 55.7% females, see table 1). For the SMS data, behavioral data from the MCC wave 5 was included as the SMS was not performed in the ECC. This yielded a total of 256 participants for the symptomatic analyses with the SMS ( $M$  age = 12.43, age range = 11.25 – 14.11, 52.6% females, see table 1).

Although our data included typically developing children without a diagnosis, one individual had a family member diagnosed with autism (included in all three datasets), and one individual was being examine for autism at the time of measurement (included in the datasets for variance ratios and SRS analyses).

**Table 1**

*Demographics*

Data	Sex	Zygoty	$N$	Mean age	SD	Age range	Handedness
Variance ratio data	53.8% F	383 MZ	674	7.99	0.65	7.02 - 9.68	85.3% (R), 11.2% (L)
	F	200 MZ	361	8.00	0.67	7.02 - 9.68	
	M	183 MZ	313	7.98	0.63	7.02 - 9.39	
SRS data	55.7% F	294 MZ	524	9.13	1.17	7.02 - 11.67	84.9% (R), 12.3% (L)
	F	156 MZ	292	9.14	1.17	7.02 - 11.67	
	M	138 MZ	232	9.12	1.16	7.03 - 11.23	
SMS data	52.6% F	140 MZ	249	12.43	0.75	11.25 - 14.11	84.7% (R), 13.3% (L)
	F	76 MZ	131	12.40	0.78	11.25 - 14.11	
	M	64 MZ	118	12.47	0.73	11.25 - 13.99	

*Note.* Demographics for the samples that were included in the current study. F = female, M = male, R = right-handed, L = left-handed.

## **Symptomatic measures**

### **SRS**

One of the core symptoms of ASD is difficulty with social communication and interaction, which can be indicated by a lack of reciprocal behavior in interaction (APA, 2013). To measure the level of social communicative skills as a reflection of ASD symptoms, we used the short version of the Social Responsiveness Scale (SRS; Constantino & Gruber, 2011; Roeyers et al., 2011). The short version of the SRS consists of 21 items that are answered on a 3-point Likert scale ranging from 0 (never true) to 3 (almost always true) by the primary parent. All items are aimed at assessing a child's ability for reciprocal social interactions through dimensions of social deficits, language deficits and repetitive and stereotypical behavior that are all characteristics of ASD (Constantino et al., 2003). This is achieved by dividing items over five subscales: social awareness, social information processing, capacity for reciprocal social responses, social use of language and stereotypical or repetitive behaviors. The SRS will yield an overall score indicating a general indication of social deficits related to ASD. A higher score represents more social deficits. The SRS hereby allows for measuring both the intensity and type of social problem behavior in children aged 4 to 18 years old (Roeyers et al., 2011).

### **SMS**

An important aspect of ASD individuals' difficulties with social communicative skills is their difficulty with nonverbal communication such as eye contact, facial expressions, and body language (APA, 2013; Brezis et al., 2017; Chita-Tegmark, 2016; Coll et al., 2020; Fitzpatrick et al., 2017; Kaur, Srinivasan & Bhar, 2017; Trevisan, Hoskyn & Birmingham, 2018). Due to their different nonverbal communication, they show less motor synchronization in interaction (Coll et al., 2020). Synchronization has been proposed to underlie the social communicative development as it fosters interaction through the embodiment of others' emotions thereby adding to social connectedness (Crompton et al., 2020; Eigsti, 2013; Morrison et al., 2019; Prochazkova & Kret, 2019). As ASD individuals already show deficits in mimicry behavior early in life, before the onset of clinical symptoms (Carter et al., 2005; van t' Hof, 2021), the lack of sensorimotor synchronization could thus be an essential contributor to the social communicative difficulties experienced by ASD individuals (McNaughton & Redcay, 2020). Therefore, sensorimotor synchronization has been suggested as an important objective measurement of ASD symptoms (Peper et al., 2016). To take this into account, we included sensorimotor synchronization as a measure of ASD symptoms. This was measured with a sensorimotor task using a computerized tapping paradigm. The Teensytap framework combines a Teensy microcontroller with a computer, an audio board and a force-sensitive resistor (Musical Instrument Digital Interface (MIDI); Schultz & Vugt, 2016; Vugt, 2020). The computer uses a python user interface so the metronome rhythm can be chosen, which is then delivered to the participant through the Teensy and audio board using headphones. The rhythm can then be indicated with finger tapping using the MIDI device, which sends the timing and pressure of the finger tapping back to the

computer for storage. Audio feedback for the responses can also be indicated on the python interface and delivered through the audio board. The task is divided into three parts: the first part required participants to tap along a regular rhythm indicated by the metronome. Participants practiced one round after which they performed two official rounds with different paces. The second part required participants to tap along a pace that deviates from the pace indicated by the metronome (anti-phase tapping). Participants again performed one practice round after which two official rounds with different paces were performed. In the first and second part, the tapping tempo was indicated for the first eight beats in every round to avoid misunderstanding of the task. The last part required participants to tap along a pace indicated by music. Participants were asked to tap along four different songs that were highly ranked in the Dutch top 40 list of 2020. In all parts, participants were instructed to tap as if they would be clapping to music in every part of the task. The task will yield a measure of sensorimotor synchronization (SMS) indicated from 0 (poor SMS) to 1 (perfect SMS; Schultz & Vugt, 2016). The average score for all three parts will be used in the current study. To stay consistent with the other symptomatic analyses for SRS, the SMS scores will be reversed coded, so a higher score will reflect a higher report of sensorimotor synchronization difficulty.

## Brain measures

### Data acquisition

MRI scans were acquired on a Philips Achieva 3.0 T MRI system with a standard whole-head coil at the Leiden University Medical Centre. Participants were first familiarized with the scanner by a practice session in the mock scanner to reduce scanner related stress (Achterberg & van der Meulen, 2019). During the actual scanning session, head motion was limited with the use of foam inserts around the children's heads, and they were asked to focus on a movie that was displayed on a screen visible through a mirror. High resolution T1 weighted images were acquired (FOV) = 224 (ap) x 177 (rl) x 168 (fh); flip angle (FA) =  $8^{\circ}$ ; voxel size = 0.875 x 0.875 x 0.875 mm). The duration of the T1-weighted scan was 296 seconds. Scan quality was checked for motion artifacts during the scan, where scans were repeated if motion was detected.

### Data processing and quality control

Data processing was performed in FreeSurfer 7.1.1, which can be used for cortical surface reconstruction and volumetric segmentation (<http://surfer.nmr.mgh.harvard.edu/>). The procedure for the structural brain reconstruction of the T1-weighted scans was performed using the longitudinal processing pipeline described in Reuter et al. (2012). The longitudinal pipeline was designed to decrease processing bias that could be introduced by treating the baseline image differently from other images. To prevent this, Reuter et al. (2012) introduces a temporally unbiased within subject template representing a subject's average anatomy over time. Using the within subject template can decrease processing bias as opposed to using one image as the baseline. After creating the within subject

Met opmerkingen [WL(1)]: Ik zou dit noemen voor de structural brain measures.

Met opmerkingen [WL(2)]: Hier ook nog iets noemen over de procedure dat de scan direct werd gecheckt en over gedaan als er beweging te zien was. En dat er een mock-scanner protocol was om

Met opmerkingen [RLv(3R2)]: In eiser et al denk ik

template, each timepoint is processed longitudinally using the template. This step consists of motion correction, skull stripping, talairach transformation, (spherical) atlas registration and parcellation (Reuter et al., 2012).

Images Parcellation of the cortex was based on the Destrieux atlas (Destrieux et al., 2010). The regions of interest (ROIs; X chromosome enriched and depleted regions) have however been identified by Mallard et al. (2021) using the HCP atlas (Glasser et al., 2016). The HCP atlas identifies a total of 180 cortical areas (per hemisphere) which were all analyzed for X chromosome enrichment or depletion (Mallard et al., 2021). For organizational reasons, these 180 cortical areas can be grouped within 22 bigger overarching regions based on architecture, task-fMRI profiles, and functional connectivity (see supplementary material Glasser et al., 2016). The Destrieux atlas identifies 74 cortical areas, causing cortical areas in the HCP atlas to be relatively smaller compared to cortical areas in the Destrieux atlas (Destrieux et al., 2010). As Mallard et al. (2021) identified functional networks of depleted and enriched regions that correspond to the functions of the overarching regions for the HCP atlas, we focused on comparing cortical areas in these overarching regions to cortical areas in the Destrieux atlas. We first determined what overarching regions to focus on, by assessing the percentage of enriched or depleted cortical areas within each overarching region regarding SA (see appendix A). We chose to include SA only as SA has a larger genetic component compared to other brain measures (CT or CV; Blakemore, 2012; Ecker, 2017; Panizzon et al., 2009), and most X chromosome effects were identified in SA (Mallard et al., 2021). If at least 30% of cortical areas that were grouped within an overarching region were considered enriched or depleted, the overarching region was included in the current study. We chose this cutoff of 30% as this is well above the average percentage of enriched or depleted cortical areas in overarching regions (16,84%), and additionally allowed for a sufficient number of regions to be included for our analyses. The overarching regions that were included correspond to the functional networks identified by Mallard et al., (2021): the visual system consisting of both the primary and early visual cortex showed a high percentage of depleted areas, and a somatosensory motor system consisting of the premotor cortex, primary motor and somatosensory cortex, the paracentral lobular and mid cingulate cortex and posterior opercular cortex that showed a high percentage of enriched areas (Glasser et al., 2016; Mallard et al., 2021). As the aim of the current study is to assess whether a difference in variance ratio exists between depleted and enriched regions, but not to provide an extensive overview of all enriched and depleted regions, we will focus on these two identified networks only. After choosing which regions are considered enriched and depleted, we identified corresponding ROIs in the Destrieux atlas for each overarching region using visual inspection and description of the overarching regions (see table 2; Destrieux et al., 2010; see also Glasser et al., 2016). This yielded five ROIs representing the depleted network, and thirteen ROIs representing the enriched network. As both hemispheres were included, this yielded a total of 36 ROIs for the current study.

**Table 2**

*ROIs from the Destrieux atlas (Destrieux et al., 2010) that correspond to identified regions by Mallard et al. (2021) based on the atlas by Glasser et al. (2016)*

<b>Status</b>	<b>Region by Glasser et al. (2016)</b>	<b>Percentage significant cortical areas</b>	<b>Corresponding ROIs in the Destrieux atlas</b>
Depleted	Primary visual cortex (1)	100%	Pole_Occipital (42) G_cuneus (11) G_oc-temp_med-Lingual (22) S_calcarine (44)
Depleted	Early visual cortex (2)	50%	S_parieto_occipital (65)
Enriched	Somatosensory and motor cortex (6)	30%	S_precentral-inf-part (68) S_precentral-sup-part (69) G_precentral (29) S_central (45) G_and_S_subcentral (4) G_postcentral (28) S_postcentral (67)
Enriched	Paracentral lobular and mid cingulate cortex (7)	50%	G_and_S_paracentral (3) G_ad_S_cingul-Mid-Ant (7) G_ad_S_cingul-Mid-Post (8)
Enriched	Premotor cortex (8)	42%	G_front_middle (15) G_front_inf-Opercular (12)

---

Enriched	Posterior opercular cortex (9)	33,3%	Lat_Fis-post (41)
----------	--------------------------------	-------	-------------------

---

*Note.* The numbers indicate the ROI according to the original atlas the ROIs are derived from.

### ***Missing data, outliers and multiple testing***

T1-weighted scans with insufficient quality or anomalous findings were excluded (25 individuals for wave 5 of the ECC, 92 for wave 1 of the MCC). To check for scan quality, the hybrid manual automatic quality assessment tool Qoala-T was used (Klapwijk et al., 2019). This tool consists of a supervised-learning model that can classify the quality of scans by training it on a subset of data that has been checked manually. It automatically classifies processed scans into two categories: the quality of the scan is sufficient for inclusion (relating to a manual rating that is either excellent, good, or doubtful), or the quality of the scan is insufficient for inclusion (relating to a manual rating that is failed). L-CID data was used as a testing set, thus all scans were checked both manually according to the supplementary material by Klapwijk et al., (2019) and with the Qoala-T tool.

As we were specifically interested in large structural brain deviations, no outliers were excluded from the analysis. Missing data was removed list-wise as analyses were not fit for missing data. In total, 697 individuals had MRI data, of which 23 individuals showed missing values in one of the regions that we included. These individuals were therefore excluded from analyses, yielding the sample size of 674 for the variance ratios (2 x 337). A total of 588 individuals had SRS scores, of which 540 had corresponding MRI data. 16 individuals showed missing values in other variables, and were therefore also excluded from analyses, yielding a sample size of 524 for the SRS analyses. A total of 311 individuals had performed the SMS task, of which 256 individuals had corresponding MRI data. 10 individuals showed missing values in other variables and were therefore excluding, yielding the sample size of 246 for the SMS analyses.

To correct for multiple comparisons, we applied the Hochberg procedure for multiple tests, which allows for control of the family wise error rate (Hochberg, 1988). We applied this to the final set of variance ratios ( $N = 36$ ), and to all estimates for the final models for all regions ( $N = 36$  for the SRS,  $N = 36$  for the SMS). As this is a rather stringent correction as we are controlling for the family wise error rate for many tests (as opposed to controlling the false-discovery rate or applying it per model), we will show results both before and after multiple comparisons correction to increase transparency.



## Analyses

All analyses were performed using R statistical software (v4.1.2; R Core Team 2021).

### *X chromosome effects on sex differences in brain structure variability*

Variance differences between males and females were assessed with variance ratios. As SA has shown mean differences between the sexes and variations with age, we controlled for mean sex differences and age using a random forest regression model (Breiman, 2001; Wierenga et al., 2014). The model yielded a predicted value of SA for each ROI per individual ( $y_i$ ). We calculated residuals based on the predicted values and the observed values ( $y_i$ ) yielding a new dataset containing all residuals.

$$r_i = y_i - \hat{y}_i$$

We then calculated variance ratios for each ROI by dividing the standard deviation (SD) of males by the SD for females. This yielded a test statistic  $T$  indicating the variance ratio for each ROI.  $T$  values were log transformed, so a positive value indicates larger variance in males compared to females and a negative value indicates larger variance in females compared to males.

$$T = SD_{\text{males}} / SD_{\text{females}}$$

For every ROI, a permutation test with 10,000 ( $B$ ) permutations was conducted, yielding a total of 10,000  $T$  values for every ROI ( $T_b$ ). This allowed us to calculate a  $P$ -value by comparing each  $T_b$  to  $T$ , yielding a value of 1 if  $T_b > T$  and a value of 0 if  $T_b < T$ . By summing all outcomes and dividing the total by  $B$ , this provided a percentage of permuted test statistics  $T_b$  that is greater than the observed  $T$  ( $P_k$ ) for every ROI ( $k$ ). This resulted in a variance ratio for every ROI overall ( $T$ ) with an accompanying  $P$ -value ( $P_k$ ) to indicate the significance of this finding.

$$P_k = \sum_{b=1}^B I(T_b - T) / B$$

Due to dependency in the data, we first calculated the variance ratios for one individual from each twin (child number 1). We then repeated this for the other individual (child number 2) and calculated the mean variance ratios for the two groups. This yielded a total of 36 average variance ratios with accompanying  $P$ -values. As calculating 36  $P$ -values increases chances of a type 1 error, we applied multiple comparisons correction. We then assessed the influence of the X chromosome by

comparing the average variance ratios for the depleted regions ( $N = 10$ ) to the average variance ratios for enriched regions ( $N = 26$ ) using a Welch t-test to account for unequal sample sizes and unequal variances between groups. To confirm that X chromosome effects relate to larger brain structure variance in males compared to females, we expect variance ratios to show a significantly higher value (indicating larger male variance) in X chromosome enriched regions compared to X chromosome depleted regions.

### *The relationship between sex differences in brain variability and ASD symptoms*

To examine whether increasing brain deviations are associated with increased ASD risk, we assessed the relationship between structural brain deviations and ASD symptoms in the same set of regions. Brain deviation was measured using Z-scores for each ROI ( $N = 36$ ) and ASD symptoms were measured using SRS scores or SMS scores. As individuals are nested within twins, a multilevel model was applied to account for within twin dependence using a random effect of family (TwinNumber). Z-scores were calculated for males and females separately for every ROI, as males generally show larger brain volumes compared to females. The interaction of ROI Z-score and sex was added to allow the relationship between ROI Z-score and ASD symptoms differs for males and females. Additionally, zygosity and sex were added as covariates. As we hypothesized the relationship between ROI Z-score and ASD symptoms to show a U-shaped relationship, a linear model was compared to a quadratic model for each ROI to assess whether the quadratic model indeed shows a better fit. Models were compared using the Bayesian Information Criterion (BIC) as an indication of model fit, as the BIC adds a penalty for the inclusion of two additional parameters in the quadratic models as both the first order and second order polynomial are added. We additionally examined the parameter estimates to see whether a U-shaped relationship, or convex curve, could be identified (indicated by a significant positive estimate for a quadratic coefficient). We used the lme4 R package to build models using full maximum likelihood (FML) as we are interested in testing fixed effects, and the lmerTest R package for significance testing of parameters (Bates et al., 2015; Kuznetsova, Brockhoff & Christensen, 2017).

$$* Y_{ij} = \gamma_{00} + \gamma_{10}(ROI\ Z - score)_{ij} + \gamma_{01}(Zygosity)_j + \gamma_{02}(Sex)_{ij} + \gamma_{12}(ROI\ Z - score)_{ij} \\ * (Sex)_{ij} + \mathcal{U}_{0j} + \mathcal{U}_{1j}(TwinNumber)_{ij} + e_{ij}$$

In the linear model,  $Y_{ij}$  denotes the symptomatic score of individual  $i$  in family  $j$  (either SRS score or SMS score), which is predicted by level 1 variable ROI Z-score, level 2 variables zygosity and sex and the interaction of ROI Z-score with sex. If the linear model provides the best fit, the relationship between symptomatic measures and brain deviations could be linear, providing evidence against our hypothesis that both low and high brain size deviations increase susceptibility for ASD.

$$* Y_{ij} = \gamma_{00} + \gamma_{10}(ROI\ Z - score)^2_{ij} + \gamma_{01}(Zygosity)_j + \gamma_{02}(Sex)_{ij} + \gamma_{12}(ROI\ Z - score)^2_{ij} \\ * (Sex)_{ij} + \mathcal{U}_{0j} + \mathcal{U}_{1j}(TwinNumber)_{ij} + e_{ij}$$

In the quadratic model a quadratic term was added for ROI Z-score, allowing a quadratic relationship between ASD symptoms and the brain deviations (Z-scores). If this model provides a better fit, our hypothesis could be confirmed if a U-shaped relationship is found between ASD symptoms and structural brain deviation.

This method was applied to both the SRS scores and SMS scores separately for each of the 36 ROIs, yielding a total of 72 comparisons of the linear and quadratic model. Multiple comparisons correction was applied per measure, so for SRS and SMS separately. Due to the large quantity of data, we created a variable indicating the best model fit by calculating a difference value (quadratic model BIC < linear model BIC), indicating 'TRUE' if the quadratic model showed a better fit, indicated by a lower BIC value. Assumptions for multilevel models (linearity, normality, homoscedasticity) were not checked due to the expectation that data is non-linear and non-homogenous, as we hypothesized that the relationship between ROI Z-score and ASD symptoms is quadratic, and we included monozygotic twins and dizygotic twins in our model which could cause heteroscedasticity.

