INTELLIGENCE AND THE STRUCTURES OF THE LINGUAL GYRI

Emil Norrman
ACKNOWLEDGEMENTS

This thesis would have been a lot harder to finish without the excellent supervision of Carola. Not only has her feedback been useful in all sorts of ways, her responses to various questions and issues have also been very rapid. For that I am more than grateful.

I am also grateful for both feedback and being able to discuss diverse aspects of my thesis with colleagues at Karolinska Institute.

Finally, I am in debt to all those that contributed to the open science initiative OpenNeuro. Without the freely available neuroimaging and behavioral data this thesis would not have been possible.
Abstract

Finding neural correlates of intelligence and cognitive abilities in the developing brain during childhood may be important in many ways, such as predicting and understanding educational abilities or making clinical evaluations of patients. Even if substantial contemporary research has established relationships between brain structures and general intelligence, little is known about the lingual gyri and their links to IQ. In this thesis it is examined (1) whether cortical thickness in the right and left lingual gyri is associated with different levels of IQ in children and (2) if the rate of change in cortical thickness located in the lingual gyri is associated with change in Performance IQ (PIQ). Neuroimaging data originated from a study by Solé-Casals and colleagues (2019) as well as a dataset from a study by Suárez-Pellicioni and colleagues (2019). Both datasets were downloaded from the OpenNeuro library of brain imaging data. Neuroimaging metrics of twenty-nine boys of approximately twelve years of age were utilized to test the hypothesis that higher IQ is related to thinner cortical thickness in lingual gyri. Neuroimaging metrics of twenty-one girls and fifteen boys under fourteen years of age utilized to examine if the rate of change in cortical thickness is related to a change in Performance IQ. Results revealed that high IQ was related to thinner cortical thickness at the age of twelve. Further results indicated that rate of thinning of the cortex in the lingual gyri is correlated to change in Performance IQ. The present thesis adds to the growing evidence that regional cortical thickness and change of cortical thickness are relevant biomarkers for intelligence. Future research with larger sample sizes and longitudinal design with additional points in time might be needed to confirm the results of the present thesis.

Keywords: IQ, cortical thickness, lingual gyri, biomarkers

Sammanfattning


Nyckelord: IQ, kortikal tjocklek, lingual gyri, biomarkörer
INTELLIGENCE AND THE STRUCTURES OF THE LINGUAL GYRI

Brief History of Intelligence and Its Constructs

Intelligence is defined by American Psychological Association as the “the ability to derive information, learn from experience, adapt to the environment, understand, and correctly utilize thought and reason.” (APA, 2015). As evident in the definition, intelligence can be considered as a relative broad ability with various theoretical perspectives. General intelligence is one of the theoretical constructs of intelligence. It is a factor derived from the notion that the results in different cognitive ability tests positively correlate with one another. (cf. Humphreys, 1979; Plucker & Shelton, 2015; Spearman, 1904; Spearman, 1927). Thus, persons with high scores on a certain test of a cognitive ability are likely to have high scores on different tests of cognitive abilities. Intelligence quotient (IQ) is a norm-referenced measure of the cognitive abilities of individuals (Deary, 2022). Historically, alternative models of intelligence, such as the Multiple Intelligences (Gardner, 1983) and the triarchic theory of intelligence (Sternberg, 1985) have also been proposed, but those theories seem not to have gained momentum in the scientific community (Gottfredson, 2003; Waterhouse, 2006; Ferrero, Vadillo, & León, 2021).