## Results

### X chromosome effects on sex differences in brain structure variability

For every ROI ( $N = 36$ ) with a high and low X chromosome expression (enrichment and depletion) an average variance ratio was calculated along with a  $P$ -value indicating the significance of the variance difference between the sexes. All results have been included in table 3. The average variance ratios were calculated over two groups, for which we additionally included the results in appendix C.

Almost all average variance ratios showed a positive value, indicating larger variance in males compared to females, except for the left inferior part of the precentral sulcus (lh\_S\_precentral.inf.part\_area), the left precentral gyrus (lh\_G\_precentral\_area), and the right central sulcus (rh\_S\_central\_area) which showed larger variance for females compared to males. Before multiple comparisons were applied, a total of ten regions showed a significantly larger variance in males compared to females (table 3, figure 2). Of these regions, three were depleted regions (3/10) while seven were enriched regions (7/26). However, none of the regions remained significant after applying multiple comparisons correction.

To test whether X chromosome effects relate to larger brain structure variance in males compared to females, we compared variance ratios for enriched regions to variance ratios for depleted regions with a Welch  $t$ -test. We found no significant differences between variance ratios for enriched regions ( $N = 26$ ,  $M = 0.21$ ,  $SD = 0.13$ ) and variance ratios for depleted regions ( $N = 10$ ,  $M = 0.22$ ,  $SD = 0.10$ ) with  $t(20.44) = 0.11$ ,  $p = .915$ .

**Table 3**

*P-values and adjusted P-values for average variance ratios of every ROI*

Status	ROI	Average variance ratio	$P$ -value	Adjusted $P$ -value
Depleted	rh_Pole_occipital_area	0.21	.104	.599
Depleted	lh_Pole_occipital_area	0.13	.171	.599
Depleted	rh_G_cuneus_area	0.08	.350	.599
Depleted	lh_G_cuneus_area	0.10	.293	.599
Depleted	rh_G_oc.temp_med.Lingual_area	0.50	.064	.599
Depleted	lh_G_oc.temp_med.Lingual_area	0.22	.076	.599
Depleted	rh_S_calcarine_area	0.27	<b>.037</b>	.599

Met opmerkingen [lvR4]: Preliminary table

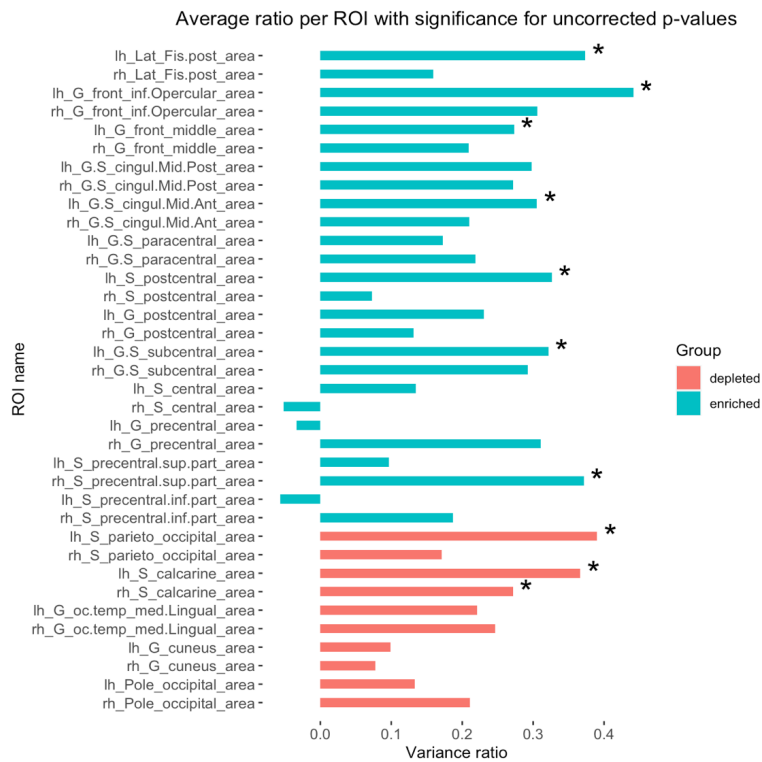
Depleted	lh_S_calcarine_area	0.37	<b>.004</b>	.140
Depleted	rh_S_parieto_occipital_area	0.17	.119	.599
Depleted	lh_S_parieto_occipital_area	0.39	<b>.004</b>	.147
Enriched	rh_S_precentral.inf.part_area	0.19	.166	.599
Enriched	lh_S_precentral.inf.part_area	-0.06	.587	.599
Enriched	rh_S_precentral.sup.part_area	0.37	<b>.006</b>	.187
Enriched	lh_S_precentral.sup.part_area	0.10	.315	.591
Enriched	rh_G_precentral_area	0.31	.061	.599
Enriched	lh_G_precentral_area	-0.03	.567	.599
Enriched	rh_S_central_area	-0.05	.599	.599
Enriched	lh_S_central_area	0.13	.280	.599
Enriched	rh_G.S_subcentral_area	0.29	.093	.599
Enriched	lh_G.S_subcentral_area	0.32	<b>.026</b>	.599
Enriched	rh_G_postcentral_area	0.13	.230	.599
Enriched	lh_G_postcentral_area	0.23	.136	.599
Enriched	rh_S_postcentral_area	0.07	.330	.599
Enriched	lh_S_postcentral_area	0.33	<b>.039</b>	.599
Enriched	rh_G.S_paracentral_area	0.22	.312	.599
Enriched	lh_G.S_paracentral_area	0.17	.163	.599
Enriched	rh_G.S_cingul.Mid.Ant_area	0.21	.127	.599
Enriched	lh_G.S_cingul.Mid.Ant_area	0.31	<b>.022</b>	.599
Enriched	rh_G.S_cingul.Mid.Post_area	0.27	.080	.599
Enriched	lh_G.S_cingul.Mid.Post_area	0.30	.112	.599
Enriched	rh_G_front_middle_area	0.21	.068	.599
Enriched	lh_G_front_middle_area	0.27	<b>.046</b>	.599
Enriched	rh_G_front_inf.Opercular_area	0.31	.064	.599
Enriched	lh_G_front_inf.Opercular_area	0.44	<b>.030</b>	.599

Enriched	rh_Lat_Fis.post_area	0.16	.162	.599
Enriched	lh_Lat_Fis.post_area	0.37	<b>.025</b>	.599

*Note.* Rh indicates the ROI is retrieved from the right hemisphere, while lh indicates the ROI is retrieved from the left hemisphere. Additionally, area indicates that the ROI corresponds to the surface area of the corresponding region. A positive ratio indicates larger variance for males. Significant values have been highlighted.

**Figure 2**

*All average variance ratios displayed by group (enriched and depleted) with indication of significance based on uncorrected P-values*



*Note.* ROIs noted with \* showed P-values < .05.

### The relationship between sex differences in brain variability and ASD symptoms

For every ROI ( $N = 36$ ) we assessed the relationship between brain structure deviation (ROI Z-score) and ASD symptoms (symptom score) by comparing the model fit of a linear and a quadratic model using the BIC. This was done for two measures of ASD symptoms: the Social Responsiveness Scale (SRS) and the sensorimotor synchronization task (SMS), yielding a total of 72 analyses. We extracted the estimates and corresponding  $P$ -values of all predictors for both the linear and quadratic models, along with the BIC values and a difference score indicating when the quadratic model showed a better (see appendix E). We interpreted estimates and  $P$ -values for the models that showed the best fit. If a better fit for the quadratic model was found, we specifically examined predictor estimates of ROI Z-score to see if a U-shaped curve could be identified (indicated by a positive estimate the quadratic term ROI Z-score).

#### SRS

For all regions, the linear model best described the relationship between brain structure deviation and SRS score (table 4). For all linear models, no significant effects of ROI Z-score or the interaction of ROI Z-score with sex were found. The main effect of sex showed a significant effect for all models before and after applying multiple comparisons,

On average, males showed a higher score on the SRS ( $M = 29.33$ ,  $SD = 5.21$ ) when compared to females ( $M = 27.68$ ,  $SD = 4.04$ ), indicated by positive estimate values for sex. The main effect of zygosity also showed significant effects for all models both before and after multiple comparisons correction, where monozygotic twins showed a generally lower SRS score ( $M = 27.73$ ,  $SD = 4.16$ ) when compared to dizygotic twins ( $M = 29.32$ ,  $SD = 5.12$ ), indicated by negative estimates for zygosity.

**Table 4**

*Results of SRS analyses before multiple comparisons*

		Depleted regions ( $N = 10$ )	Enriched regions ( $N = 26$ )
Better quadratic fit	Significant ROI Z-score	0	<b>0</b>
	No significant ROI Z-score	0	0
Better linear fit	Significant ROI Z-score	0	0
	No significant ROI Z-score	10	26

*Note.* Results in the bolded category would confirm that brain structure deviations in regions with a high expression of the X chromosome relate to increased social communicative difficulties.

### SMS

For almost all regions, the linear model best described the relationship between brain structure deviation and SMS score, except for the left parieto-occipital sulcus (lh\_S\_parieto\_occipital\_area; table 5). Of the linear models, some regions showed a significant main effect of ROI Z-score before applying multiple comparisons correction: the left precentral gyrus (lh\_G\_precentral\_area,  $\beta = -0.04$ ,  $t = -3.28$ ,  $p = .001$ ), the right posterior lateral sulcus (rh\_Lat\_Fis.post\_area,  $\beta = -0.04$ ,  $t = -2.61$ ,  $p = .010$ ), the left superior precentral sulcus (lh\_S\_precentral\_sup.part\_area,  $\beta = -0.03$ ,  $t = -2.48$ ,  $p = .014$ ), the right subcentral gyrus and anterior, the posterior subcentral sulci (rh\_G.S\_subcentral\_area,  $\beta = -0.03$ ,  $t = -2.32$ ,  $p = .021$ ), and the right lingual gyrus (rh\_G\_oc.temp\_med.Lingual\_area,  $\beta = -0.03$ ,  $t = -2.09$ ,  $p = .038$ ). All betas show negative values, indicating that for a more positive ROI Z-score and thus a larger brain structure within these regions, the SMS score decreases. After applying multiple comparisons correction, only the main effect of ROI Z-score on SMS score for the left precentral gyrus survived (figure 3).

**Met opmerkingen [RLv(5)]:** Is it useful to calculate effect sizes (cohen's d) for the group differences?

**Table 5**

*Results of SMS analyses before multiple comparison*

		Depleted regions (N = 10)	Enriched regions (N = 26)
Better quadratic fit	Significant ROI Z-score	1	<b>0</b>
	No significant ROI Z-score	0	0
Better linear fit	Significant ROI Z-score	1	4
	No significant ROI Z-score	8	22

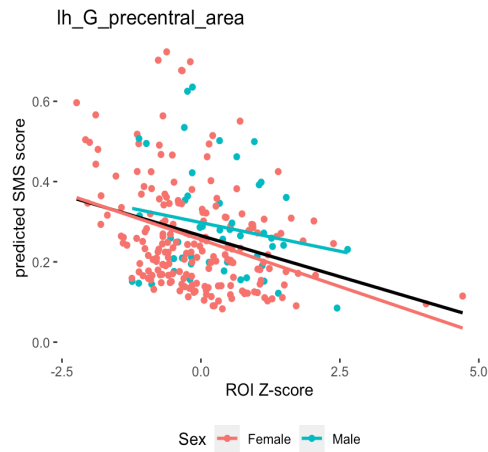
*Note.* Results in the bolded category would confirm that brain structure deviations in regions with a high expression of the X chromosome relate to increased sensorimotor difficulties.

The main effects of sex and zygosity did not show significant effects for any of the linear models. Before applying multiple comparisons correction, the interaction of sex and ROI Z-score showed a significant effect for the left posterior lateral sulcus (lh\_Lat\_Fis.post\_area,  $\beta = -0.06$ ,  $t = -2.13$ ,  $p = .034$ , figure 4a) and the right precentral gyrus (rh\_G\_precentral\_area,  $\beta = 0.06$ ,  $t = 2.02$ ,  $p = .045$ , figure 4b). The former region shows a negative relationship between ROI Z-score and SMS score for males, and a positive one for females, while the latter region shows a positive relationship between ROI Z-score and SMS score for males, and a negative one for females. However, these effects did not survive multiple comparisons correction

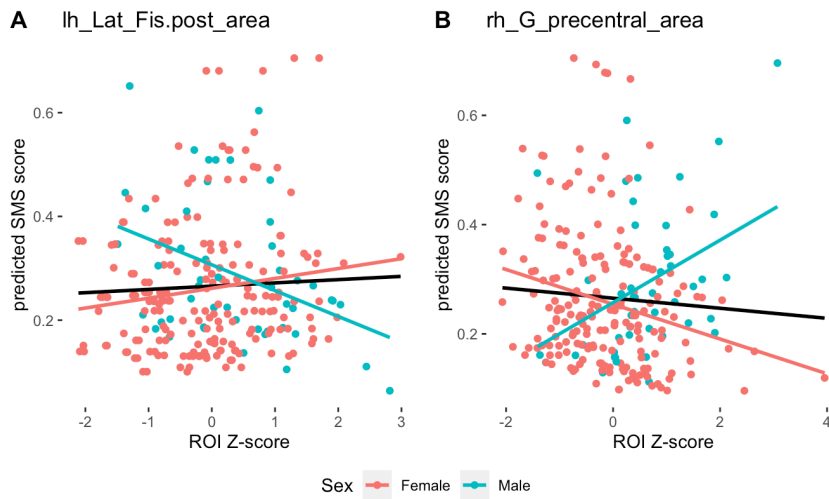


**Figure 3**

*The relationship between ROI Z-score and predicted SMS score for the left precentral gyrus*

**Figure 4**

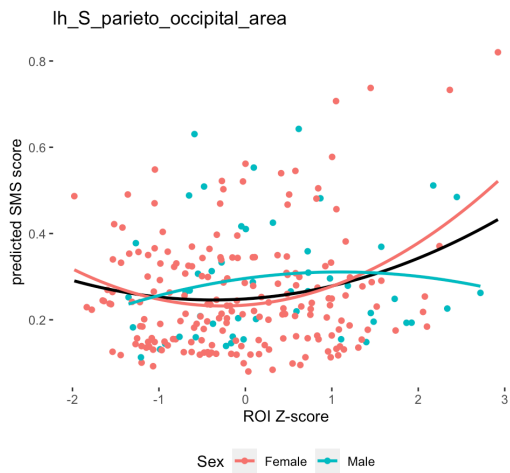
*The interaction effect of ROI Z-score and sex for predicted SMS score for the left posterior lateral sulcus and the right precentral gyrus*



The left parieto-occipital sulcus was the only region for which the quadratic model showed a better fit. Before we applied multiple comparisons correction, the main effect of ROI Z-score on SMS score (beta = 0.55,  $t = 3.11$ ,  $p = .002$ , figure 5) and the interaction of sex and ROI Z-score on SMS score (beta = -1.22,  $t = -3.10$ ,  $p = .002$ , figure 5) showed significant effects. This indicates that the relationship between ROI Z-score and SMS score differs between the sexes, showing a positive relationship for females and a negative relationship for males. These effects did not survive multiple comparisons correction.

**Figure 5**

*The relationship between ROI Z-score and SMS score and the interaction with sex for the left parieto-occipital sulcus*



### Discussion

The current study examined to what extent X chromosome effects on the brain are related to increased brain structure variance in males compared to females, and whether a larger deviance of brain structure relates to increased reporting of ASD symptoms (see figure 1 for a visualization). Variance ratios did not show a significant difference for X chromosome enriched regions compared to X chromosome depleted regions. Our results thus do not support the hypothesis that the X chromosome imbalance relates to larger brain structure variance in males compared to females. Additionally, no brain regions showed a U-shaped quadratic relationship. Our findings thus do not support the hypothesis that increased deviance of brain structure is related to an increase in ASD symptoms.

#### **The direct effect of the X chromosome imbalance and sex differences in brain structure variability**

We first aimed to replicate increased brain structure variance in males compared females, as was reported in previous studies (Ritchie et al., 2018; Wierenga et al., 2018, 2019, 2020). Whereas Wierenga et al., (2018) and Ritchie et al., (2018) only assessed overall brain volume of the cerebral cortex, Wierenga et al., (2019, 2020) provided an analysis of all cortical regions of SA. Wierenga et al. (2019) found a few SA regions to show significantly greater variance for males, while Wierenga et al., (2020) found all SA regions to show significantly greater variance for males. Although we did not test overall brain structure, we were unable to replicate the significantly greater variance for males compared to females in any region of SA.

To assess whether X chromosome effects relate to larger brain structure variance in males compared to females, we needed variance ratios to be positive for enriched regions (indicating larger male variance), and significantly more positive when compared to variance ratios in depleted regions. Contrary to our hypothesis, 3/26 enriched regions showed larger variance for females compared to males, indicated by negative variance ratios. None of the variance ratios for depleted regions showed negative values. We can however not draw any conclusions from these findings as none of the regions showed significant variance ratios after applying multiple comparisons correction. Most importantly, variance ratios were not significantly larger for X chromosome enriched regions. X chromosome effects are thus not related to increased brain structure variance in males compared to females. We could therefore not confirm a direct effect of the X chromosome imbalance on brain structure variance differences between the sexes. Our findings might suggest that the mere effect of the X chromosome imbalance is not sufficient to explain sex differences in brain structure variance. The current study focused on the effect of the X chromosome imbalance as the X chromosome holds more than 800 protein coding genes (excluding nonfunctional genes) while the Y chromosome holds only 48 known protein coding genes (Cabrera Zapata et al., 2022). Although most genes on the Y chromosome are involved in the male sexual differentiation, the Y chromosome also holds several homolog genes,

which are genes that show similarity to X linked genes as the X and Y chromosomes originate from the same autosomal genes (Cabrera Zapata et al., 2022). Recent animal studies have suggested some of these homolog genes might be involved in specific male brain development (Cabrera Zapata et al., 2022). Raznahan et al. (2016) has examined the potential effects of the Y chromosome on global and local SA regions within the brain. Increasing the number of Y chromosomes caused an overall increase in global SA. Additionally, increasing the number of Y chromosomes caused larger SA in medial posterior regions, including regions of the visual system that were assumed to be X chromosome depleted in the current study (Raznahan et al., 2016). If Y chromosome expression can cause larger SA in these X chromosome depleted regions, males can still show larger brain structure variance in these regions when compared to females (as found by Wierenga et al., 2020). The fact that variance ratios for depleted regions were not significantly lower than variance ratios for enriched regions could therefore depend on Y chromosome effects. Only taking the X chromosome effects into account might not be sufficient to explain sex differences in brain structure variance.

Besides a potential direct effect of the Y chromosome (and therefore its imbalance) on brain structure, the Y chromosome also initiates downstream mechanisms that determine the male sexual development, and thereby influence sex differences (Ferri, Abel & Brodtkin, 2018). Although a direct effect of the X chromosome imbalance on brain structure variance could still exist, downstream mechanisms of the sexual differentiation could additionally influence brain structure variance. An influence of hormones later in life seems unlikely as sex differences in variance have been observed in children aged 3 and were shown to remain stable across development (Wierenga et al., 2018, 2020). However, early mechanisms of the sexual differentiation such as prenatal hormones have often been implicated in brain structure development, where testosterone is considered an important contributor (Baron-Cohen, 2011; Ostatníková et al., 2021). Although brain structure and specifically SA is mostly determined by genetics, some brain regions show higher susceptibility to environmental influences (Hegarty et al., 2020; Peper et al., 2007; van der Meulen, 2020). Downstream mechanisms of the sexual differentiation such as prenatal hormones might therefore differentially affect brain structure regions in males and females, providing a way for downstream effects of the sexual differentiation and direct effects of the chromosomal imbalance to interact. This susceptibility to environmental influences might even be different for the sexes through sex differences in epigenetic mechanisms caused by the mechanisms of X inactivation, although this idea is still in its infancy and has not been studied much yet (Arnold, 2022; Cabrera Zapata et al., 2022).

Although our findings do not provide any conclusive results, they suggest that the relationship between the direct effect of the X chromosome imbalance and brain structure variance is not as straightforward as we hypothesized. The lack of a direct association between X chromosome effects and sex differences in brain structure variance could imply that sex differences in brain structure variance are formed by more complicated interactions of both the X and Y chromosome imbalance, and the downstream effects initiated for the sexual differentiation.

### **Sex differences in brain structure variability and sex differences in ASD**

To assess whether a larger brain structure deviation relates to increased ASD risk, we examined the relationship between brain structure deviation (as ROI Z-score) and ASD symptoms (as social communicative skills (SRS) and sensorimotor synchronization (SMS)). We examined this relationship in the same set of regions that we assessed the variance ratios in so potential relationships could be linked to X chromosome effects. To confirm both our hypotheses, an association between brain structure deviation and ASD symptoms was expected at least within enriched regions. However, no regions showed that larger brain structure deviance was associated with increased ASD symptoms, both for the social communicative difficulties as for the sensorimotor synchronization difficulties. Models including social communicative difficulties did not show U-shaped relationships (indicated by both a preference for the quadratic model and a positive estimate for ROI Z-score). Even within the linear models, brain structure deviance did not significantly explain social communicative scores. This suggests that no relationship exists between brain structure size within these regions and social communicative difficulties. For sensorimotor synchronization, the model including the left parieto-occipital sulcus showed a preference for the quadratic model. However, this region has shown a depletion of X chromosome effects and did not show a significant U-shaped relationship (Mallard et al., 2021). So, when the quadratic model provided a better fit, brain structure deviance also did not explain sensorimotor synchronization. Within the linear models, there was only one significant effect: brain structure deviance showed a negative association with sensorimotor synchronization within the left precentral gyrus. Individuals with a larger surface area within this region thus show better sensorimotor performance. This finding could be explained by the function of the precentral gyrus as it encompasses the primary motor cortex, a region that holds pyramidal neurons that extend down into the spinal cord to control various contralateral body parts including the fingers (see the cortical homunculus; Penfield & Rasmussen, 1950; Roux et al., 2020). As the SMS task requires finger tapping, better motor control due to an increased number of neurons could relate to better performance on the SMS, thereby explaining how larger SA in this region could relate to better sensorimotor performance. Although no studies have examined the relationship between primary motor cortex structure and sensorimotor synchronization in children, studies using functional brain activity imply that there is a relationship between the primary motor cortex' activity and sensorimotor synchronization (Criscuolo et al., 2022; Witt et al., 2008). The fact that 84.7% of participants in the SMS sample were right-handed (table 1) additionally explains why this effect would only be observed for the left precentral gyrus, as the region controls the contralateral body parts including the right hand. This finding indicates that increased lateralization to the right precentral gyrus, causing decreased lateralization to the left precentral gyrus, might relate to lower sensorimotor synchronization abilities. This aligns with the findings that ASD individuals show decreased sensorimotor synchronization (McNaughton & Redcay, 2020) and brain structure asymmetry in favor of the right

hemisphere (Carper et al., 2015). Further research is needed to replicate this finding and examine its potential relationship to ASD.

Overall, our findings thus do not support a relationship between larger brain structure deviation and ASD symptoms. It could be that a sex difference in brain variance does not affect a sex difference in ASD risk. Another explanation could lie in the highly individualized brain structure deviations found for individuals with ASD (Zabihi et al., 2019). As indicated by the broad autism phenotype, ASD has a high variability in the expression of symptoms (Sucksmith & Hoekstra, 2011). This means that every individual can show their own profile of certain symptoms. If brain structure deviations are functionally related to ASD symptoms, we would additionally expect each individual to show their own specific profile of brain deviations (Dawson et al., 2002; Zabihi et al., 2019). It might therefore be hard to assess a one-on-one relationship between large brain structure deviations in specific regions and ASD symptoms. Brain structure deviations related to ASD symptoms might instead be reflected in a network of regions associated with their individualized expression of ASD symptoms. This could explain why no association was found for specific regional brain structure deviations and ASD symptoms in the set of regions we analyzed, as these do not necessarily comprehend regions associated with the social communicative difficulties that are often observed for individuals with ASD or the broad autism phenotype (Zabihi et al., 2019; Losh et al., 2009). A more individualized pattern of brain structure deviations also leaves room for potential environmental influences to affect an individuals' brain structure, as environmental influences can differ per individual and could therefore contribute to individualized patterns of brain structure deviations. Although overall sex differences in brain structure variance might be essentially determined by genetics, environmental factors could shape an individuals' brain structure over time (Gu & Kanai, 2014). As we included structural brain data of individuals aged 7-9, potential environmental effects could have influenced individuals' brain structure. This could make it more complicated to assess a direct relationship between brain structure deviations in specific regions and ASD symptoms.

Although both our hypotheses could thus not be confirmed, the proposed relationships might not be as straightforward as we hypothesized. Brain structure variance can be determined by an interplay of multiple factors, and the mere effect of the X chromosome imbalance might not be sufficient to explain sex differences in brain structure variance. Additionally, profiles of brain structure variance associated with ASD could instead be reflected in patterns over multiple regions that relate to the individuals' symptoms, thereby showing no overall association between specific regions and ASD symptoms.

### **Strengths and limitations**

Overall, the current study provides an assessment of the relationships between X chromosome expression and brain structure variance, and brain structure deviation and ASD symptoms. The study holds several limitations. First, we did not consider other potential influences such as the downstream

effects of the sexual differentiation environmental effects. Second, our analyses were chosen as the data included dependency induced by twins. These analyses are however not fit to calculate a power analysis. We can therefore not be sure whether we reached sufficient power to detect an effect. Although the study by Wierenga et al. (2019) used a similar sample size to assess sex differences in brain structure variance, they found less regions to show a significant variance ratio when compared to the study by Wierenga et al. (2020), where a considerably larger sample size was included (299 participants and 16,683 participants respectively). Our relatively smaller sample size compared to the study by Wierenga et al. (2020) could therefore have influenced our power to detect an effect. Although we were additionally unable to replicate the effects found in the study by Wierenga et al. (2019), this might be attributable to our stricter correction for multiple comparisons, as this was not used in previous studies. Third, we analyzed regions according to the Destrieux atlas due to availability of this data (Destrieux et al., 2010). The Destrieux atlas contains larger regions when compared to the HCP atlas used by Mallard et al. (2021) to identify X chromosome enriched and depleted regions (Glasser et al., 2016). Therefore, the regions according to the Destrieux atlas might include additional SA that do not show X chromosome effects. Mallard et al., (2021) found X chromosome enrichment to operate very locally, as the overall total brain showed a depletion of X chromosome expression for SA. Potential effects of the X chromosome might be cancelled out due to the inclusion of additional SA that did not show enrichment, thereby affecting our power to find an effect within the regions we studied. Lastly, although the broad autism phenotype extends to the general population, differences between typically developing individuals and ASD individuals in brain structure susceptibility to genetic (and environmental influences) have been found (Hegarty et al., 2020). This suggests that limiting our study to typically developing individuals might decrease representation of the ASD population, and therefore decrease generalizability of these findings.

## Conclusion

No relationships between the X chromosome imbalance, sex differences in brain structure variance, and sex differences in ASD risk were found. Our hypotheses could therefore not be confirmed. This implies that the hypothesized relationships are absent, or that our study limitations prevented us from finding an effect. A challenge lies in the complexity of assessing a single genetic effect while not controlling for all other factors that could be involved. Both brain structure variance and ASD risk are likely influenced by an interplay of multiple mechanisms including multiple genetic and environmental effects. Limited research on potential contributions of other genetic mechanisms like the Y chromosome or difficulties in studying prenatal hormones make it even more complex to take such potential contributors into account. An unexplored pathway for future research lies in determining how the interplay of these and other involved mechanisms can contribute to sex differences in brain structure variance and sex differences in ASD risk, and to what extent the X chromosome imbalance might be involved.

Met opmerkingen [WL(6): Kan je ook nog redenen bedenken voor de SRS resultaten? E.g. we hebben bijvoorbeeld een normale populatie wat zowel een pre als een con is

### References

- Achterberg, M., & van der Meulen, M. (2019a). Genetic and environmental influences on MRI scan quantity and quality. *Developmental Cognitive Neuroscience*, 38, 100667. <https://doi.org/10.1016/j.dcn.2019.100667>
- Achterberg, M., & van der Meulen, M. (2019b). Genetic and environmental influences on MRI scan quantity and quality. *Developmental Cognitive Neuroscience*, 38, 100667. <https://doi.org/10.1016/j.dcn.2019.100667>
- Arnold, A. P. (2012). The end of gonad-centric sex determination in mammals. *Trends in Genetics: TIG*, 28(2), 55–61. <https://doi.org/10.1016/j.tig.2011.10.004>
- Arnold, A. P. (2022). X chromosome agents of sexual differentiation. *Nature Reviews Endocrinology*, 18(9), 574–583. <https://doi.org/10.1038/s41574-022-00697-0>
- Baron-Cohen, S., Lombardo, M. V., Auyeung, B., Ashwin, E., Chakrabarti, B., & Knickmeyer, R. (2011). Why are autism spectrum conditions more prevalent in males? *PLoS Biology*, 9(6), e1001081. <https://doi.org/10.1371/journal.pbio.1001081>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brezis, R.-S., Noy, L., Alony, T., Gotlieb, R., Cohen, R., Golland, Y., & Levit-Binnun, N. (2017). Patterns of Joint Improvisation in Adults with Autism Spectrum Disorder. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01790>
- Cabrera Zapata, L. E., Garcia-Segura, L. M., Cambiasso, M. J., & Arevalo, M. A. (2022). Genetics and Epigenetics of the X and Y Chromosomes in the Sexual Differentiation of the Brain. *International Journal of Molecular Sciences*, 23(20), 12288. <https://doi.org/10.3390/ijms232012288>
- Carper, R. A., Treiber, J. M., DeJesus, S. Y., & Müller, R.-A. (2016). Reduced Hemispheric Asymmetry of White Matter Microstructure in Autism Spectrum Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 55(12), 1073–1080. <https://doi.org/10.1016/j.jaac.2016.09.491>
- Carter, A., Davis, N., Klin, A., & Volkmar, F. (2013). Social Development in Autism. In *Handbook of Autism and Pervasive Developmental Disorders. Volume 1* (pp. 312–334). <https://doi.org/10.1002/9780470939345.ch11>
- Chita-Tegmark, M. (2016). Social attention in ASD: A review and meta-analysis of eye-tracking studies. *Research in Developmental Disabilities*, 48, 79–93. <https://doi.org/10.1016/j.ridd.2015.10.011>
- Coll, S.-M., Foster, N. E. V., Meilleur, A., Brambati, S. M., & Hyde, K. L. (2020a). Sensorimotor skills in autism spectrum disorder: A meta-analysis. *Research in Autism Spectrum Disorders*, 76, 101570. <https://doi.org/10.1016/j.rasd.2020.101570>



- Coll, S.-M., Foster, N., Meilleur, A., Brambati, S., & Hyde, K. (2020b). Sensorimotor skills in autism spectrum disorder: A meta-analysis. *Research in Autism Spectrum Disorders*, 76, 101570. <https://doi.org/10.1016/j.rasd.2020.101570>
- Constantino, J. N. (2013). Social Responsiveness Scale. In F. R. Volkmar (Ed.), *Encyclopedia of Autism Spectrum Disorders* (pp. 2919–2929). Springer. [https://doi.org/10.1007/978-1-4419-1698-3\\_296](https://doi.org/10.1007/978-1-4419-1698-3_296)
- Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., Metzger, L. M., Shoushtari, C. S., Splinter, R., & Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders*, 33(4), 427–433. <https://doi.org/10.1023/a:1025014929212>
- Crisuolo, A., Pando-Naude, V., Bonetti, L., Vuust, P., & Brattico, E. (2022). An ALE meta-analytic review of musical expertise. *Scientific Reports*, 12(1), Article 1. <https://doi.org/10.1038/s41598-022-14959-4>
- Crompton, C. J., Sharp, M., Axbey, H., Fletcher-Watson, S., Flynn, E. G., & Ropar, D. (2020). Neurotype-Matching, but Not Being Autistic, Influences Self and Observer Ratings of Interpersonal Rapport. *Frontiers in Psychology*, 11. <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.586171>
- Crone, E. A., M. Achterberg, M., Dobbelaar, S., Euser, S., Bulk, B. G. van den, Meulen, M. van der, Drunen, L. van, Wierenga, L. M., Bakermans-Kranenburg, M. J., & IJzendoorn, M. H. van. (2020). Neural and behavioral signatures of social evaluation and adaptation in childhood and adolescence: The Leiden consortium on individual development (L-CID). *Developmental Cognitive Neuroscience*, 45, 100805. <https://doi.org/10.1016/j.dcn.2020.100805>
- Dawson, G., Webb, S. J., Schellenberg, G., Dager, S., Friedman, S., Aylward, E., & Richards, T. (2002). Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology*, 14, 581–611. <https://doi.org/10.1017/S0954579402003103>
- Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, 53(1), 1–15. <https://doi.org/10.1016/j.neuroimage.2010.06.010>
- Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed* (pp. xlv, 947). (2013). American Psychiatric Publishing, Inc. <https://doi.org/10.1176/appi.books.9780890425596>
- Eigsti, I.-M. (2013). A Review of Embodiment in Autism Spectrum Disorders. *Frontiers in Psychology*, 4. <https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00224>
- Euser, S., Bakermans-Kranenburg, M. J., van den Bulk, B. G., Linting, M., Damsteegt, R. C., Vrijhof, C. I., van Wijk, I. C., Crone, E. A., & van IJzendoorn, M. H. (2016). Efficacy of the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline in Twin Families (VIPP-Twins): Study protocol for a randomized controlled trial. *BMC Psychology*, 4(1), 33. <https://doi.org/10.1186/s40359-016-0139-y>
- Ferri, S. L., Abel, T., & Brodtkin, E. S. (2018). Sex Differences in Autism Spectrum Disorder: A Review. *Current Psychiatry Reports*, 20(2), 9. <https://doi.org/10.1007/s11920-018-0874-2>

- Fitzpatrick, P., Romero, V., Amaral, J. L., Duncan, A., Barnard, H., Richardson, M. J., & Schmidt, R. C. (2017). Evaluating the importance of social motor synchronization and motor skill for understanding autism. *Autism Research, 10*(10), 1687–1699. <https://doi.org/10.1002/aur.1808>
- Fombonne, E. (2005). The Changing Epidemiology of Autism. *Journal of Applied Research in Intellectual Disabilities, 18*(4), 281–294. <https://doi.org/10.1111/j.1468-3148.2005.00266.x>
- Gu, J., & Kanai, R. (2014). What contributes to individual differences in brain structure? *Frontiers in Human Neuroscience, 8*. <https://doi.org/10.3389/fnhum.2014.00262>
- Hegarty, J. P., Pegoraro, L. F. L., Lazzeroni, L. C., Raman, M. M., Hallmayer, J. F., Monterrey, J. C., Cleveland, S. C., Wolke, O. N., Phillips, J. M., Reiss, A. L., & Hardan, A. Y. (2020). Genetic and environmental influences on structural brain measures in twins with autism spectrum disorder. *Molecular Psychiatry, 25*(10), 2556–2566. <https://doi.org/10.1038/s41380-018-0330-z>
- Herting, M. M., Gautam, P., Spielberg, J. M., Kan, E., Dahl, R. E., & Sowell, E. R. (2014). The role of testosterone and estradiol in brain volume changes across adolescence: A longitudinal structural MRI study. *Human Brain Mapping, 35*(11), 5633–5645. <https://doi.org/10.1002/hbm.22575>
- HOCHBERG, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75*(4), 800–802. <https://doi.org/10.1093/biomet/75.4.800>
- Hyman, S. L., Levy, S. E., Myers, S. M., & COUNCIL ON CHILDREN WITH DISABILITIES, SECTION ON DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS. (2020). Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. *Pediatrics, 145*(1), e20193447. <https://doi.org/10.1542/peds.2019-3447>
- Hyseni, F., Blanken, L. M. E., Muetzel, R., Verhulst, F. C., Tiemeier, H., & White, T. (2019). Autistic traits and neuropsychological performance in 6- to-10-year-old children: A population-based study. *Child Neuropsychology, 25*(3), 352–369. <https://doi.org/10.1080/09297049.2018.1465543>
- Kaur, M., Srinivasan, S., & Bhat, A. (2017). Comparing motor performance, praxis, coordination, and interpersonal synchrony between children with and without Autism Spectrum Disorder (ASD). *Research in Developmental Disabilities, 72*, 79–95. <https://doi.org/10.1016/j.ridd.2017.10.025>
- Klapwijk, E. T., van de Kamp, F., van der Meulen, M., Peters, S., & Wierenga, L. M. (2019). Qoala-T: A supervised-learning tool for quality control of FreeSurfer segmented MRI data. *NeuroImage, 189*, 116–129. <https://doi.org/10.1016/j.neuroimage.2019.01.014>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*, 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Li, J., Wang, L., Guo, H., Shi, L., Zhang, K., Tang, M., Hu, S., Dong, S., Liu, Y., Wang, T., Yu, P., He, X., Hu, Z., Zhao, J., Liu, C., Sun, Z. S., & Xia, K. (2017). Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. *Molecular Psychiatry, 22*(9), 1282–1290. <https://doi.org/10.1038/mp.2017.140>

- Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). Neuropsychological profile of autism and the broad autism phenotype. *Archives of General Psychiatry*, *66*(5), 518–526. <https://doi.org/10.1001/archgenpsychiatry.2009.34>
- Mallard, T. T., Liu, S., Seidlitz, J., Ma, Z., Moraczewski, D., Thomas, A., & Raznahan, A. (2021). X-chromosome influences on neuroanatomical variation in humans. *Nature Neuroscience*, *24*(9), Article 9. <https://doi.org/10.1038/s41593-021-00890-w>
- Marco, E. J., & Skuse, D. H. (2006). Autism-lessons from the X chromosome. *Social Cognitive and Affective Neuroscience*, *1*(3), 183–193. <https://doi.org/10.1093/scan/nsi028>
- McNaughton, K. A., & Redcay, E. (2020). Interpersonal Synchrony in Autism. *Current Psychiatry Reports*, *22*(3), 12. <https://doi.org/10.1007/s11920-020-1135-8>
- Meulen, M. van der, Wierenga, L. M., Achterberg, M., Drenth, N., IJzendoorn, M. H. van, & Crone, E. a. M. (2020a). Genetic and environmental influences on structure of the social brain in childhood. *Developmental Cognitive Neuroscience*, *44*, 100782. <https://doi.org/10.1016/j.dcn.2020.100782>
- Meulen, M. van der, Wierenga, L. M., Achterberg, M., Drenth, N., IJzendoorn, M. H. van, & Crone, E. a. M. (2020b). Genetic and environmental influences on structure of the social brain in childhood. *Developmental Cognitive Neuroscience*, *44*, 100782. <https://doi.org/10.1016/j.dcn.2020.100782>
- Morrison, K. E., DeBrabander, K. M., Jones, D. R., Ackerman, R. A., & Sasson, N. J. (2020). Social Cognition, Social Skill, and Social Motivation Minimally Predict Social Interaction Outcomes for Autistic and Non-Autistic Adults. *Frontiers in Psychology*, *11*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.591100>
- Nguyen, D. K., & Disteche, C. M. (2006). Dosage compensation of the active X chromosome in mammals. *Nature Genetics*, *38*(1), 47–53. <https://doi.org/10.1038/ng1705>
- Ostatníková, D., Lakatsova, S., Babková, J., Hodosy, J., & Celec, P. (2021). Testosterone and the brain: From cognition to autism. *Physiological Research*, *69*, S403–S419. <https://doi.org/10.33549/physiolres.934592>
- Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M. J., Grant, M. D., Franz, C. E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., & Kremen, W. S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cerebral Cortex (New York, N.Y.: 1991)*, *19*(11), 2728–2735. <https://doi.org/10.1093/cercor/bhp026>
- Pasterski, V., Gilligan, L., & Curtis, R. (2014). Traits of autism spectrum disorders in adults with gender dysphoria. *Archives of Sexual Behavior*, *43*(2), 387–393. <https://doi.org/10.1007/s10508-013-0154-5>
- Peper, C. E., van der Wal, S. J., & Begeer, S. (2016). Autism in Action: Reduced Bodily Connectedness during Social Interactions? *Frontiers in Psychology*, *7*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01862>