Measurements of intelligence and the theorized underlying g factor has been criticized. For instance, IQ has been proposed to be more of an academic ability measure than a true measure of cognitive abilities and thus ignore practical intelligence (Wagner & Sternberg, 1986). Another argument is that the measures are constructed in a way that favors abilities that Western society view as important, so-called cultural bias (Gould, 1996, p. 344; Greenfield, 1997). However, a recent study showed that the g-factor seems to be universal and cross-cultural (Warne & Burningham, 2019). Moreover, several studies have shown that IQ (partly) can be predicted using global structural brain differences as well as changes of structures and differences within structures (e.g., Price et al., 2013; Shaw et al., 2006). Another strength of the g factor is that tasks that are very different from standardized test batteries can measure g if they share variance with other cognitive tests (c.f., Cucina & Howardson, 2017; Spearman, 1927). As a summary, although historically criticized, the g factor is in a contemporary context supported by biological associations as well as cross-cultural links and positive correlations to various cognitive tasks.

In 1941 Catell proposed that the g factor could be split into two parts, fluid intelligence (gf) and crystallized intelligence (gc) (Cattell, 1963). Fluid intelligence (gf) refers to the ability to reason, recognize patterns and solve problems that is unrelated to earlier experiences or knowledge (Cochrane et al., 2019). Crystallized intelligence (gc), on the other hand, refers to obtained knowledge such as language, information and different concepts (McGrew, 2009). Although there are differences in what abilities in specific tests are called, many modern IQ tests contain parts aimed to measure both fluid and crystallized intelligence (Benson et. al., 2010; Kanne, 2011). Some tests are, however, more narrow: one such example is Raven's Progressive Matrices that is a non-verbal test more aimed towards fluid intelligence (Bilker et al., 2012; McLeod, 2021). In the present thesis Performance IQ (PIQ) is used as a measure of fluid intelligence, which is common practise when utilizing Wechler’s intelligence test batteries (e.g., Fitzpatrick et al., 2020; Griffiths et al., 2000).

Intelligence As a Predictor of Various Life Outcomes

Originally measurements of intelligence were developed to predict educational success (cf. Binet & Simon, 1904; Mackintosh, 2011). It is therefore, perhaps, not surprising that a meta-analysis with a sample of a total of 105,185 individuals showed a positive population correlation between school grades and IQ ($\rho = .54$) (Roth, et al., 2015). Compared to
exclusively nonverbal tests, an even stronger relationship between grades was evident with verbal and mixed intelligence tests (Roth et al. 2015). Based on the results, Roth (2015) and colleagues concluded that independent of sex, tests containing both verbal and nonverbal intelligence tests are the best predictor for school grades. Even though measurements of intelligence were intended to predict outcomes relevant to education, individual differences in IQ have shown to correlate with a wide array of outcome variables, in different cultural settings and time epochs (cf. Anger & Heineck, 2010; Bergman, et al., 2014; Glaser, et al., 2011; Sörberg, et al., 2014; Templer, et al., 2007). Low IQ in childhood has been linked to risk of several health-related pathologies such as schizophrenia as well as severe variants of depression and anxiety disorders (Koenen, et al., 2009). On the other end of the IQ spectrum, high IQ in early adulthood has been associated with a higher risk of Parkinson’s disease (Fardell, et al., 2020). Considering the substantial evidence that IQ predicts a wide variety of life outcome variables, finding neural correlates of intelligence may be of importance in various fields such as educational science, cognitive neuroscience, and the medical sciences.

**Brain Anatomy, Lesions, Features of Specific Areas and Cognitive Abilities**

The idea of the brain as an anatomical structure related to behavior dates back to at least the 1600s, as the French philosopher Descartes suggested that the brain controls behavior (Kolb & Whishaw, 2017). Examining links between brain anatomy and intelligence has been prevalent for many years. As an example, collecting, mapping, and analyzing brains from prominent Russians was a part of Russian scientific efforts already in the 1920s (Vein & Maat-Schieman, 2008). Similarly in the western world, Einstein’s brain was both examined and photographed postmortem (Falk et al., 2013). An early suggestion, dating at least back to Francis Galton in the late 19th century, was the positive correlation between brain volume and intelligence (Galton, 1889). Albeit weakly correlated, the link between brain volume and intelligence is nowadays well known and has been so for some time, (Andreasen, et al., 1993; Thoma, et al., 2005; Wickett, et al., 2000; Nave, et al., 2019). Overall, not only has interest in links between brain anatomy and intelligence been longstanding; some of the early assumptions nowadays have strong scientific support but given the dynamic development of imaging techniques new insights have also emerged.

Nowadays, a number of non-invasive neuroimaging techniques exist (Kim et al., 2021). For many years, structural magnetic resonance imaging (sMRI) has been a common method to obtain cerebral structural data, such as volumetrics, cortical surface area, distinguishing white matter from grey matter and various other properties (cf. Symms, et al., 2004; Viard, et al., 2021). Another non-invasive neuroimaging technique commonly used is functional magnetic resonance imaging (fMRI; Kim et al., 2021). While different types of fMRI exist, the most common form is the blood-oxygen-level dependent (BOLD) contrast (Kim et al., 2021; Specht, 2020). The BOLD-contrast is a representation of differences in oxygenation of the blood, how it flows and blood volume of the brain (Logothetis & Pfeuffer, 2004). For example, contrasting functional brain activity when a specific task is performed and when a control task is performed results in a BOLD signal used for subsequent statistical analysis (Gore et al., 2003; Kim et al., 2021). More intricate contrasts and complex designs with both tasks of different difficulties and control tasks with different groups are possible (as an example of relevance to intelligence research, see Dunst et al., 2014).

Variations of regional grey matter thickness in healthy adults has been shown to correlate with intelligence as early as 2006 (Narr et al., 2006). Narr and colleagues (2006) points out that prior to their study, only the study by Shaw et al., (2006) had investigated the relation between IQ and cortical thickness. Shaw et al., (2006) concluded that IQ and cortical thickness in cerebral cortex correlates negatively in younger children (3.8 to 8.4 years), but
positively in older subjects (8.6 to 29 years). More recent studies (Bajaj et al., 2018; Schnack, et al., 2015) have provided evidence that: (1) cortical thickness in different parts of the human brain are linked to IQ, and (2) a faster rate of thinning of cortical thickness in childhood is associated with higher intelligence. In adults, thicker cortical grey matter has been linked to higher IQ (Bajaj, et al., 2018). Bajaj and colleagues (2018) found several areas in where cortical thickness relates to Full Scale IQ (FSIQ) mainly, but not exclusively in the fronto-parietal regions. However, Bajaj et al. (2018) also highlighted their findings in relation to the findings in the study by Schnack et al (2015) which only found that thinner cortex in the left hemisphere in children are related to higher IQ (as will be described in a paragraph below). Summarizing the studies above, regional cortical thickness is related to IQ, but in different ways depending on age groups and birth weight.

It should be noted that there is no consensus regarding what or which processes that are causing the cortices to become thinner or to be assessed as thinner by the image processing itself (Fjell, et al., 2015; Natu, et al., 2019). Even though Natu et. al. (2019) suggest it is due to myelinization that affects image interpretation (the grey and white matter borders are depending on degree of myelination) they also point out that such a change is not excluding other microstructural changes such as pruning (elimination of extra synapses). In the present thesis, cortical thickness is therefore defined as the thickness as measured by the software used, regardless if the cortex is objectively of a certain thickness due to processes such as myelinisation.

The Lingual Gyri, Related Brain Structures and Cognitive Functions

The lingual gyri are bilateral anatomical structures within the medial occipitotemporal gyri (e.g., Chau, Stewart, & Gragnaniello, 2014; see Fig. S1 & Fig. S2a in Supplementary material for a visualization of the specific localization of the lingual gyri). The lingual gyrus is a posterior continuation of the parahippocampal gyrus (Donkelaar, et al., 2018). Anatomically, the medial occipitotemporal gyrus (of which the lingual gyri is a part) is said to be an “extremely interconnected system, supporting its ability to perform coordinated basic visual processing, but also serves as a center for many long-range association fibers, supporting its importance in nonvisual functions, such as language and memory” (Palejwala et al., 2021). The researchers highlights that the white matter of the medial occipitotemporal gyrus projects both to itself and adjacent cortices such as the temporal and the frontal lobes. It could therefore be argued that this area might be important for several higher cognitive functions implicated in intelligence.