- Peper, J. S., Brouwer, R. M., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (2007). Genetic influences on human brain structure: A review of brain imaging studies in twins. *Human Brain Mapping*, 28(6), 464–473. <https://doi.org/10.1002/hbm.20398>
- Prochazkova, E., & Kret, M. E. (2017). Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neuroscience & Biobehavioral Reviews*, 80, 99–114. <https://doi.org/10.1016/j.neubiorev.2017.05.013>
- Raznahan, A., & Distèche, C. M. (2021). X-chromosome regulation and sex differences in brain anatomy. *Neuroscience and Biobehavioral Reviews*, 120, 28–47. <https://doi.org/10.1016/j.neubiorev.2020.10.024>
- Raznahan, A., Lee, N. R., Greenstein, D., Wallace, G. L., Blumenthal, J. D., Clasen, L. S., & Giedd, J. N. (2016). Globally Divergent but Locally Convergent X- and Y-Chromosome Influences on Cortical Development. *Cerebral Cortex*, 26(1), 70–79. <https://doi.org/10.1093/cercor/bhu174>
- Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage*, 61(4), 1402–1418. <https://doi.org/10.1016/j.neuroimage.2012.02.084>
- Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., Harris, M. A., Alderson, H. L., Hunter, S., Neilson, E., Liewald, D. C. M., Auyeung, B., Whalley, H. C., Lawrie, S. M., Gale, C. R., Bastin, M. E., McIntosh, A. M., & Deary, I. J. (2018). Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. *Cerebral Cortex (New York, N.Y.: 1991)*, 28(8), 2959–2975. <https://doi.org/10.1093/cercor/bhy109>
- Roeyers, H., Thys, M., Druart, C., De Schryver, M., & Schittekatte, M. (2011). *SRS, Screeningslijst voor autismespectrumstoornissen*. <http://hdl.handle.net/1854/LU-2067366>
- Roux, F.-E., Niare, M., Charni, S., Giussani, C., & Durand, J.-B. (2020). Functional architecture of the motor homunculus detected by electrostimulation. *The Journal of Physiology*, 598(23), 5487–5504. <https://doi.org/10.1113/JP280156>
- Ruigrok, A. N. V., 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 and Biobehavioral Reviews*, 39(100), 34–50. <https://doi.org/10.1016/j.neubiorev.2013.12.004>
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Schieber, M. H. (2020). Modern coordinates for the motor homunculus. *The Journal of Physiology*, 598(23), 5305–5306. <https://doi.org/10.1113/JP280712>
- Schultz, B. G., & van Vugt, F. T. (2016). Tap Arduino: An Arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments. *Behavior Research Methods*, 48, 1591–1607. <https://doi.org/10.3758/s13428-015-0671-3>

- Schulz, K. M., & Sisk, C. L. (2016). The organizing actions of adolescent gonadal steroid hormones on brain and behavioral development. *Neuroscience and Biobehavioral Reviews*, *70*, 148–158. <https://doi.org/10.1016/j.neubiorev.2016.07.036>
- Skuse, D. (2000). Imprinting, the X-Chromosome, and the Male Brain: Explaining Sex Differences in the Liability to Autism. *Pediatric Research*, *47*, 9–16. <https://doi.org/10.1203/00006450-200001000-00006>
- Sucksmith, E., Roth, I., & Hoekstra, R. A. (2011). Autistic traits below the clinical threshold: Re-examining the broader autism phenotype in the 21st century. *Neuropsychology Review*, *21*(4), 360–389. <https://doi.org/10.1007/s11065-011-9183-9>
- Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijdsdijk, F. (2016). Heritability of autism spectrum disorders: A meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *57*(5), 585–595. <https://doi.org/10.1111/jcpp.12499>
- Trevisan, D. A., Hoskyn, M., & Birmingham, E. (2018). Facial Expression Production in Autism: A Meta-Analysis. *Autism Research: Official Journal of the International Society for Autism Research*, *11*(12), 1586–1601. <https://doi.org/10.1002/aur.2037>
- van der Meulen, M., Wierenga, L. M., Achterberg, M., Drenth, N., van IJzendoorn, M. H., & Crone, E. A. (2020a). Genetic and environmental influences on structure of the social brain in childhood. *Developmental Cognitive Neuroscience*, *44*, 100782. <https://doi.org/10.1016/j.dcn.2020.100782>
- van der Meulen, M., Wierenga, L. M., Achterberg, M., Drenth, N., van IJzendoorn, M. H., & Crone, E. A. (2020b). Genetic and environmental influences on structure of the social brain in childhood. *Developmental Cognitive Neuroscience*, *44*, 100782. <https://doi.org/10.1016/j.dcn.2020.100782>
- van 't Hof, M., Tisseur, C., van Berckelaar-Onnes, I., van Nieuwenhuizen, A., Daniels, A. M., Deen, M., Hoek, H. W., & Ester, W. A. (2021). Age at autism spectrum disorder diagnosis: A systematic review and meta-analysis from 2012 to 2019. *Autism: The International Journal of Research and Practice*, *25*(4), 862–873. <https://doi.org/10.1177/1362361320971107>
- van Vugt, F. T. (2020). The TeensyTap Framework for Sensorimotor Synchronization Experiments. *Advances in Cognitive Psychology*, *16*(4), 302–308. <https://doi.org/10.5709/acp-0304-y>
- Werling, D. M., & Geschwind, D. H. (2013). Sex differences in autism spectrum disorders. *Current Opinion in Neurology*, *26*(2), 146–153. <https://doi.org/10.1097/WCO.0b013e32835ee548>
- Wierenga, L. M., Bos, M. G. N., Rossenberg, F. van, & Crone, E. A. (2019). Sex effects on development of brain structure and executive functions: Greater variance than mean effects. *Journal of Cognitive Neuroscience*, *31*(5), 730–753. [https://doi.org/10.1162/jocn\\_a\\_01375](https://doi.org/10.1162/jocn_a_01375)
- Wierenga, L. M., Bos, M. G. N., Schreuders, E., Kamp, F. van der, Peper, J. S., Tamnes, C. K., & Crone, E. A. (2018). Unraveling age, puberty and testosterone effects on subcortical brain development across adolescence. *Psychoneuroendocrinology*, *91*, 105–114. <https://doi.org/10.1016/j.psyneuen.2018.02.034>

- Wierenga, L. M., Doucet, G. E., Dima, D., Agartz, I., Aghajani, M., Akudjedu, T. N., Albajes-Eizagirre, A., D, A., Alpert, K. I., Andreassen, O. A., Anticevic, A., Asherson, P., Banaschewski, T., Bargallo, N., Baumeister, S., Baur-Streubel, R., Bertolino, A., Bonvino, A., Boomsma, D. I., ... Tamnes, C. K. (2020). Greater male than female variability in regional brain structure across the lifespan. *Human Brain Mapping*. <https://doi.org/10.1002/hbm.25204>
- Wierenga, L. M., Langen, M., Oranje, B., & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *NeuroImage*, *87*, 120–126. <https://doi.org/10.1016/j.neuroimage.2013.11.010>
- Wierenga, L. M., Sexton, J. A., Laake, P., Giedd, J. N., Tamnes, C. K., & Pediatric Imaging, Neurocognition, and Genetics Study. (2018). A Key Characteristic of Sex Differences in the Developing Brain: Greater Variability in Brain Structure of Boys than Girls. *Cerebral Cortex (New York, N.Y.: 1991)*, *28*(8), 2741–2751. <https://doi.org/10.1093/cercor/bhx154>
- Witt, S. T., Laird, A. R., & Meyerand, M. E. (2008). Functional neuroimaging correlates of finger-tapping task variations: An ALE meta-analysis. *NeuroImage*, *42*(1), 343–356. <https://doi.org/10.1016/j.neuroimage.2008.04.025>
- Zabihi, M., Oldehinkel, M., Wolfers, T., Frouin, V., Goyard, D., Loth, E., Charman, T., Tillmann, J., Banaschewski, T., Dumas, G., Holt, R., Baron-Cohen, S., Durston, S., Bölte, S., Murphy, D., Ecker, C., Buitelaar, J. K., Beckmann, C. F., & Marquand, A. F. (2019). Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using Normative Models. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *4*(6), 567–578. <https://doi.org/10.1016/j.bpsc.2018.11.013>

## Appendix A

**Table 6**

*The percentage of significant cortical areas per overarching region based on the HCP atlas by Glasser et al. (2016), as found by Mallard et al. (2021)*

<b>Region</b>	<b>Percentage of significant depleted cortical areas</b>	<b>Percentage of significant enriched cortical areas</b>
Primary visual cortex (1)	100%	0%
Early visual cortex (2)	50%	0%
Dorsal stream visual cortex (3)	8,3%	8,3%
Ventral stream visual cortex (6)	7%	0%
MT+ complex and neighbouring visual areas (5)	5,6%	5,6%
Somatosensory and motor cortex (6)	10%	30%
Paracentral lobular and mid cingulate cortex (7)	6,3%	50%
Premotor cortex (8)	8,3%	42,7%
Posterior opercular cortex (9)	0%	33,3%
Early auditory cortex (10)	0%	20%
Auditory association cortex (11)	6,3%	6,3%
Insular and frontal opercular cortex (12)	7,7%	11,5%
Medial temporal cortex (13)	21,4%	7,1%

Lateral temporal cortex (14)	0%	5,6%
Tempo-parietal occipital junction (15)	0%	30%
Superior parietal cortex (16)	0%	20%
Inferior parietal cortex (17)	0%	30%
Posterior cingulate cortex (18)	3,6%	14,3%
Anterior cingulate and medial prefrontal cortex (19)	20%	10%
Orbital and polar frontal cortex (20)	4,5%	0%
Inferior frontal cortex (21)	6,3%	12,5%
Dorsolateral prefrontal cortex (22)	0%	3,8%

*Note.* Percentages were calculated over both hemispheres, every region thus included both the left and right areas. Only regions that showed a minimum of 30% of significant regions were included in the current study.



## Appendix B

Table 7

Overview of the overarching regions from the HCP atlas and the corresponding ROIs from the Destrieux atlas that were included in the current study

Status	Region by Glasser et al. (2016)	Significant cortical areas (L)	Significant cortical areas (R)	Significant cortical areas	Corresponding ROIs in the Destrieux atlas
Depleted	Primary visual cortex (1)	100%	100%	100%	<p><b>Pole_Occipital (42)</b> The most lateral and posterior part of the Occipital lobe, representing parts of the primary and association visual cortices.</p> <p><b>G_cuneus, (11)</b> The medial part superior to the calcarine fissure and inferior to the parieto-occipital sulcus, therefore representing the medial aspects of the primary and association visual cortices.</p> <p><b>G_oc-temp_med-Lingual (22)</b> Lingual gyrus, inferior to the calcarine fissure and therefore representing the medial and inferior aspects of the visual cortices.</p>
Depleted	Early visual cortex (2)	25%	75%	50%	<p><b>S_calcarine (44)</b> The calcarine sulcus that separates the medial part of the occipital lobe into the cuneus and the lingual gyrus.</p> <p><b>S-parieto_occipital (65)</b> The parieto-occipital sulcus or fissure that superiorly limits the cuneus.</p>

---

				<p><b>S_precentral-inf-part (68)</b> Inferior part of the precentral sulcus, distinguishes premotor cortex from primary motor cortex.</p> <p><b>S_precentral-sup-part (69)</b> Superior part of the precentral sulcus, distinguishes supplementary motor area from primary motor cortex.</p> <p><b>G_precentral (29)</b> Precentral gyrus, corresponds to primary motor cortex.</p> <p><b>S_central (45)</b> Central sulcus, distinguishes the primary motor cortex precentral from primary somatosensory cortex post central.</p> <p><b>G_and_S_subcentral (4)</b> Subcentral sulcus and gyrus, corresponds to inferior part of primary motor and somatosensory cortex.</p> <p><b>G_postcentral (28)</b> The poscentral gyrus, corresponds to the somatosensory cortex, anteriorly bounded by the central sulcus and posteriorly by the post central sulcus.</p> <p><b>S_postcentral (67)</b> The post central sulcus limits the somatosensory cortex posteriorly.</p>
Enriched	Somatosensory and motor cortex (6)	40%	20%	30%

---

					<p><b>G_and_S_paracentral (3)</b> The paracentral lobule and sulcus, representing the paracentral lobule, that medially surrounds the central sulcus.</p>
Enriched	Paracentral lobular and mid cingulate cortex (7)	37,5%	62,5%	50%	<p><b>G_ad_S_cingul-Mid-Ant (7)</b> The middle anterior part of the cingulate gyrus and sulcus (aMCC). In combination with region 8 it forms the medial part of the cingulate cortex.</p> <p><b>G_ad_S_cingul-Mid-Post (8)</b> The middle posterior part of the cingulate gyrus and sulcus (pMCC). In combination with region 7 it forms the medial part of the cingulate cortex.</p>
Enriched	Premotor cortex (8)	16,7%	66,7%	42%	<p><b>G_front_middle (15)</b> The middle frontal gyrus, representing the middle lateral part of the frontal lobe encompassing both the middle premotor cortex and the dorsal lateral prefrontal cortex.</p> <p><b>G_front_inf-Opercular (12)</b> The opercular part of the inferior frontal gyrus, representing the lateral inferior part of the premotor cortex.</p> <p><i>*The middle frontal gyrus encompassing both the premotor region and the DLPFC will be analyzed, although caution should be taken in ascribing findings solely to the premotor cortex.</i></p>

---

Enriched	Posterior opercular cortex (9)	33,3%	33,3%	33,3%	<b>Lat_Fis-post (41)</b> Posterior segment of the lateral sulcus. represents the posterior part of the lateral sulcus, causing the region to find inside the brain forming the opercular cortex.
----------	--------------------------------	-------	-------	-------	---

---

*Note.* The overarching regions are indicated with numbers corresponding to numbering by Glasser et al. (2016). The same applies to regions from the Destrieux atlas indicated between brackets (Destrieux et al., 2010).

## Appendix C

Table 8

*P-values and adjusted P-values for variance ratios of every ROI for data containing all first individuals from each twin (child number = 1)*

Status	ROI	Variance Ratio	P-values	Adjusted P-values
Depleted	rh_Pole_occipital_area	0.03	.420	.587
Depleted	lh_Pole_occipital_area	-0.02	.537	.587
Depleted	rh_G_cuneus_area	0.17	.132	.587
Depleted	lh_G_cuneus_area	0.12	.205	.587
Depleted	rh_G_oc.temp.med.Lingual_area	0.26	<b>.041</b>	.587
Depleted	lh_G_oc.temp.med.Lingual_area	0.19	.090	.587
Depleted	rh_S_calcarine_area	0.25	.088	.587
Depleted	lh_S_calcarine_area	0.37	<b>.016</b>	.530
Depleted	rh_S_parieto_occipital_area	0.27	<b>.028</b>	.587
Depleted	lh_S_parieto_occipital_area	0.48	<b>&lt; .001</b>	<b>.004</b>
Enriched	rh_S_precentral.inf.part_area	0.07	.365	.587
Enriched	lh_S_precentral.inf.part_area	0.03	.444	.587
Enriched	rh_S_precentral.sup.part_area	0.43	<b>.019</b>	.587
Enriched	lh_S_precentral.sup.part_area	0.10	.266	.587
Enriched	rh_G_precentral_area	0.42	.116	.587
Enriched	lh_G_precentral_area	-0.08	.585	.587
Enriched	rh_S_central_area	0.03	.443	.587
Enriched	lh_S_central_area	0.19	.188	.587
Enriched	rh_G.S_subcentral_area	0.08	.338	.587
Enriched	lh_G.S_subcentral_area	0.24	.110	.587

Enriched	rh_G_postcentral_area	0.30	.179	.587
Enriched	lh_G_postcentral_area	0.30	<b>.036</b>	.587
Enriched	rh_S_postcentral_area	0.02	.440	.587
Enriched	lh_S_postcentral_area	0.05	.392	.587
Enriched	rh_G.S_paracentral_area	0.16	.189	.587
Enriched	lh_G.S_paracentral_area	0.14	.167	.587
Enriched	rh_G.S_cingul.Mid.Ant_area	0.12	.271	.587
Enriched	lh_G.S_cingul.Mid.Ant_area	0.27	.136	.587
Enriched	rh_G.S_cingul.Mid.Post_area	0.13	.227	.587
Enriched	lh_G.S_cingul.Mid.Post_area	-0.02	.538	.587
Enriched	rh_G_front_middle_area	0.08	.310	.587
Enriched	lh_G_front_middle_area	0.13	.250	.587
Enriched	rh_G_front_inf.Opercular_area	0.05	.398	.587
Enriched	lh_G_front_inf.Opercular_area	0.44	<b>.038</b>	.587
Enriched	rh_Lat_Fis.post_area	0.06	.318	.587
Enriched	lh_Lat_Fis.post_area	0.33	<b>.012</b>	.424

*Note.* Rh indicates the ROI is retrieved from the right hemisphere, while lh indicates the ROI is retrieved from the left hemisphere. Additionally, area indicates that the ROI corresponds to the surface area of the corresponding region. A positive ratio indicates larger variance for males. Significant values have been highlighted.