Divergent thinking can be defined as the potential to think creatively and/or solve problems (Acar, et al., 2020). Divergent thinking has been negatively correlated to the cortical thickness of the lingual gyri (Jung, et al., 2010). Original verbal associations have also been shown to bilaterally activate the lingual gyri (Benedek et al., 2020). Additionally, forming metaphors showed higher activation in the left lingual gyrus compared to when the subjects were performing a control task (Benedek, et al., 2014). In a general activation likelihood estimation meta-analysis, it was noted that the right lingual gyrus showed activation during verbal creativity tasks (Boccia et al., 2015). Moreover, several types of recognition have been linked to the lingual gyri. For example, lesions in the lingual gyrus have been shown to negatively affect recognition of objects such as landscapes and buildings (Takahashi & Kawamura, 2002). Biological motion perception refers to the ability to perceive and understand the movements performed by others (Johansson, 1973; Fox & McDaniel, 1982). An fMRI study found that recognition of biological motion was positively linked to activity within a region of the lingual gyrus (Servos, et al., 2002). More recently, a study found that the lingual gyri were involved in recognition of objects that were gradually revealed from visual noise.
Thus, visual recognition of different types is linked to the lingual gyri. When it comes to mathematics and activity in the lingual gyri, an fMRI study found bilateral activation in the lingual gyri during arithmetic calculations in adults as compared to a resting condition without stimuli processing (Rickard et al., 2000). Another study showed activation in the lingual gyri of adults during arithmetic operations, and more so during multiplication than during addition (Rosenberg-Lee et al., 2011). When it comes to younger children (range 7.1 to 9.4 years) the study by Davis et al., (2009) found higher activation in both left and right lingual gyrus during exact calculations compared to performing a control task (Davis et al., 2009). However, the activation in the study was lower for children than for adults.

Summarizing the paragraph above, it has been shown that several dimensions of cognitive performance such as divergent thinking, forming metaphors, recognition and specific mathematical abilities are linked to recruitment of the lingual gyri. In the current thesis it is hypothesized that variability in lingual gyri thickness is related to variability in intelligence.

Short Overview of The Current Thesis

The current thesis uses imaging data from a database (see methods section for details). Two different datasets will be used based on some predefined selection criteria related to the aim of the thesis (see section “Dataset Description, Data Acquisition, and Data Handling” for details). In the first study by Solé-Casals et al., (2019) participants (n=29) were scanned once, at an age of approximately twelve years. They were divided into two IQ groups (High IQ, n = 15, Lower IQ, n =14) which also will be used in the current thesis but with a different research question. In the second study, by Suárez-Pellicioni et al. (2019), they had a longitudinal design (n=63), spaced approximately two years apart, which enables detection of changes in cortical thickness in the same individuals over time. In the context of the present thesis, the main focus is solely related to the change in cortical thickness in lingual gyri and PIQ in children (n= 36) that participated both at the first (T1) and second wave (T2) and fulfilled the criteria for the current thesis (see Study 2 “Participants” for details). Hereafter, the study by Solé-Casals et al. (2019) will be denoted as Study I, and the study by Suárez-Pellicioni et al. (2019) will be denoted as Study 2. Below are the aims and expected predictions from each of the datasets.

Aim and Predictions

As shown above, cognitive abilities that are associated with intelligence have, in different ways, been associated with both the left and right lingual gyrus. To the best of my knowledge, no research or previously published articles have investigated the association between intelligence and the lingual gyri on a structural level. The aim of this exploratory study is to examine if such an association exists when it comes to cortical thickness. It is well-established that creativity, divergent thinking, recognition, and mathematical abilities are positively correlated with intelligence. As described in previous sections, it is also well-established that all those abilities have associations with the lingual gyri when it comes to functional brain activity (i.e., fMRI). If the rate of cortical thinning in the lingual gyri is related to higher intelligence, then IQ differences of groups should also reflect differences in cortical thickness of the groups (given that the process has been ongoing long and fast enough in time). If the process is timed in a similar manner as mentioned by Schnack et al., (2015) the cortical thickness of the lingual gyri is expected to differ depending on IQ from approximately an age of 10. As the sample in the study 1 of the present thesis had a mean age of approximately 12 years the prediction for that dataset is as follows:
• Prediction 1: Children with higher IQ will have thinner cortices in the lingual gyri compared to the group of children with lower IQ.

If the rate of change is linked in a similar manner for the lingual gyri as for the left hemisphere in Schnack et al. (2015) then it follows that the rate of change would also be linked to change in PIQ. Since the dataset in study 2 is longitudinal a prediction of rate of change is possible. The longitudinal design was spaced approximately two years apart whereas the first occasion is labeled as T1, and the second as T2. Rate of change in cortical thickness in the context of this thesis is defined as the result of the following calculation: thickness at T2 subtracted by thickness at T1. The resulting thickness difference divided by the difference of age at T2 subtracted by age at T1 (i.e., change in cortical thickness divided by change in time equals rate of change). Change in PIQ is defined as the PIQ at T2 subtracted by the PIQ at T1. Furthermore, as some suggestions have been made that it might be differences in how IQ is expressed structurally in the female and the male brain (Haier, et al., 2005), the following prediction includes controlling for sex.

• Prediction 2: If grey matter volume and sex is controlled for, children with faster cortical thinning will show greater change in PIQ (i.e., rate of change in cortical thickness is negatively correlated to higher IQ, in other words rate of thinning is positively correlated to higher IQ).

Materials and Methods

As the current thesis uses two independent samples from two different research projects, this section will first introduce the common features for the datasets and how cortical thickness was assessed followed by a description of each study separately (i.e., Study 1 and Study 2). Finally, ethical consideration for both studies will be presented.

Dataset Description, Data Acquisition, and Data Handling

To be able to test the hypothesis presented above, the current thesis used already collected data available at an open platform of imaging data. The data was found via an extensive search in the OpenNeuro library (openneuro.org), which is an archive that contains brains from various neuroimaging research projects (for specific details about OpenNeuro, see Markiewicz, et al., 2021). Datasets from OpenNeuro have recently been utilized in peer reviewed studies (e.g., Joo, et al., 2021; Yeung, et al., 2021). The aim of the extensive search was to obtain data relevant to cortical thinning and IQ. Specifically, data sets with T1-weighted images of children below the age of 14 and with measured IQ-scores, preferable with data from possible subtests of intelligence (for the rationale behind the age of 14, see under the section “Participants study 2”). A suitable dataset was identified with accession number ds002726. Another suitable dataset was identified with accession number ds001486.

The dataset ds002726 (Solé-Casals et al., 2019) contains T1-weighted neuroimages of twenty-nine subjects. All images were obtained using a 3-Tesla MRI scanner. MPRAGE 3D protocol was utilized with repetition time (TR) set to 2300 ms, echo time (TE) set to 3 ms, inversion time (TI) 900 ms, with 1 mm × 1mm × 1mm voxel size, and Field of View set to 244 × 244 mm². More details of the acquisition of the data can be found in the original study (Solé-Casals et al., 2019). Henceforth the study in this paragraph will be entitled ‘Study 1’.

The longitudinal dataset ds001486 (Suárez-Pellicioni et al., 2019) contains T1-weighted neuroimages of 132 subjects at time point 1 as well as 63 of the initial subjects at time point 2. All images were obtained using a 3-Tesla MRI scanner. MPRAGE 3D protocol
was utilized with repetition time (TR) set to 2300 ms, echo time (TE) set to 3.36 ms, with 1 mm × 1mm × 1mm voxel size, matrix size 256 × 256 m. and field of view set to 240 mm. More details about the acquisition of the images and the dataset itself can be found in the original study (Suárez-Pellicioni et al., 2019). Henceforth the study in this paragraph will be entitled ‘Study 2’.