**Table 9**

*P-values and adjusted P-values for variance ratios of every ROI for data containing all first individuals from each twin (child number = 2)*

<b>Status</b>	<b>ROI</b>	<b>Variance Ratio</b>	<b>P-value</b>	<b>Adjusted P-value</b>
Depleted	rh_Pole_occipital_area	0.39	<b>.01</b>	.272
Depleted	lh_Pole_occipital_area	0.26	<b>.036</b>	.722
Depleted	rh_G_cuneus_area	-0.03	.557	.728
Depleted	lh_G_cuneus_area	0.07	.358	.728
Depleted	rh_G_oc.temp_med.Lingual_area	0.21	.095	.728
Depleted	lh_G_oc.temp_med.Lingual_area	0.26	<b>.048</b>	.728
Depleted	rh_S_calcarine_area	0.28	<b>.036</b>	.722
Depleted	lh_S_calcarine_area	0.37	<b>.005</b>	.149
Depleted	rh_S_parieto_occipital_area	0.08	.300	.728
Depleted	lh_S_parieto_occipital_area	0.31	<b>.020</b>	.453
Enriched	rh_S_precentral.inf.part_area	0.32	<b>.041</b>	.728
Enriched	lh_S_precentral.inf.part_area	-0.16	.708	.728
Enriched	rh_S_precentral.sup.part_area	0.32	<b>.014</b>	.353
Enriched	lh_S_precentral.sup.part_area	0.09	.323	.728
Enriched	rh_G_precentral_area	0.22	.136	.728
Enriched	lh_G_precentral_area	0.01	.494	.728
Enriched	rh_S_central_area	-0.13	.728	.728
Enriched	lh_S_central_area	0.06	.392	.728
Enriched	rh_G.S_subcentral_area	0.53	<b>.002</b>	.061
Enriched	lh_G.S_subcentral_area	0.41	<b>.006</b>	.180
Enriched	rh_G_postcentral_area	-0.03	.5595	.728

Enriched	lh_G_postcentral_area	0.13	.2497	.728
Enriched	rh_S_postcentral_area	0.14	.1953	.728
Enriched	lh_S_postcentral_area	0.58	<b>&lt; .001</b>	<b>.032</b>
Enriched	rh_G.S_paracentral_area	0.27	.245	.728
Enriched	lh_G.S_paracentral_area	0.21	.106	.728
Enriched	rh_G.S_cingul.Mid.Ant_area	0.29	.060	.728
Enriched	lh_G.S_cingul.Mid.Ant_area	0.34	<b>.012</b>	.312
Enriched	rh_G.S_cingul.Mid.Post_area	0.41	<b>.019</b>	.451
Enriched	lh_G.S_cingul.Mid.Post_area	0.63	<b>.002</b>	.066
Enriched	rh_G_front_middle_area	0.33	<b>.008</b>	.235
Enriched	lh_G_front_middle_area	0.41	<b>.005</b>	.149
Enriched	rh_G_front_inf.Opercular_area	0.57	<b>.002</b>	.060
Enriched	lh_G_front_inf.Opercular_area	0.44	<b>.024</b>	.537
Enriched	rh_Lat_Fis.post_area	0.26	<b>.012</b>	.313
Enriched	lh_Lat_Fis.post_area	0.41	.053	.728

*Note.* Rh indicates the ROI is retrieved from the right hemisphere, while lh indicates the ROI is retrieved from the left hemisphere. Additionally, area indicates that the ROI corresponds to the surface area of the corresponding region. A positive ratio indicates larger variance for males. Significant values have been highlighted.



## Appendix D

### Data preparation

#### Set working directory

```
setwd("~/Users/lottevanrijn/Library/Mobile Documents/com~apple~CloudDocs/Documents/thesis/Methods/data")
```

#### Load libraries

```
library(readxl)
library(plyr)
library(dplyr)
library(ggplot2)
library(randomForest)
library(tidyr)
library(reshape2)
library(lme4)
library(lmerTest)
library(gridExtra)
library(ggpubr)
```

#### Load data

```
lcid_ecc_MRIdata_ses1 <- read.csv2("lcid_ecc_MRIdata_child1ses1_voorLotte.csv")
lcid_ecc_behavioral <- read.csv2("lcid_ecc_voorLotte2.csv")
lcid_mcc_MRIdata_ses1 <- read.csv2("lcid_mcc_MRIdata_child1ses1_voorLotte.csv")
lcid_mcc_behavioral <- read.csv2("lcid_mcc_voorLotte2.csv")
```

#### Remove unnecessary variables

```
lcid_ecc_behavioral = subset(lcid_ecc_behavioral, select = -c(mcc_childnr, ecc_dropout, ecc_dropout_before, ecc_w05_participation, ecc_wppsi_tiq_child, ecc_handedness, ecc_w05_stop_total_correctGO, ecc_w05_stop_total_correctNOGO, ecc_w05_stop_total_median_corGO_rt, ecc_w05_stop_total_mean_corGO_rt, ecc_w05_stop_total_mean_ssd, ecc_w05_stop_total_ssr))
```

```
lcid_mcc_behavioral = subset(lcid_mcc_behavioral, select = -c(mcc_dropout, mcc_dropout_before, mcc_w01_participation, mcc_w03_participation, mcc_w05_participation, mcc_w01_wisc_tiq_child, mcc_w01_health_disability, mcc_w01_health_disability_type, mcc_w01_stop_total_mean_corGO_rt, mcc_w03_health_changes, mcc_w03_health_changes_remarks, mcc_w05_health_changes, mcc_w05_health_changes_remarks, MCC_W0, MCC_W1, MCC_W2, mcc_handedness, mcc_w01_stop_total_correctGO, mcc_w01_stop_total_correctNOGO, mcc_w01_stop_total_median_corGO_rt, mcc_w01_stop_total_mean_ssd, mcc_w01_stop_total_ssr, mcc_w05_seq_totalsum_attentionhyper_pp, mcc_w05_seq_attention_sum_pp, mcc_w05_seq_hyper_sum_pp, mcc_w05_seq_impulsivity_sum_pp, mcc_w05_seq_q_autism_totalsum_pp))
```

```
lcid_ecc_MRIdata_ses1 = subset(lcid_ecc_MRIdata_ses1, select = c(visit_id, rh_Pole_occipital_area, lh_Pole_occipital_area, rh_G_cuneus_area, lh_G_cuneus_area, rh_G_oc.temp_med.Lingual_area, lh_G_oc.temp_med.Lingual_area, rh_S_calcarine_area, lh_S_calcarine_area, rh_S_parieto_occipital_area, lh_S_parieto_occipital_area, rh_S_precentral.inf.part_area, lh_S_precentral.inf.part_area, rh_S_precentral.sup.part_area, lh_S_precentral.sup.part_area, rh_G_precentral_area, lh_G_precentral_area, rh_S_central_area, lh_S_central_area, rh_G.S_subcentral_area, lh_G.S_subcentral_area, rh_G_postcentral_area, lh_G_postcentral_area, rh_S_postcentral_area, lh_S_postcentral_area, rh_G.S_paracentral_area, lh_G.S_paracentral_area, rh_G.S_cingul.Mid.Ant_area, lh_G.S_cingul.Mid.Ant_area, rh_G.S_cingul.
```

```
Mid.Post_area, lh_G.S_cingul.Mid.Post_area, rh_G_front_middle_area, lh_G_front_middle_area, rh_G_front_inf.Opercular_area, lh_G_front_inf.Opercular_area, rh_Lat_Fis.post_area, lh_Lat_Fis.post_area))
```

```
lcid_mcc_MRIdata_ses1 = subset(lcid_mcc_MRIdata_ses1, select = c(visit_id, rh_Pole_occipital_area, lh_Pole_occipital_area, rh_G_cuneus_area, lh_G_cuneus_area, rh_G_oc.temp_med.Lingual_area, lh_G_oc.temp_med.Lingual_area, rh_S_calcarine_area, lh_S_calcarine_area, rh_S_parieto_occipital_area, lh_S_parieto_occipital_area, rh_S_precentral.inf.part_area, lh_S_precentral.inf.part_area, rh_S_precentral.sup.part_area, lh_S_precentral.sup.part_area, rh_G_precentral_area, lh_G_precentral_area, rh_S_central_area, lh_S_central_area, rh_G.S_subcentral_area, lh_G.S_subcentral_area, rh_G_postcentral_area, lh_G_postcentral_area, rh_S_postcentral_area, lh_S_postcentral_area, rh_G.S_paracentral_area, lh_G.S_paracentral_area, rh_G.S_cingul.Mid.Ant_area, lh_G.S_cingul.Mid.Ant_area, rh_G.S_cingul.Mid.Post_area, lh_G.S_cingul.Mid.Post_area, rh_G_front_middle_area, lh_G_front_middle_area, rh_G_front_inf.Opercular_area, lh_G_front_inf.Opercular_area, rh_Lat_Fis.post_area, lh_Lat_Fis.post_area))
```

## Creating a dataset for the analyses with the Social Responsiveness Scale (SRS)

### Check whether missings occur in the data for SRS

```
is.na(lcid_ecc_behavioral$ecc_w05_srs_sum)
is.na(lcid_mcc_behavioral$mcc_w03_srs_sum)
```

### Recode factor "NA" to missing

```
lcid_ecc_behavioral[lcid_ecc_behavioral == "NA"] <- NA
lcid_mcc_behavioral[lcid_mcc_behavioral == "NA"] <- NA
```

### Drop individuals with missing values for SRS

```
SRS_data_ecc <- drop_na(lcid_ecc_behavioral, ecc_w05_srs_sum)
SRS_data_mcc <- drop_na(lcid_mcc_behavioral, mcc_w03_srs_sum)
```

### Remove additional unnecessary variables for SRS data

```
SRS_data_mcc = subset(SRS_data_mcc, select = -c(MCC5_RV_TRIAL_Music_easy_mean, mcc_w01_age_child, mcc_w05_age_child))
```

### Creating similar visit\_id's in every relevant dataset for merging of the behavioral data and MRI data relating to the SRS analyses

```
for (row in 1:nrow(SRS_data_ecc)) {
  ID <- substr(SRS_data_ecc$visit_id, 5, 13)
}
for (row in 1:nrow(SRS_data_ecc)) {
  SRS_data_ecc$visit_id <- ID
}
for (row in 1:nrow(SRS_data_mcc)) {
  ID <- substr(SRS_data_mcc$sub_id, 5, 13)
}
for (row in 1:nrow(SRS_data_mcc)) {
  SRS_data_mcc$sub_id <- ID
}
SRS_data_mcc <- rename(SRS_data_mcc, "visit_id" = "sub_id")

for (row in 1:nrow(lcid_ecc_MRIdata_ses1)) {
  ID <- substr(lcid_ecc_MRIdata_ses1$visit_id, 5, 13)
```

```

}
for (row in 1:nrow(lcid_ecc_MRIdata_ses1)) {
  lcid_ecc_MRIdata_ses1$visit_id <- ID
}
for (row in 1:nrow(lcid_mcc_MRIdata_ses1)) {
  ID <- substr(lcid_mcc_MRIdata_ses1$visit_id, 5, 13)
}
for (row in 1:nrow(lcid_mcc_MRIdata_ses1)) {
  lcid_mcc_MRIdata_ses1$visit_id <- ID
}

```

### Giving variables for the cohorts similar names for merging of behavioral data

```

SRS_data_ecc <- rename(SRS_data_ecc, sex_child = ecc_sex_child)
SRS_data_ecc <- rename(SRS_data_ecc, primaryparent_twin = ecc_w01_primaryparent)
SRS_data_ecc <- rename(SRS_data_ecc, age_child = ecc_w05_age_child)
SRS_data_ecc <- rename(SRS_data_ecc, ethnicity_twin = ecc_ethnicity_twin)
SRS_data_ecc <- rename(SRS_data_ecc, ses_twin = ecc_w01_ses)
SRS_data_ecc <- rename(SRS_data_ecc, zygosity_twin = ecc_zygosity)
SRS_data_ecc <- rename(SRS_data_ecc, health_disorder_child = ecc_w01_health_disorder)
SRS_data_ecc <- rename(SRS_data_ecc, health_disorder_type_child = ecc_w01_health_disorder_type)
)
SRS_data_ecc <- rename(SRS_data_ecc, SRS_sum = ecc_w05_srs_sum)
SRS_data_mcc <- rename(SRS_data_mcc, sex_child = mcc_sex_child)
SRS_data_mcc <- rename(SRS_data_mcc, primaryparent_twin = mcc_primaryparent)
SRS_data_mcc <- rename(SRS_data_mcc, age_child = mcc_w03_age_child)
SRS_data_mcc <- rename(SRS_data_mcc, ethnicity_twin = mcc_ethnicity)
SRS_data_mcc <- rename(SRS_data_mcc, ses_twin = mcc_w01_ses)
SRS_data_mcc <- rename(SRS_data_mcc, zygosity_twin = mcc_zygosity)
SRS_data_mcc <- rename(SRS_data_mcc, health_disorder_child = mcc_w01_health_disorder)
SRS_data_mcc <- rename(SRS_data_mcc, health_disorder_type_child = mcc_w01_health_disorder_t
ype)
SRS_data_mcc <- rename(SRS_data_mcc, SRS_sum = mcc_w03_srs_sum)

```

### Then we merge the behavioral data available for the SRS in the ECC cohort and the MCC cohort

```
SRS_data <- rbind(SRS_data_ecc, SRS_data_mcc)
```

### Then we merge the behavioral data available for the SRS with corresponding MRI data

```

MRI_data <- rbind(lcid_ecc_MRIdata_ses1, lcid_mcc_MRIdata_ses1)
SRS_MRI_data <- merge(SRS_data, MRI_data, by.x = "visit_id", by.y = "visit_id")

```

### Creating a dataset for the analyses with the Sensorimotor synchronization task (SMS)

#### Check whether missings occur in the data for the SMS

```
is.na(lcid_mcc_behavioral$MCC5_RV_TRIAL_Music_easy_mean)
```

#### Drop individuals with missing values for SMS

```
SMS_data <- drop_na(lcid_mcc_behavioral, MCC5_RV_TRIAL_Music_easy_mean)
```

#### Remove additional unnecessary variables for SMS data

```
SMS_data = subset(SMS_data, select = -c(mcc_w03_srs_sum, mcc_w01_age_child, mcc_w03_age_c
hild))
```

### Creating similar Visit\_id's in every relevant dataset for merging of the behavioral data and MRI data relating to the SMS analyses

```
for (row in 1:nrow(SMS_data)) {
  ID <- substr(SMS_data$sub_id, 5, 13)
}
for (row in 1:nrow(SMS_data)) {
  SMS_data$sub_id <- ID
}
SMS_data <- rename(SMS_data, "visit_id" = "sub_id")
```

### Giving variables for the cohorts similar names for merging of behavioral data and MRI data

```
SMS_data <- rename(SMS_data, sex_child = mcc_sex_child)
SMS_data <- rename(SMS_data, primaryparent_twin = mcc_primaryparent)
SMS_data <- rename(SMS_data, age_child = mcc_w05_age_child)
SMS_data <- rename(SMS_data, ethnicity_twin = mcc_ethnicity)
SMS_data <- rename(SMS_data, ses_twin = mcc_w01_ses)
SMS_data <- rename(SMS_data, zygosity_twin = mcc_zygosity)
SMS_data <- rename(SMS_data, health_disorder_child = mcc_w01_health_disorder)
SMS_data <- rename(SMS_data, health_disorder_type_child = mcc_w01_health_disorder_type)
SMS_data <- rename(SMS_data, SMS_score = MCC5_RV_TRIAL_Music_easy_mean)
```

### Then we merge the behavioral data available for the SMS with corresponding MRI data

```
SMS_MRI_data <- merge(SMS_data, MRI_data, by.x = "visit_id", by.y = "visit_id")
SMS_MRI_data$SMS_score <- as.numeric(SMS_MRI_data$SMS_score)
```

### Recode SMS score so a higher score indicates more symptoms

```
SMS_MRI_data$SMS_score <- (1 - SMS_MRI_data$SMS_score)
save(SMS_MRI_data, file = "SMS_MRI_data.RData")
```

## Assessing variance ratios

Creating dataset 1 for the assessment of the variance ratios

### Creating a subject ID to subset data for one individual so we can remove dependency from the data

```
for (row in 1:nrow(MRI_data)) {
  subjnr <- substr(MRI_data$visit_id, 8, 9)
  MRI_data$SubjNumber <- subjnr
}
```

### Select either individuals 1 or 2

```
Ratio_MRI_data <- MRI_data[MRI_data$SubjNumber == "01", ]
```

### Creating a dataset with the demographic variables with ID (visit\_id), age (age\_child) and sex (sex\_child)

```
Ratio_demo_data_ecc <- subset(lcid_ecc_behavioral, select = c(visit_id, ecc_sex_child, ecc_w05_age_child))
Ratio_demo_data_ecc <- rename(Ratio_demo_data_ecc, sex_child = ecc_sex_child)
Ratio_demo_data_ecc <- rename(Ratio_demo_data_ecc, age_child = ecc_w05_age_child)
Ratio_demo_data_mcc <- subset(lcid_mcc_behavioral, select = c(sub_id, mcc_sex_child, mcc_w01_age_child))
Ratio_demo_data_mcc <- rename(Ratio_demo_data_mcc, visit_id = sub_id)
Ratio_demo_data_mcc <- rename(Ratio_demo_data_mcc, sex_child = mcc_sex_child)
Ratio_demo_data_mcc <- rename(Ratio_demo_data_mcc, age_child = mcc_w01_age_child)
```

### Merge data from the ECC and MCC

```
Ratio_demo_data <- rbind(Ratio_demo_data_ecc, Ratio_demo_data_mcc)
```

### Create a similar ID variable as in the MRI data for merging

```
for (row in 1:nrow(Ratio_demo_data)) {
  ID <- substr(Ratio_demo_data$visit_id, 5, 13)
}
for (row in 1:nrow(Ratio_demo_data)) {
  Ratio_demo_data$visit_id <- ID
}
```

### Merge demographic data and MRI data to create dataset with all available data for variance ratios (childr = 1)

```
Ratio_data <- merge(Ratio_demo_data, Ratio_MRI_data, by.x = "visit_id", by.y = "visit_id")
Ratio_data <- Ratio_data[complete.cases(Ratio_data[,c(4:39)]),]
```

### Give the correct variable type to all variables

```
Ratio_data$visit_id <- as.factor(Ratio_data$visit_id)
Ratio_data$sex_child <- as.factor(Ratio_data$sex_child)
Ratio_data$SubjNumber <- as.factor(Ratio_data$SubjNumber)
Ratio_data <- Ratio_data %>% mutate_if(is.character, as.numeric)
Ratio_data <- Ratio_data %>% mutate_if(is.integer, as.numeric)
```

### Assessing the variance ratios for dataset 1

### Create a dataframe with residuals for all ROIs while controlling for gender and age

```
vnames=names(Ratio_data[,c(4:39)])
resid_DF=NULL
for(vname in vnames) {
  df=data.frame(y=Ratio_data[[vname]],gender=Ratio_data$sex_child, age=Ratio_data$age_child)
  ff=randomForest(y~.,data=df,ntrees=500,proximity=TRUE,importance=TRUE,do.trace=100)
  residuals=ff$y -predict(ff)
  resid_DF=cbind(resid_DF,residuals)
}
resid_DF <- as.data.frame(resid_DF, row.names = vnames)
names(resid_DF)=vnames
```

### Make separate vectors of same length for gender and age

```
gender <- Ratio_data$sex_child
age <- Ratio_data$age_child
```

### Permute data for variance ratios

```
B=10000
perm=NULL
n=nrow(resid_DF)
for(b in 1:B) {
  gender_permute=gender[sample(n,n)]
  aa=aggregate(resid_DF,by=list(gr=gender_permute),var)
  ratio=log(aa[aa$gr=="Male",-1]/aa[aa$gr=="Female",-1])
  perm=rbind(perm,ratio)
}
```

### Calculate an overall variance ratio for each ROI

```
aa=aggregate(resid_DF,by=list(gr=gender),var)
ratio_obs=log(aa[aa$gr=="Male",-1]/aa[aa$gr=="Female",-1])
```

### create P-values for each ROI

```
p_value=NULL
for(k in 1:length(ratio_obs)) {
  p_value=c(p_value,mean(perm[,k]>=as.numeric(ratio_obs[k])))
}
```

### Apply multiple comparisons check using the hochberg procedure

```
Hochberg_p_value <- rbind(ratio_obs, p_value)
Hochberg_p_value <- Hochberg_p_value[-c(1),]
Hochberg_p_value <- melt(Hochberg_p_value)
Hochberg_p_value <- rename(Hochberg_p_value, Area = variable)
Hochberg_p_value = Hochberg_p_value[order(Hochberg_p_value$value),]
Hochberg_p_value$Hochberg <- p.adjust(Hochberg_p_value$value, method = "hochberg")
```

### Creating dataset 2 for the assessment of the variance ratios

#### Select remaining individuals that were not included in dataset 1 (childnr = 2)

```
Ratio_MRI_data_2 <- MRI_data[MRI_data$SubjNumber == "02", ]
```

#### Merge demographic data and MRI data to create dataset with all available data for variance ratios (childnr = 1)

```
Ratio_data_2 <- merge(Ratio_demo_data, Ratio_MRI_data_2, by.x = "visit_id", by.y = "visit_id")
Ratio_data_2 <- Ratio_data_2[complete.cases(Ratio_data_2[,c(4:39)]),]
```

#### Give the correct variable type to all variables

```
Ratio_data_2$visit_id <- as.factor(Ratio_data_2$visit_id)
Ratio_data_2$sex_child <- as.factor(Ratio_data_2$sex_child)
Ratio_data_2$SubjNumber <- as.factor(Ratio_data_2$SubjNumber)
Ratio_data_2 <- Ratio_data_2 %>% mutate_if(is.character, as.numeric)
Ratio_data_2 <- Ratio_data_2 %>% mutate_if(is.integer, as.numeric)
```

### Assessing the variance ratios for dataset 2

#### Create a dataframe with residuals for all ROIs while controlling for gender and age

```

vnames=names(Ratio_data_2[c(4:39)])
resid_DF=NULL
for(vname in vnames) {
  df=data.frame(y=Ratio_data_2[[vname]],gender=Ratio_data_2$sex_child,age=Ratio_data_2$age_child)
  ff=randomForest(y~.,data=df,ntrees=500,proximity=TRUE,importance=TRUE,do.trace=100)
  residuals=ff$y -predict(ff)
  resid_DF=cbind(resid_DF,residuals)
}
resid_DF <- as.data.frame(resid_DF, row.names = vnames)
names(resid_DF)=vnames

```

### Make separate vectors of same length for gender and age

```

gender <- Ratio_data_2$sex_child
age <- Ratio_data_2$age_child

```

### Permute data for variance ratios

```

B=10000
perm=NULL
n=nrow(resid_DF)
for(b in 1:B) {
  gender_permute=gender[sample(n,n)]
  aa=aggregate(resid_DF,by=list(gr=gender_permute),var)
  ratio=log(aa[aa$gr=="Male",-1]/aa[aa$gr=="Female",-1])
  perm=rbind(perm,ratio)
}

```

### Calculate an overall variance ratio for each ROI

```

aa=aggregate(resid_DF,by=list(gr=gender),var)
ratio_obs_2=log(aa[aa$gr=="Male",-1]/aa[aa$gr=="Female",-1])

```

### create P-values for each ROI

```

p_value_2=NULL
for(k in 1:length(ratio_obs_2)) {
  p_value_2=c(p_value_2,mean(perm[,k]>=as.numeric(ratio_obs_2[k])))
}

```

### Apply multiple comparisons check using the hochberg procedure

```

Hochberg_p_value_2 <- rbind(ratio_obs_2, p_value_2)
Hochberg_p_value_2 <- Hochberg_p_value_2[-c(1),]
Hochberg_p_value_2 <- melt(Hochberg_p_value_2)
Hochberg_p_value_2 <- rename(Hochberg_p_value_2, Area = variable)
Hochberg_p_value_2 = Hochberg_p_value_2[order(Hochberg_p_value_2$value),]
Hochberg_p_value_2$Hochberg <- p.adjust(Hochberg_p_value_2$value, method = "hochberg")

```

### Assessing the average variance ratios and overall P-values

**Merge the files containing the variance ratios for dataset 1 (chilnr = 1) and dataset 2 (chilnr = 2)**

```
average_Ratios <- rbind(ratio_obs, ratio_obs_2)
save(average_Ratios, file = "Average_Ratios.RData")
average_Ratios <- colMeans(average_Ratios)
```

### Create overall *P*-values for every ROI

```
p_value=NULL
for(k in 1:length(average_Ratios)) {
  p_value=c(p_value,mean(perm[,k]>=as.numeric(average_Ratios[k])))
}
```

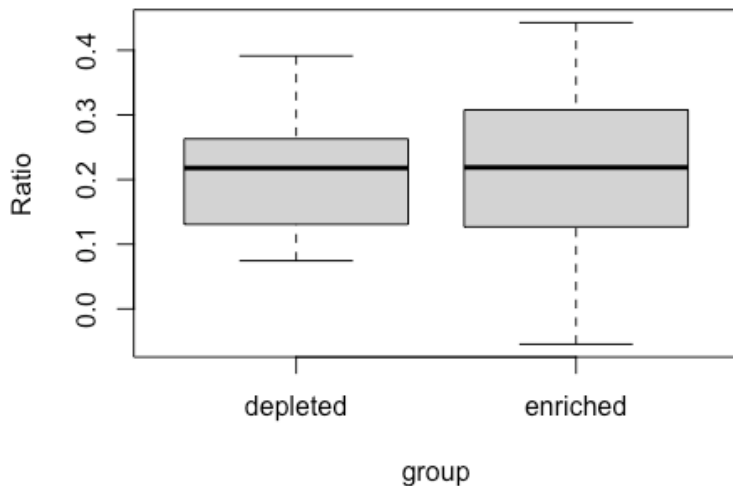
### apply multiple comparisons check to overall *P*-values

```
Hochberg_p_value_M <- rbind(average_Ratios, p_value)
Hochberg_p_value_M = t(Hochberg_p_value_M)
Hochberg_p_value_M <- as.data.frame(Hochberg_p_value_M)
Hochberg_p_value_M = Hochberg_p_value_M[order(Hochberg_p_value_M$p_value),]
Hochberg_p_value_M$Hochberg <- p.adjust(Hochberg_p_value_M$p_value, method = "hochberg")
```

### Comparing depleted and enriched regions with welsh t-test

```
vector_region <- c("depleted", "depleted", "depleted", "depleted", "depleted", "depleted", "depleted", "
depleted", "depleted", "depleted", "enriched", "enriched", "enriched", "enriched", "enriched", "enric
hed", "enriched", "enriched", "enriched", "enriched", "enriched", "enriched", "enriched", "enriche
d", "enriched", "enriched", "enriched", "enriched", "enriched", "enriched", "enriched", "enriched",
"enriched", "enriched", "enriched", "enriched")
average_Ratios <- as.data.frame(average_Ratios)
average_Ratios <- rename(average_Ratios, Ratio = average_Ratios)
average_Ratios$group <- vector_region
average_Ratios$group <- as.factor(average_Ratios$group)
boxplot(Ratio~group, data=average_Ratios)
```





```
t.test(Ratio~group, data=average_Ratios)
ME <- average_Ratios[average_Ratios$group == "enriched", ]
mean(ME$Ratio)
sd(ME$Ratio)
MD <- average_Ratios[average_Ratios$group == "depleted", ]
mean(MD$Ratio)
sd(MD$Ratio)
```

## Behavioral analyses SRS

### Create dataframe with Z values of ROI sizes

```
SRS_MRI_data <- SRS_MRI_data %>% mutate_if(is.integer, as.numeric)
SRS_MRI_data$age_child <- as.numeric(SRS_MRI_data$age_child)
SRS_MRI_data <- SRS_MRI_data %>% mutate_if(is.character, as.factor)

Z_SRS_MRI_data_male <- SRS_MRI_data[SRS_MRI_data$sex_child == "Male", ]
Z_SRS_MRI_data_male <- Z_SRS_MRI_data_male %>% mutate_at(c(11:46), funs(c(scale(.))))

Z_SRS_MRI_data_female <- SRS_MRI_data[SRS_MRI_data$sex_child == "Female", ]
Z_SRS_MRI_data_female <- Z_SRS_MRI_data_female %>% mutate_at(c(11:46), funs(c(scale(.))))

Z_SRS_MRI_data <- rbind(Z_SRS_MRI_data_male, Z_SRS_MRI_data_female)
Z_SRS_MRI_data <- Z_SRS_MRI_data[complete.cases(Z_SRS_MRI_data),]
```

### Create TwinNumber to take dependency into account

```

for (row in 1:nrow(Z_SRS_MRI_data)) {
  twinnr <- substr(Z_SRS_MRI_data$visit_id, 4, 7)
  Z_SRS_MRI_data$TwinNumber <- twinnr
}
Z_SRS_MRI_data$TwinNumber <- as.factor(Z_SRS_MRI_data$TwinNumber)

```

### Make dataset for the predicted values of SRS sum

```
data1_pred <- Z_SRS_MRI_data
```

### Loop over all ROIs to create Linear and quadratic model predicting SRS score with ROI size

```

Lvector = c()
Qvector = c()
LvectorAIC = c()
QvectorAIC = c()
SRS_L_ROI_Est = c()
SRS_L_ROI_t = c()
SRS_L_ROI_p = c()
SRS_L_sex_Est = c()
SRS_L_sex_t = c()
SRS_L_sex_p = c()
SRS_L_zyg_Est = c()
SRS_L_zyg_t = c()
SRS_L_zyg_p = c()
SRS_L_int_Est = c()
SRS_L_int_t = c()
SRS_L_int_p = c()
SRS_Q_ROI_Est = c()
SRS_Q_ROI_t = c()
SRS_Q_ROI_p = c()
SRS_Q_sex_Est = c()
SRS_Q_sex_t = c()
SRS_Q_sex_p = c()
SRS_Q_zyg_Est = c()
SRS_Q_zyg_t = c()
SRS_Q_zyg_p = c()
SRS_Q_int_Est = c()
SRS_Q_int_t = c()
SRS_Q_int_p = c()
BIC_srs_DF = NULL
count = 0
#loop over all rois
ROInames = names(Z_SRS_MRI_data[c(11:46)])
for(ROIname in ROInames) {
  #linear model
  count = count + 1
  Model1 = lmer(SRS_sum ~ Z_SRS_MRI_data[[ROIname]]*sex_child + zygotity_twin + (1|TwinNumber), REML=F, data=Z_SRS_MRI_data)
  Lvector <- c(Lvector, BIC(Model1))
  LvectorAIC <- c(LvectorAIC, AIC(Model1))
  #quadratic model
  Model2 = lmer(SRS_sum ~ poly(Z_SRS_MRI_data[[ROIname]], 2)*sex_child + zygotity_twin + (1|TwinNumber), REML=F, data=Z_SRS_MRI_data)

```

```

Qvector <- c(Qvector, BIC(Model2))
QvectorAIC <- c(QvectorAIC, AIC(Model2))
#get coefficients
SRS_L_ROI_Est <- c(SRS_L_ROI_Est, summary(Model1)$coefficients["Z_SRS_MRI_data[[ROIName]]", "Estimate"])
SRS_L_ROI_t <- c(SRS_L_ROI_t, summary(Model1)$coefficients["Z_SRS_MRI_data[[ROIName]]", "t value"])
SRS_L_ROI_p <- c(SRS_L_ROI_p, summary(Model1)$coefficients["Z_SRS_MRI_data[[ROIName]]", "Pr(>lt)"])
SRS_L_sex_Est <- c(SRS_L_sex_Est, summary(Model1)$coefficients["sex_childMale", "Estimate"])
)
SRS_L_sex_t <- c(SRS_L_sex_t, summary(Model1)$coefficients["sex_childMale", "t value"])
SRS_L_sex_p <- c(SRS_L_sex_p, summary(Model1)$coefficients["sex_childMale", "Pr(>lt)"])
SRS_L_zyg_Est <- c(SRS_L_zyg_Est, summary(Model1)$coefficients["zygosity_twinMZ", "Estimate"])
)
SRS_L_zyg_t <- c(SRS_L_zyg_t, summary(Model1)$coefficients["zygosity_twinMZ", "t value"])
SRS_L_zyg_p <- c(SRS_L_zyg_p, summary(Model1)$coefficients["zygosity_twinMZ", "Pr(>lt)"])
SRS_L_int_Est <- c(SRS_L_int_Est, summary(Model1)$coefficients["Z_SRS_MRI_data[[ROIName]]:sex_childMale", "Estimate"])
SRS_L_int_t <- c(SRS_L_int_t, summary(Model1)$coefficients["Z_SRS_MRI_data[[ROIName]]:sex_childMale", "t value"])
SRS_L_int_p <- c(SRS_L_int_p, summary(Model1)$coefficients["Z_SRS_MRI_data[[ROIName]]:sex_childMale", "Pr(>lt)"])
#quad
SRS_Q_ROI_Est <- c(SRS_Q_ROI_Est, summary(Model2)$coefficients["poly(Z_SRS_MRI_data[[ROIName]], 2)2", "Estimate"])
SRS_Q_ROI_t <- c(SRS_Q_ROI_t, summary(Model2)$coefficients["poly(Z_SRS_MRI_data[[ROIName]], 2)2", "t value"])
SRS_Q_ROI_p <- c(SRS_Q_ROI_p, summary(Model2)$coefficients["poly(Z_SRS_MRI_data[[ROIName]], 2)2", "Pr(>lt)"])
SRS_Q_sex_Est <- c(SRS_Q_sex_Est, summary(Model2)$coefficients["sex_childMale", "Estimate"])
)
SRS_Q_sex_t <- c(SRS_Q_sex_t, summary(Model2)$coefficients["sex_childMale", "t value"])
SRS_Q_sex_p <- c(SRS_Q_sex_p, summary(Model2)$coefficients["sex_childMale", "Pr(>lt)"])
SRS_Q_zyg_Est <- c(SRS_Q_zyg_Est, summary(Model2)$coefficients["zygosity_twinMZ", "Estimate"])
)
SRS_Q_zyg_t <- c(SRS_Q_zyg_t, summary(Model2)$coefficients["zygosity_twinMZ", "t value"])
SRS_Q_zyg_p <- c(SRS_Q_zyg_p, summary(Model2)$coefficients["zygosity_twinMZ", "Pr(>lt)"])
SRS_Q_int_Est <- c(SRS_Q_int_Est, summary(Model2)$coefficients["poly(Z_SRS_MRI_data[[ROIName]], 2)2:sex_childMale", "Estimate"])
SRS_Q_int_t <- c(SRS_Q_int_t, summary(Model2)$coefficients["poly(Z_SRS_MRI_data[[ROIName]], 2)2:sex_childMale", "t value"])
SRS_Q_int_p <- c(SRS_Q_int_p, summary(Model2)$coefficients["poly(Z_SRS_MRI_data[[ROIName]], 2)2:sex_childMale", "Pr(>lt)"])
#make predicted values
data1.pred[[as.symbol(paste0('pred', count))]] <- predict(Model2, newdata=data1.pred, level=0, na.action="na.omit")
}

```

### Create dataframe with BIC values, quadratic coefficient, and its *P*-value

```

BIC_srs_DF <- data.frame(Lvector, Qvector, SRS_L_ROI_Est, SRS_L_ROI_t, SRS_L_ROI_p, SRS_L_sex_Est, SRS_L_sex_t, SRS_L_sex_p, SRS_L_zyg_Est, SRS_L_zyg_t, SRS_L_zyg_p, SRS_L_int_Est, SRS_L_int_t, SRS_L_int_p, SRS_Q_ROI_Est, SRS_Q_ROI_t, SRS_Q_ROI_p, SRS_Q_sex_Est, SRS_Q_sex_t, SRS_Q_sex_p, SRS_Q_zyg_Est, SRS_Q_zyg_t, SRS_Q_zyg_p, SRS_Q_int_Est, SRS_Q_int_t, SRS_Q_int_p)

```

```
S_Q_int_t, SRS_Q_int_p)
BIC_srs_DF <- rename(BIC_srs_DF, c("LinearBIC" = "Lvector"))
BIC_srs_DF <- rename(BIC_srs_DF, c("QuadraticBIC" = "Qvector"))
BIC_srs_DF$Difference <- (BIC_srs_DF$QuadraticBIC < BIC_srs_DF$LinearBIC)
row.names(BIC_srs_DF) <- ROInames
```

### Add adjusted *P*-values to the dataframe

```
BIC_srs_DF = BIC_srs_DF[order(BIC_srs_DF$SRS_L_zyg_p),]
BIC_srs_DF$SRS_L_zyg_p_H <- p.adjust(BIC_srs_DF$SRS_L_zyg_p, method = "hochberg")
BIC_srs_DF = BIC_srs_DF[order(BIC_srs_DF$SRS_L_sex_p),]
BIC_srs_DF$SRS_L_sex_p_H <- p.adjust(BIC_srs_DF$SRS_L_sex_p, method = "hochberg")
BIC_srs_DF = BIC_srs_DF[order(BIC_srs_DF$SRS_Q_zyg_p),]
BIC_srs_DF$SRS_Q_zyg_p_H <- p.adjust(BIC_srs_DF$SRS_Q_zyg_p, method = "hochberg")
BIC_srs_DF = BIC_srs_DF[order(BIC_srs_DF$SRS_Q_sex_p),]
BIC_srs_DF$SRS_Q_sex_p_H <- p.adjust(BIC_srs_DF$SRS_Q_sex_p, method = "hochberg")
```

### Request means for groups

```
#calculate means for groups of sex and zygosity
aggregate(x = SRS_data$SRS_sum, by = list(SRS_data$sex_child), FUN = mean)
aggregate(x = SRS_data$SRS_sum, by = list(SRS_data$zygosity_twin), FUN = mean)
```

### Plot all quadratic models (SRS)

#### rh\_Pole\_occipital\_Area

```
SRS_plot1 <- ggplot(data=data1.pred, aes(rh_Pole_occipital_area, pred1, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") +
  ggtitle("rh_Pole_occipital_Area") + theme(panel.grid.major = element_blank(), panel.grid.minor = el
  ement_blank(),
  panel.background = element_blank())
```

#### lh\_Pole\_occipital area

```
SRS_plot2 <- ggplot(data=data1.pred, aes(lh_Pole_occipital_area, pred2, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") +
  ggtitle("lh_Pole_occipital area ") + theme(panel.grid.major = element_blank(), panel.grid.minor = ele
  ment_blank(),
  panel.background = element_blank())
```

#### rh\_G\_cuneus\_area

```
SRS_plot3 <- ggplot(data=data1.pred, aes(rh_G_cuneus_area, pred3, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") +
  ggtitle("rh_G_cuneus_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = elemen
  t_blank(),
  panel.background = element_blank())
```

**lh\_G\_cuneus\_area**

```
SRS_plot4 <- ggplot(data=data1.pred, aes(lh_G_cuneus_area, pred4, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") +
  ggtitle("lh_G_cuneus_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = elemen
  t_blank(),
  panel.background = element_blank())
```

**rh\_G\_oc.temp\_med.Lingual\_area**

```
SRS_plot5 <- ggplot(data=data1.pred, aes(rh_G_oc.temp_med.Lingual_area, pred5, color = sex_child)
) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") +
  ggtitle("rh_G_oc.temp_med.Lingual_area") + theme(panel.grid.major = element_blank(), panel.grid.
  minor = element_blank(),
  panel.background = element_blank())
```

**lh\_G\_oc.temp\_med.Lingual\_area**

```
SRS_plot6 <- ggplot(data=data1.pred, aes(lh_G_oc.temp_med.Lingual_area, pred6, color = sex_child)
) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") +
  ggtitle("lh_G_oc.temp_med.Lingual_area") + theme(panel.grid.major = element_blank(), panel.grid.
  minor = element_blank(),
  panel.background = element_blank())
```

**rh\_S\_calcarine\_area**

```
SRS_plot7 <- ggplot(data=data1.pred, aes(rh_S_calcarine_area, pred7, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_S_calcarine_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

**lh\_S\_calcarine\_area**