The datasets were downloaded from the OpenNeuro website. After the data was acquired and unpacked, the brains were visually quality checked in MRicroGL (https://www.nitrc.org/projects/microgl). The MRicroGL software has been exemplified as a viable tool for quality checking structural MRI images (Backhausen, et al., 2016) and has been previously used for visualization of structural MRI data (Bouton, et al., 2016). In the downloaded dataset each image was checked for artefacts in the sagittal plane as well as the coronal plane. In study 1, no significant artefacts were present that were believed to compromise further data analysis. The quality check did therefore not render in any exclusions of participants in study 1. However, in study 2 a T1-weighted image of a participant (subject 099, second scan) was not correctly labelled, dimensionally and appeared ‘upside down’ in the sagittal plane as well as in the coronal plane. Due to the nature of 3d images, it was not possible to distinguish if the image was inverted and/or rotated (see further handling under the paragraph ‘Measurements of Cortical Thickness’). Another subject had clear and obvious motion artefacts from the first scan (subject 024). Both subjects were excluded. A third subject (male) was excluded in a later stage (identified as an outlier in study 2).

Measurements of Cortical Thickness

All brain images were analyzed using the vol2Brain pipeline (Manjón, et al., 2022). Different versions of the pipeline have been utilized in several previous studies (e.g., Coupé, et al., 2019; Zamani, et al., 2022). Manjón et. al., (2022) describes the following steps in the preprocessing of the images: (1) denoising the imaging using Spatially Adaptive Non-Local Means, (2) correcting inhomogeneities, (3) registering the images to the Montreal Neurological Institute (MNI) space, (4) letting SPM8 based toolbox correct inhomogeneities, and, finally, (5) the images intensity normalized so the pictures that are to be segmented share geometrical and intensity space.

After preprocessing, and in line with Manjón et. al., (2022) the vol2Brain pipeline continues with multiscale labeling and cortical thickness estimation in the following order: (1) intracranial cavity extraction (ICC), (2) full brain structure segmentation in which neuroanatomical structures of the brain are identified (and labelled), (3) systematic errors are corrected, (4) labels are combined to create a tissue-type segmentation map, and, finally, (5) cortical thickness in different areas of the cortex is estimated by computing the distance between the inner and the outer parts of the cortical grey matter. When the pipeline has completed a brain, it produces an output of several files. Of relevance to the present thesis, a .csv file with estimated volumes, thicknesses and asymmetry ratios is then downloadable from the website as well as a report file with visual segmentation and various estimates also included in the .csv file. Regarding the inverted and/or rotated neuroimage for one subject, previously mentioned, it was incorrectly segmented. It was therefore excluded from further data analysis.

Study 1 (Solé-Casals et al., 2019)

Participants

Twenty-nine righthanded boys (Solé-Casals, et al., 2019) divided into two groups with different IQ scores were part of the study. All subjects were scanned during approximately the
age of twelve (control group age = 12.53, SD = 0.77, High IQ group age = 12.03, SD = 0.54). The group with higher IQ consisted of fifteen boys with High IQ (IQ = 148.80, SD = 2.93), hereafter denoted as the High IQ group. The control group consisted of 14 boys with somewhat lower IQ than the High IQ group (IQ = 122.71, SD = 3.89), hereafter denoted as the control group. The original researchers of the study were contacted with the purpose to obtain individual data from the different parts of the IQ tests. However, none of them did respond.