```
SRS_plot8 <- ggplot(data=data1.pred, aes(lh_S_calcarine_area, pred8, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("lh_S_calcarine_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

**rh\_S\_parieto\_occipital\_area**

```
SRS_plot9 <- ggplot(data=data1.pred, aes(rh_S_parieto_occipital_area, pred9, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_S_parieto_occipital_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

### lh\_S\_parieto\_occipital\_area

```
SRS_plot10 <- ggplot(data=data1.pred, aes(lh_S_parieto_occipital_area, pred10, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("lh_S_parieto_occipital_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

### rh\_S\_precentral.inf.part\_area

```
SRS_plot11 <- ggplot(data=data1.pred, aes(rh_S_precentral.inf.part_area, pred11, color = sex_child))
+
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_S_precentral.inf.part_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

### lh\_S\_precentral.inf.part\_area

```
SRS_plot12 <- ggplot(data=data1.pred, aes(lh_S_precentral.inf.part_area, pred12, color = sex_child))
+
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("lh_S_precentral.inf.part_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

### rh\_S\_precentral.sup.part\_area

```
SRS_plot13 <- ggplot(data=data1.pred, aes(rh_S_precentral.sup.part_area, pred13, color = sex_child))
+
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_S_precentral.sup.part_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
  panel.background = element_blank())
```

### lh\_S\_precentral.sup.part\_area

```
SRS_plot14 <- ggplot(data=data1.pred, aes(lh_S_precentral.sup.part_area, pred14, color = sex_child))
+
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
```

```
stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("lh_S_precentral.sup.part_area") +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank())
```

### rh\_G\_precentral\_area

```
SRS_plot15 <- ggplot(data=data1.pred, aes(rh_G_precentral_area, pred15, color = sex_child)) +
geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") +
stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("rh_G_precentral_area") +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank())
```

### lh\_G\_precentral\_area

```
SRS_plot16 <- ggplot(data=data1.pred, aes(lh_G_precentral_area, pred16, color = sex_child)) +
geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") +
stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("lh_G_precentral_area") +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank())
```

### rh\_S\_central\_area

```
SRS_plot17 <- ggplot(data=data1.pred, aes(rh_S_central_area, pred17, color = sex_child)) +
geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") +
stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("rh_S_central_area") +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank())
```

### lh\_S\_central\_area

```
SRS_plot18 <- ggplot(data=data1.pred, aes(lh_S_central_area, pred18, color = sex_child)) +
geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") +
stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("lh_S_central_area") +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
panel.background = element_blank())
```

### rh\_G.S\_subcentral\_area

```
SRS_plot19 <- ggplot(data=data1.pred, aes(rh_G.S_subcentral_area, pred19, color = sex_child)) +
geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") + xlab("ROI size") +
stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + ylab("predicted SRS score
") + ggtitle("rh_G.S_subcentral_area") +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### lh\_G.S\_subcentral\_area

```
SRS_plot20 <- ggplot(data=data1.pred, aes(lh_G.S_subcentral_area, pred20, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") +
  ggtitle("lh_G.S_subcentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = el
ement_blank(),
panel.background = element_blank())
```

### rh\_G\_postcentral\_area

```
SRS_plot21 <- ggplot(data=data1.pred, aes(rh_G_postcentral_area, pred21, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("rh_G_postcentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### lh\_G\_postcentral\_area

```
SRS_plot22 <- ggplot(data=data1.pred, aes(lh_G_postcentral_area, pred22, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("lh_G_postcentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### rh\_S\_postcentral\_area

```
SRS_plot23 <- ggplot(data=data1.pred, aes(rh_S_postcentral_area, pred23, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("rh_S_postcentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = e
lement_blank())
```

### lh\_S\_postcentral\_area

```
SRS_plot24 <- ggplot(data=data1.pred, aes(lh_S_postcentral_area, pred24, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SRS score") + ggtitle("lh_S_postcentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### rh\_G.S\_paracentral\_area

```
SRS_plot25 <- ggplot(data=data1.pred, aes(rh_G.S_paracentral_area, pred25, color = sex_child)) + ge
om_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI si
ze") + ylab("predicted SRS score") + ggtitle("rh_G.S_paracentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```



**lh\_G.S\_paracentral\_area**

```
SRS_plot26 <- ggplot(data=data1.pred, aes(lh_G.S_paracentral_area, pred26, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") +
  ylab("predicted SRS score") + ggtitle("lh_G.S_paracentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

**rh\_G.S\_cingul.Mid.Ant\_area**

```
SRS_plot27 <- ggplot(data=data1.pred, aes(rh_G.S_cingul.Mid.Ant_area, pred27, color = sex_child))
+
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_G.S_cingul.Mid.Ant_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

**lh\_G.S\_cingul.Mid.Ant\_area**

```
SRS_plot28 <- ggplot(data=data1.pred, aes(lh_G.S_cingul.Mid.Ant_area, pred28, color = sex_child))
+ geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
  OI size") + ylab("predicted SRS score") + ggtitle("lh_G.S_cingul.Mid.Ant_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

**rh\_G.S\_cingul.Mid.Post\_area**

```
SRS_plot29 <- ggplot(data=data1.pred, aes(rh_G.S_cingul.Mid.Post_area, pred29, color = sex_child))
+ geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE)
, colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
  OI size") + ylab("predicted SRS score") + ggtitle("rh_G.S_cingul.Mid.Post_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

**lh\_G.S\_cingul.Mid.Post\_area**

```
SRS_plot30 <- ggplot(data=data1.pred, aes(lh_G.S_cingul.Mid.Post_area, pred30, color = sex_child))
+ geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE)
, colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
  OI size") + ylab("predicted SRS score") + ggtitle("lh_G.S_cingul.Mid.Post_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

**rh\_G\_front\_middle\_area**

```
SRS_plot31 <- ggplot(data=data1.pred, aes(rh_G_front_middle_area, pred31, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("RO
  I size") + ylab("predicted SRS score") + ggtitle("rh_G_front_middle_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

**lh\_G\_front\_middle\_area**

```
SRS_plot32 <- ggplot(data=data1.pred, aes(lh_G_front_middle_area, pred31, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("lh_G_front_middle_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

### rh\_G\_front\_inf.Opercular\_area

```
SRS_plot33 <- ggplot(data=data1.pred, aes(rh_G_front_inf.Opercular_area, pred33, color = sex_child)
) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_G_front_inf.Opercular_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = e
  lement_blank())
```

### lh\_G\_front\_inf.Opercular\_area

```
SRS_plot34 <- ggplot(data=data1.pred, aes(lh_G_front_inf.Opercular_area, pred34, color = sex_child)
) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("lh_G_front_inf.Opercular_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

### rh\_Lat\_Fis.post\_area

```
SRS_plot35 <- ggplot(data=data1.pred, aes(rh_Lat_Fis.post_area, pred35, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("rh_Lat_Fis.post_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
  element_blank())
```

### lh\_Lat\_Fis.post\_area

```
SRS_plot36 <- ggplot(data=data1.pred, aes(lh_Lat_Fis.post_area, pred36, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
  predicted SRS score") + ggtitle("lh_Lat_Fis.post_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = e
  lement_blank())
```

### visualize all graphs in a grid

```
grid <- grid.arrange(SRS_plot1, SRS_plot2, SRS_plot3, SRS_plot4, nrow = 2)
```

```
grid2 <- grid.arrange(SRS_plot5, SRS_plot6, SRS_plot7, SRS_plot8, nrow = 2)
```

```
grid3 <- grid.arrange(SRS_plot9, SRS_plot10, SRS_plot11, SRS_plot12, nrow = 2)
```

```

grid4 <- grid.arrange(SRS_plot13, SRS_plot14, SRS_plot15, SRS_plot16, nrow = 2)
grid5 <- grid.arrange(SRS_plot17, SRS_plot18, SRS_plot19, SRS_plot20, nrow = 2)
grid6 <- grid.arrange(SRS_plot21, SRS_plot22, SRS_plot23, SRS_plot24, nrow = 2)
grid7 <- grid.arrange(SRS_plot25, SRS_plot26, SRS_plot27, SRS_plot28, nrow = 2)
grid8 <- grid.arrange(SRS_plot29, SRS_plot30, SRS_plot31, SRS_plot32, nrow = 2)
grid9 <- grid.arrange(SRS_plot33, SRS_plot34, SRS_plot35, SRS_plot36, nrow = 2)

```

## Behavioral analyses SMS

### Create dataframe with Z values of ROI sizes

```

SMS_MRI_data$age_child <- as.numeric(SMS_MRI_data$age_child)
SMS_MRI_data <- SMS_MRI_data %>% mutate_if(is.integer, as.numeric)
SMS_MRI_data <- SMS_MRI_data %>% mutate_if(is.character, as.factor)

Z_SMS_MRI_data_male <- SMS_MRI_data[SRS_MRI_data$sex_child == "Male", ]
Z_SMS_MRI_data_male <- Z_SMS_MRI_data_male %>% mutate_at(c(11:46), funs(c(scale(.))))

Z_SMS_MRI_data_female <- SMS_MRI_data[SMS_MRI_data$sex_child == "Female", ]
Z_SMS_MRI_data_female <- Z_SMS_MRI_data_female %>% mutate_at(c(11:46), funs(c(scale(.))))

Z_SMS_MRI_data <- rbind(Z_SMS_MRI_data_male, Z_SMS_MRI_data_female)
Z_SMS_MRI_data <- Z_SMS_MRI_data[complete.cases(Z_SMS_MRI_data),]

```

### Create TwinNumber to take dependency into account

```

for (row in 1:nrow(Z_SMS_MRI_data)) {
  twinnr <- substr(Z_SMS_MRI_data$visit_id, 4, 7)
  Z_SMS_MRI_data$TwinNumber <- twinnr
}
Z_SMS_MRI_data$TwinNumber <- as.factor(Z_SMS_MRI_data$TwinNumber)

```

### Create dataset for predicted values of SRS

```
data2_pred <- Z_SMS_MRI_data
```

### Loop over all ROIs to create Linear and quadratic model predicting SRS score with ROI size

```

Lvector2 = c()
Qvector2 = c()
LvectorAIC2 = c()
QvectorAIC2 = c()
anova2 = c()
SMS_L_ROI_Est = c()
SMS_L_ROI_t = c()
SMS_L_ROI_p = c()
SMS_L_sex_Est = c()
SMS_L_sex_t = c()
SMS_L_sex_p = c()
SMS_L_zyg_Est = c()

```

```

SMS_L_zyg_t = c()
SMS_L_zyg_p = c()
SMS_L_int_Est = c()
SMS_L_int_t = c()
SMS_L_int_p = c()
SMS_Q_ROI_Est = c()
SMS_Q_ROI_t = c()
SMS_Q_ROI_p = c()
SMS_Q_sex_Est = c()
SMS_Q_sex_t = c()
SMS_Q_sex_p = c()
SMS_Q_zyg_Est = c()
SMS_Q_zyg_t = c()
SMS_Q_zyg_p = c()
SMS_Q_int_Est = c()
SMS_Q_int_t = c()
SMS_Q_int_p = c()
BIC_SMS_DF = NULL
count = 0
#loop over all rois
ROInames = names(Z_SMS_MRI_data[c(11:46)])
for(ROIname in ROInames) {
  #linear model
  count = count + 1
  Model1 = lmer(SMS_score ~ Z_SMS_MRI_data[[ROIname]]*sex_child + zygosity_twin + (1|Twin
Number) , REML=F , data=Z_SMS_MRI_data)
  Lvector2 <- c(Lvector2, BIC(Model1))
  LvectorAIC2 <- c(LvectorAIC2, AIC(Model1))
  #quadratic model
  Model2 = lmer(SMS_score ~ poly(Z_SMS_MRI_data[[ROIname]], 2)*sex_child + zygosity_twin +
(1|TwinNumber) , REML=F , data=Z_SMS_MRI_data)
  Qvector2 <- c(Qvector2, BIC(Model2))
  QvectorAIC2 <- c(QvectorAIC2, AIC(Model2))
  #get coefficients
  SMS_L_ROI_Est <- c(SMS_L_ROI_Est, summary(Model1)$coefficients["Z_SMS_MRI_data[[ROI
name]]", "Estimate"])
  SMS_L_ROI_t <- c(SMS_L_ROI_t, summary(Model1)$coefficients["Z_SMS_MRI_data[[ROIname
]]", "t value"])
  SMS_L_ROI_p <- c(SMS_L_ROI_p, summary(Model1)$coefficients["Z_SMS_MRI_data[[ROInam
e]]", "Pr(>|t|)"])
  SMS_L_sex_Est <- c(SMS_L_sex_Est, summary(Model1)$coefficients["sex_childMale", "Estimate"
])
  SMS_L_sex_t <- c(SMS_L_sex_t, summary(Model1)$coefficients["sex_childMale", "t value"])
  SMS_L_sex_p <- c(SMS_L_sex_p, summary(Model1)$coefficients["sex_childMale", "Pr(>|t|)"])
  SMS_L_zyg_Est <- c(SMS_L_zyg_Est, summary(Model1)$coefficients["zygosity_twinMZ", "Estim
ate"])
  SMS_L_zyg_t <- c(SMS_L_zyg_t, summary(Model1)$coefficients["zygosity_twinMZ", "t value"])
  SMS_L_zyg_p <- c(SMS_L_zyg_p, summary(Model1)$coefficients["zygosity_twinMZ", "Pr(>|t|)"])
  SMS_L_int_Est <- c(SMS_L_int_Est, summary(Model1)$coefficients["Z_SMS_MRI_data[[ROIname]]:sex_childMale", "Estimate"])
  SMS_L_int_t <- c(SMS_L_int_t, summary(Model1)$coefficients["Z_SMS_MRI_data[[ROIname]]:s
ex_childMale", "t value"])
  SMS_L_int_p <- c(SMS_L_int_p, summary(Model1)$coefficients["Z_SMS_MRI_data[[ROIname]]:
sex_childMale", "Pr(>|t|)"])
  #quad

```

```

SMS_Q_ROI_Est <- c(SMS_Q_ROI_Est, summary(Model2)$coefficients["poly(Z_SMS_MRI_data[
ROIname]], 2)2", "Estimate"])
SMS_Q_ROI_t <- c(SMS_Q_ROI_t, summary(Model2)$coefficients["poly(Z_SMS_MRI_data[[ROI
name]], 2)2", "t value"])
SMS_Q_ROI_p <- c(SMS_Q_ROI_p, summary(Model2)$coefficients["poly(Z_SMS_MRI_data[[R
OIname]], 2)2", "Pr(>|t|)"])
SMS_Q_sex_Est <- c(SMS_Q_sex_Est, summary(Model2)$coefficients["sex_childMale", "Estimate
"])
SMS_Q_sex_t <- c(SMS_Q_sex_t, summary(Model2)$coefficients["sex_childMale", "t value"])
SMS_Q_sex_p <- c(SMS_Q_sex_p, summary(Model2)$coefficients["sex_childMale", "Pr(>|t|)"])
SMS_Q_zyg_Est <- c(SMS_Q_zyg_Est, summary(Model2)$coefficients["zygosity_twinMZ", "Esti
mate"])
SMS_Q_zyg_t <- c(SMS_Q_zyg_t, summary(Model2)$coefficients["zygosity_twinMZ", "t value"])
SMS_Q_zyg_p <- c(SMS_Q_zyg_p, summary(Model2)$coefficients["zygosity_twinMZ", "Pr(>|t|)"]
)
SMS_Q_int_Est <- c(SMS_Q_int_Est, summary(Model2)$coefficients["poly(Z_SMS_MRI_data[[R
OIname]], 2)2:sex_childMale", "Estimate"])
SMS_Q_int_t <- c(SMS_Q_int_t, summary(Model2)$coefficients["poly(Z_SMS_MRI_data[[ROIname
]], 2)2:sex_childMale", "t value"])
SMS_Q_int_p <- c(SMS_Q_int_p, summary(Model2)$coefficients["poly(Z_SMS_MRI_data[[ROIname
]], 2)2:sex_childMale", "Pr(>|t|)"])
#make predicted values
data2.pred[[as.symbol(paste0('pred', count))]] <- predict(Model2, newdata=data2.pred, level=0, na.ac
tion="na.omit")
}

```

### Create dataframe with BIC values, quadratic coefficient, and it's P-value

```

BIC_SMS_DF <- data.frame(Lvector2, Qvector2, SMS_L_ROI_Est, SMS_L_ROI_t, SMS_L_ROI_p,
SMS_L_sex_Est, SMS_L_sex_t, SMS_L_sex_p, SMS_L_zyg_Est, SMS_L_zyg_t, SMS_L_zyg_p, S
MS_L_int_Est, SMS_L_int_t, SMS_L_int_p, SMS_Q_ROI_Est, SMS_Q_ROI_t, SMS_Q_ROI_p, S
MS_Q_sex_Est, SMS_Q_sex_t, SMS_Q_sex_p, SMS_Q_zyg_Est, SMS_Q_zyg_t, SMS_Q_zyg_p, S
MS_Q_int_Est, SMS_Q_int_t, SMS_Q_int_p)
BIC_SMS_DF <- rename(BIC_SMS_DF, c("LinearBIC" = "Lvector2"))
BIC_SMS_DF <- rename(BIC_SMS_DF, c("QuadraticBIC" = "Qvector2"))
BIC_SMS_DF$Difference <- (BIC_SMS_DF$QuadraticBIC < BIC_SMS_DF$LinearBIC)
row.names(BIC_SMS_DF) <- ROInames

```

### Add adjusted P-values to the dataframe

```

BIC_SMS_DF = BIC_SMS_DF[order(BIC_SMS_DF$SMS_L_ROI_p),]
BIC_SMS_DF$SMS_L_ROI_p_H <- p.adjust(BIC_SMS_DF$SMS_L_ROI_p, method = "hochberg")
BIC_SMS_DF = BIC_SMS_DF[order(BIC_SMS_DF$SMS_L_int_p),]
BIC_SMS_DF$SMS_L_int_p_H <- p.adjust(BIC_SMS_DF$SMS_L_int_p, method = "hochberg")
BIC_SMS_DF = BIC_SMS_DF[order(BIC_SMS_DF$SMS_Q_ROI_p),]
BIC_SMS_DF$SMS_Q_ROI_p_H <- p.adjust(BIC_SMS_DF$SMS_Q_ROI_p, method = "hochberg"
)
BIC_SMS_DF = BIC_SMS_DF[order(BIC_SMS_DF$SMS_Q_int_p),]
BIC_SMS_DF$SMS_Q_int_p_H <- p.adjust(BIC_SMS_DF$SMS_Q_int_p, method = "hochberg")

```

### Plot all quadratic models (SRS)

#### rh\_Pole\_occipital\_Area

```

SMS_plot1 <- ggplot(data=data2.pred, aes(rh_Pole_occipital_area, pred1, color = sex_child)) +
geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c

```

```
colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SMS score") + ggtitle("rh_Pole_occipital_Area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### lh\_Pole\_occipital\_Area

```
SMS_plot2 <- ggplot(data=data2.pred, aes(lh_Pole_occipital_area, pred2, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("RO
I size") + ylab("predicted SMS score") + ggtitle("lh_Pole_occipital area ") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### rh\_G\_cuneus\_area

```
SMS_plot3 <- ggplot(data=data2.pred, aes(rh_G_cuneus_area, pred3, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SMS score") + ggtitle("rh_G_cuneus_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### lh\_G\_cuneus\_area

```
SMS_plot4 <- ggplot(data=data2.pred, aes(lh_G_cuneus_area, pred4, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),c
olour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("
predicted SMS score") + ggtitle("lh_G_cuneus_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### rh\_G\_oc.temp\_med.Lingual\_area

```
SMS_plot5 <- ggplot(data=data2.pred, aes(rh_G_oc.temp_med.Lingual_area, pred5, color = sex_child
)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRU
E), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("
ROI size") + ylab("predicted SMS score") + ggtitle("rh_G_oc.temp_med.Lingual_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = el
ement_blank())
```

### lh\_G\_oc.temp\_med.Lingual\_area

```
SMS_plot6 <- ggplot(data=data2.pred, aes(lh_G_oc.temp_med.Lingual_area, pred6, color = sex_child
)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE
), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
OI size") + ylab("predicted SMS score") + ggtitle("lh_G_oc.temp_med.Lingual_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())
```

### rh\_S\_calcarine\_area

```
SMS_plot7 <- ggplot(data=data2.pred, aes(rh_S_calcarine_area, pred7, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),co
```

```

colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI
size") + ylab("predicted SMS score") + ggtitle("rh_S_calcarine_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())

```

### lh\_S\_calcarine\_area

```

SMS_plot8 <- ggplot(data=data2.pred, aes(lh_S_calcarine_area, pred8, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), c
olour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI
size") + ylab("predicted SMS score") + ggtitle("lh_S_calcarine_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())

```

### rh\_S\_parieto\_occipital\_area

```

SMS_plot9 <- ggplot(data=data2.pred, aes(rh_S_parieto_occipital_area, pred9, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), c
olour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI
size") + ylab("predicted SMS score") + ggtitle("rh_S_parieto_occipital_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())

```

### lh\_S\_parieto\_occipital\_area

```

SMS_plot10 <- ggplot(data=data2.pred, aes(lh_S_parieto_occipital_area, pred10, color = sex_child))
+ geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
OI size") + ylab("predicted SMS score") + ggtitle("lh_S_parieto_occipital_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())

```

### rh\_S\_precentral.inf.part\_area

```

SMS_plot11 <- ggplot(data=data2.pred, aes(rh_S_precentral.inf.part_area, pred11, color = sex_child))
+ geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
OI size") + ylab("predicted SMS score") + ggtitle("rh_S_precentral.inf.part_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())

```

### lh\_S\_precentral.inf.part\_area

```

SMS_plot12 <- ggplot(data=data2.pred, aes(lh_S_precentral.inf.part_area, pred11, color = sex_child))
+ geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
OI size") + ylab("predicted SMS score") + ggtitle("lh_S_precentral.inf.part_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background =
element_blank())

```

### rh\_S\_precentral.sup.part\_area

```

SMS_plot13 <- ggplot(data=data2.pred, aes(rh_S_precentral.sup.part_area, pred13, color = sex_child)
) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE)
), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("R
OI size") + ylab("predicted SMS score") + ggtitle("rh_S_precentral.sup.part_area") +

```

```
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_S\_precentral.sup.part\_area

```
SMS_plot14 <- ggplot(data=data2.pred, aes(lh_S_precentral.sup.part_area, pred14, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") +
  ylab("predicted SMS score") + ggtitle("lh_S_precentral.sup.part_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G\_precentral\_area

```
SMS_plot15 <- ggplot(data=data2.pred, aes(rh_G_precentral_area, pred15, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") +
  ylab("predicted SMS score") + ggtitle("rh_G_precentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G\_precentral\_area

```
SMS_plot16 <- ggplot(data=data2.pred, aes(lh_G_precentral_area, pred16, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") +
  ylab("predicted SMS score") + ggtitle("lh_G_precentral_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_S\_central\_area

```
SMS_plot17 <- ggplot(data=data2.pred, aes(rh_S_central_area, pred17, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") +
  ggtitle("rh_S_central_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_S\_central\_area

```
SMS_plot18 <- ggplot(data=data2.pred, aes(lh_S_central_area, pred18, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") +
  stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") +
  ggtitle("lh_S_central_area") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G.S\_subcentral\_area

```
SMS_plot19 <- ggplot(data=data2.pred, aes(rh_G.S_subcentral_area, pred19, color = sex_child)) +
  geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE),
  colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") +
  ylab("predicted SMS score") + ggtitle("rh_G.S_subcentral_area") +
```



```
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G.S\_subcentral\_area

```
SMS_plot20 <- ggplot(data=data2.pred, aes(lh_G.S_subcentral_area, pred20, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G.S_subcentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G\_postcentral\_area

```
SMS_plot21 <- ggplot(data=data2.pred, aes(rh_G_postcentral_area, pred21, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_G_postcentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G\_postcentral\_area

```
SMS_plot22 <- ggplot(data=data2.pred, aes(lh_G_postcentral_area, pred22, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G_postcentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_S\_postcentral\_area

```
SMS_plot23 <- ggplot(data=data2.pred, aes(rh_S_postcentral_area, pred23, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_S_postcentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_S\_postcentral\_area

```
SMS_plot24 <- ggplot(data=data2.pred, aes(lh_S_postcentral_area, pred24, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_S_postcentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G.S\_paracentral\_area

```
SMS_plot25 <- ggplot(data=data2.pred, aes(rh_G.S_paracentral_area, pred25, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_G.S_paracentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G.S\_paracentral\_area

```
SMS_plot26 <- ggplot(data=data2.pred, aes(lh_G.S_paracentral_area, pred26, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G.S_paracentral_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G.S\_cingul.Mid.Ant\_area

```
SMS_plot27 <- ggplot(data=data2.pred, aes(rh_G.S_cingul.Mid.Ant_area, pred27, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_G.S_cingul.Mid.Ant_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G.S\_cingul.Mid.Ant\_area

```
SMS_plot28 <- ggplot(data=data2.pred, aes(lh_G.S_cingul.Mid.Ant_area, pred28, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G.S_cingul.Mid.Ant_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G.S\_cingul.Mid.Post\_area

```
SMS_plot29 <- ggplot(data=data2.pred, aes(rh_G.S_cingul.Mid.Post_area, pred29, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_G.S_cingul.Mid.Post_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G.S\_cingul.Mid.Post\_area

```
SMS_plot30 <- ggplot(data=data2.pred, aes(lh_G.S_cingul.Mid.Post_area, pred30, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G.S_cingul.Mid.Post_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_G\_front\_middle\_area

```
SMS_plot31 <- ggplot(data=data2.pred, aes(rh_G_front_middle_area, pred31, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_G_front_middle_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G\_front\_middle\_area

```
SMS_plot32 <- ggplot(data=data2.pred, aes(lh_G_front_middle_area, pred31, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G_front_middle_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### #rh\_G\_front\_inf.Opercular\_area

```
SMS_plot33 <- ggplot(data=data2.pred, aes(rh_G_front_inf.Opercular_area, pred33, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_G_front_inf.Opercular_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_G\_front\_inf.Opercular\_area

```
SMS_plot34 <- ggplot(data=data2.pred, aes(lh_G_front_inf.Opercular_area, pred34, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_G_front_inf.Opercular_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### rh\_Lat\_Fis.post\_area

```
SMS_plot35 <- ggplot(data=data2.pred, aes(rh_Lat_Fis.post_area, pred35, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("rh_Lat_Fis.post_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### lh\_Lat\_Fis.post\_area

```
SMS_plot36 <- ggplot(data=data2.pred, aes(lh_Lat_Fis.post_area, pred36, color = sex_child)) + geom_point() + stat_smooth(method="lm", se=TRUE, fill=NA, formula=y ~ poly(x, 2, raw=TRUE), colour="black") + stat_smooth(method="gam", formula=y ~ poly(x, 2, raw=TRUE), se=F) + xlab("ROI size") + ylab("predicted SMS score") + ggtitle("lh_Lat_Fis.post_area") + theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), panel.background = element_blank())
```

### Visualize all graphs in a grid

```
SMS_grid <- grid.arrange(SMS_plot1, SMS_plot2, SMS_plot3, SMS_plot4, nrow = 2)
```

```
SMS_grid2 <- grid.arrange(SMS_plot5, SMS_plot6, SMS_plot7, SMS_plot8, nrow = 2)
```

```
SMS_grid3 <- grid.arrange(SMS_plot9, SMS_plot10, SMS_plot11, SMS_plot12, nrow = 2)
```

```
SMS_grid4 <- grid.arrange(SMS_plot13, SMS_plot14, SMS_plot15, SMS_plot16, nrow = 2)
```

```

SMS_grid5 <- grid.arrange(SMS_plot17, SMS_plot18, SMS_plot19, SMS_plot20, nrow = 2)
SMS_grid6 <- grid.arrange(SMS_plot21, SMS_plot22, SMS_plot23, SMS_plot24, nrow = 2)
SMS_grid7 <- grid.arrange(SMS_plot25, SMS_plot26, SMS_plot27, SMS_plot28, nrow = 2)
SMS_grid8 <- grid.arrange(SMS_plot29, SMS_plot30, SMS_plot31, SMS_plot32, nrow = 2)
SMS_grid9 <- grid.arrange(SMS_plot33, SMS_plot34, SMS_plot35, SMS_plot36, nrow = 2)
SMS_grid_SigRatioDepl <- grid.arrange(SMS_plot7, SMS_plot8, SMS_plot10, nrow = 2)
SMS_grid_SigRatioEnr <- grid.arrange(SMS_plot13, SMS_plot19, SMS_plot24, SMS_plot28, SMS_
plot32, SMS_plot34, SMS_plot36, nrow = 3)

combine1 <- SMS_plot36 + labs(colour = "Sex")
combine2 <- SMS_plot15 + labs(colour = "Sex")
ggarrange(combine1, combine2, labels = c("A", "B"),
common.legend = TRUE, legend = "bottom")

```

## Create graph for variance ratios

```

dev.off()
Hochberg_p_value_M <- rename(Hochberg_p_value_M, "Ratio" = "average_Ratios")
average_Ratios <- rename(average_Ratios, "Group" = "group")
hist_file <- merge(average_Ratios, Hochberg_p_value_M, by.x = c("row.names", "Ratio"), by.y = c("r
ow.names", "Ratio"))

p <- ggplot(hist_file, aes(x = factor(Row.names), y = Ratio, width = 0.5)) +
  geom_col(aes(fill = Group), width = 0.7) + coord_flip() +
  geom_text(label = ifelse(hist_file$p_value < 0.05, "*", ""), size = 8, position = position_dodge(width
= 2), hjust = -0.5) +
  xlab("Ratio") + ylab("ROI") +
  ggtitle("Average ratio per ROI") +
  theme(plot.title = element_text(hjust = 0.5), axis.text = element_text(size=10)) +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank())

```

## Post-hoc analyses

### Looking at AIC model fit values as this is less stringent when compared to the BIC

```

SRS_AIC <- data.frame(LvectorAIC, QvectorAIC)
SRS_AIC$DifferenceAIC <- (SRS_AIC$QvectorAIC < SRS_AIC$LvectorAIC)
row.names(SRS_AIC) <- ROInames
SRS_AIC
SRS_AIC_DF <- merge(SRS_AIC, BIC_srs_DF, by=0)
row.names(SRS_AIC_DF) <- SRS_AIC_DF$Row.names
SRS_AIC_DF = subset(SRS_AIC_DF, select = -c(LinearBIC, QuadraticBIC, Difference, Row.names)
)
SMS_AIC <- data.frame(LvectorAIC2, QvectorAIC2)
SMS_AIC$DifferenceAIC <- (SMS_AIC$QvectorAIC2 < SMS_AIC$LvectorAIC2)
row.names(SMS_AIC) <- ROInames
SMS_AIC
SMS_AIC_DF <- merge(SMS_AIC, BIC_SMS_DF, by=0)

```

```
row.names(SMS_AIC_DF) <- SMS_AIC_DF$Row.names
SMS_AIC_DF = subset(SMS_AIC_DF, select = -c(LinearBIC, QuadraticBIC, Difference, Row.names
))
```

### Getting handedness for confirming the finding of the left precentral gyrus

```
lcid_ecc_behavioral <- read.csv2("lcid_ecc_voorLotte2.csv")
lcid_mcc_behavioral <- read.csv2("lcid_mcc_voorLotte2.csv")
lcid_ecc_hand = subset(lcid_ecc_behavioral, select = c(visit_id, ecc_handedness))
lcid_mcc_hand = subset(lcid_mcc_behavioral, select = c(sub_id, mcc_handedness))
lcid_mcc_hand <- rename(lcid_mcc_hand, visit_id = sub_id)
lcid_ecc_hand <- rename(lcid_ecc_hand, handedness = ecc_handedness)
lcid_mcc_hand <- rename(lcid_mcc_hand, handedness = mcc_handedness)
Hand_data <- rbind(lcid_ecc_hand, lcid_mcc_hand)
```

```
for (row in 1:nrow(Hand_data)) {
  ID <- substr(Hand_data$visit_id, 5, 13)
}
for (row in 1:nrow(Hand_data)) {
  Hand_data$visit_id <- ID
}
```

```
Hand_data$handedness <- as.factor(Hand_data$handedness)
levels(Hand_data$handedness)
Hand_SMS_MRI_data <- merge(SMS_MRI_data, Hand_data, by.x = "visit_id", by.y = "visit_id")
table(Hand_SMS_MRI_data$handedness)/sum(table(Hand_SMS_MRI_data$handedness))
```

### Testing whether there is a mean difference between monozygotic and dizygotic twins srs score when accounting for variance differences

```
var.test(SRS_sum ~ zygosity_twin, SRS_MRI_data)
var.test(SRS_sum ~ sex_child, SRS_MRI_data)

t.test(SRS_sum~zygosity_twin, data=SRS_MRI_data)
t.test(SRS_sum~sex_child, data=SRS_MRI_data)

t.test(SMS_score~zygosity_twin, data=SMS_MRI_data)
t.test(SMS_score~sex_child, data=SMS_MRI_data)
```

### Assess correlation between SRS and SMS

```
SMS_score = subset(SMS_MRI_data, select = c(visit_id, SMS_score, sex_child))
SRS_sum = subset(SRS_MRI_data, select = c(visit_id, SRS_sum))
Corr_DF = merge(SMS_score, SRS_sum, by.x = "visit_id", by.y = "visit_id")
for (row in 1:nrow(Corr_DF)) {
  twinnr <- substr(Corr_DF$visit_id, 9, 11)
  Corr_DF$TwinNumber <- twinnr
}
Corr_DF1 <- Corr_DF[Corr_DF$TwinNumber == "1", ]
cor.test(Corr_DF1$SMS_score, Corr_DF1$SRS_sum, method = "pearson")

Corr_DF2 <- Corr_DF[Corr_DF$TwinNumber == "2", ]
cor.test(Corr_DF2$SMS_score, Corr_DF2$SRS_sum, method = "pearson")
```

### Creating a histogram for variance ratios

```

hist_file <- inner_join(average_Ratios, Hochberg_p_value_M)
row.names(hist_file) <- row.names(average_Ratios)
hist_file$ROInames <- row.names(hist_file)

hist_file$ROInames <- factor(hist_file$ROInames, levels = hist_file$ROInames[order(hist_file$Group
)])

p <- ggplot(hist_file, aes(x = factor(ROInames), y = Ratio, width = 0.5)) +
  geom_col(aes(fill = Group), width = 0.7) + coord_flip() +
  geom_text(label = ifelse(hist_file$p_value < 0.05, "*", ""), size = 8, position = position_dodge(width
= 2), hjust = -0.5) +
  xlab("ROI name") + ylab("Variance ratio") +
  ggtitle("Average ratio per ROI with significance for uncorrected p-values") +
  theme(plot.title = element_text(hjust = 0.5), axis.text = element_text(size=10)) +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank())

```

## Demographics

### Variance ratios

```

demodata <- rbind(Ratio_data, Ratio_data_2)
demodata$age_child <- as.numeric(demodata$age_child)
mean(demodata$age_child)
range(demodata$age_child)
sd(demodata$age_child)
table(demodata$sex_child)
for (row in 1:nrow(lcid_ecc_behavioral)) {
  ID <- substr(lcid_ecc_behavioral$visit_id, 5, 13)
}
for (row in 1:nrow(lcid_ecc_behavioral)) {
  lcid_ecc_behavioral$visit_id <- ID
}

for (row in 1:nrow(lcid_mcc_behavioral)) {
  ID <- substr(lcid_mcc_behavioral$sub_id, 5, 13)
}
for (row in 1:nrow(lcid_mcc_behavioral)) {
  lcid_mcc_behavioral$visit_id <- ID
}

Zyg_data_ecc = subset(lcid_ecc_behavioral, select = c(visit_id, ecc_zygosity))
Zyg_data_mcc = subset(lcid_mcc_behavioral, select = c(visit_id, mcc_zygosity))
Zyg_data_ecc <- rename(Zyg_data_ecc, zygosity = ecc_zygosity)
Zyg_data_mcc <- rename(Zyg_data_mcc, zygosity = mcc_zygosity)
zyg_data <- rbind(Zyg_data_ecc, Zyg_data_mcc)
demodata <- merge(demodata, zyg_data, by.x = "visit_id", by.y = "visit_id")
table(demodata$zygosity)

VF <- demodata[demodata$sex_child == "Female", ]
mean(VF$age_child)
range(VF$age_child)
sd(VF$age_child)
table(VF$zygosity_twin)

```

```
VM <- demodata[demodata$sex_child == "Male", ]
mean(VM$age_child)
range(VM$age_child)
sd(VM$age_child)
table(VM$zygosity_twin)
```

### SRS data

```
mean(Z_SRS_MRI_data$age_child, na.rm = TRUE)
range(Z_SRS_MRI_data$age_child, na.rm = TRUE)
sd(Z_SRS_MRI_data$age_child)
table(Z_SRS_MRI_data$sex_child)
table(Z_SRS_MRI_data$zygosity_twin)
table(Z_SRS_MRI_data$zygosity, Z_SRS_MRI_data$sex_child)
range(Z_SRS_MRI_data$SRS_sum)

VF <- Z_SRS_MRI_data[Z_SRS_MRI_data$sex_child == "Female", ]
mean(VF$age_child, na.rm = TRUE)
range(VF$age_child, na.rm = TRUE)
sd(VF$age_child)
sd(VF$SRS_sum)

VM <- Z_SRS_MRI_data[Z_SRS_MRI_data$sex_child == "Male", ]
mean(VM$age_child)
range(VM$age_child)
sd(VM$age_child)
sd(VM$SRS_sum)

VMZ <- Z_SRS_MRI_data[Z_SRS_MRI_data$zygosity_twin == "MZ", ]
range(VMZ$SRS_sum)
var(VMZ$SRS_sum)
sd(VMZ$SRS_sum)

VDZ <- Z_SRS_MRI_data[Z_SRS_MRI_data$zygosity_twin == "DZ", ]
range(VDZ$SRS_sum)
var(VDZ$SRS_sum)
sd(VDZ$SRS_sum)
```

### SMS data

```
mean(Z_SMS_MRI_data$age_child, na.rm = TRUE)
range(Z_SMS_MRI_data$age_child, na.rm = TRUE)
sd(Z_SMS_MRI_data$age_child)
table(Z_SMS_MRI_data$sex_child)
table(Z_SMS_MRI_data$zygosity_twin)

table(Z_SMS_MRI_data$zygosity, Z_SMS_MRI_data$sex_child)

VF <- Z_SMS_MRI_data[Z_SMS_MRI_data$sex_child == "Female", ]
mean(VF$age_child, na.rm = TRUE)
range(VF$age_child, na.rm = TRUE)
sd(VF$age_child)

VM <- Z_SMS_MRI_data[Z_SMS_MRI_data$sex_child == "Male", ]
mean(VM$age_child)
```

```
range(VM$age_child)
sd(VM$age_child)
```