**Instruments**

**Cognitive Abilities**

IQ of all participants (N=29) was measured using a Spanish cognitive test battery (Santamaría, et al., 2014; Santamaría & Pereña, n.d.). The battery evaluates the following abilities: (1) spatial, (2) numerical, (3) abstract reasoning and (4) verbal. Three global dimensions of the individual are then computed as a composite score: (1) General intelligence, (2) Non-verbal intelligence and (3) Verbal intelligence. In the present thesis, the composite General intelligence is being utilized. The composite score is combining spatial, numerical, abstract reasoning and verbal abilities. The validity of this measure was found to be 0.75, in terms of the median correlation to other tests measuring general intelligence (Consejo General de Colegios Oficiales de Psicologos, n.d., p. 11) As a comparison, the non-verbal Raven’s progressive matrices and WAIS full scale IQ was correlated moderately (0.69) in a Dutch twin study (Rijsdijk, et al., 2002).

As mentioned in the Participants section, the dataset consists of two predefined groups with different IQ scores. The High IQ group, consisting of 15 boys (IQ = 148.80 ± 2.93) and the control group of 14 boys (IQ = 122.71 ± 3.89). Additional criteria for the High IQ group were performances above the 90th percentile in the following abilities: “spatial, numerical, abstract reasoning, verbal reasoning and memory” (Solé-Casals, et al., 2019, p. 2374).

**Statistical Analysis**

All data from the vol2Brain pipeline were downloaded in .csv format. The data of each brain comes in a separate file. A power query was therefore conducted in Excel (Microsoft Corporation) to create a coherent file including all brains. The resulting Excel file was then imported to IBM SPSS Statistics (Version 28). All statistical analyses were then performed in SPSS. Outlier detection (when it came to cortical thickness of the gyri, separately as well as individually) was conducted in line with recommendations (Tabachnick & Fidell, 2013) by using a z-score of ±3.29 as threshold for potentially excluding participants.

Due to the relatively small sample size, independent samples t-tests were conducted. The test was originally developed to be able to handle small sample sizes (c.f. Student, 1908; Zabell, 2008). Furthermore, t-tests are common in the field of cognitive neuroscience (Szucs & Ioannidis, 2017). The authors report that a search in 18 journals from January 2011 to August 2014 rendered over 3800 papers. First a t-test was conducted regarding total bilateral cortical thickness of the lingual gyri to examine whether the High IQ group had thinner cortices than the control group. Schnack et al., (2015) concludes that the rate of cortical thinning in the left hemisphere is linked to higher IQ. It could be argued that a similar pattern in the lingual gyri is driven by changes in the left hemisphere in which the left lingual gyrus is part, and that changes within the lingual gyri is therefore not directly related to intelligence, but to the general change of the left hemisphere. Therefore, two independent t-tests were conducted regarding cortical thickness of the left and the right gyrus respectively in order to reject a left-hand side driven mean total thickness (i.e., that the possible group differences is exclusively or almost
exclusively in the left lingual gyrus). For the outcome of those analysis, see Supplementary material under ‘Control Analysis study 1’.

**Procedure**

In the original peer reviewed article Solé-Casals et al., (2019) does not clarify how the FSIQ scores were obtained, except what has been described in the paragraph ‘Instruments of study 1’. The original article does not contain a ‘Procedure’ section. For details regarding structural MRI, see section ‘Dataset description, data acquisition, and data handling’ above.

**Study 2** (Suárez-Pellicioni et al., 2019).

**Participants**

Of the participants in the original data collection (70 right-handed females, 62 right-handed males) 63 of the subjects were followed up and scanned a second time approximately two years later (35 females, 28 males) (Suárez-Pellicioni et al., 2019). According to Schnack et. al., (2015) the relation between thickness change in the left hemisphere and Performance IQ is significant for ages under 14.

To account for the restraints mentioned above the following criteria was used for inclusion in the data analysis of prediction 2: (1) under 14 years of age at T2 with (2) a calculated Performance IQ score at both T1 and T2. After selecting according to the criteria, twenty-one females (age at T1 = 9.97, SD = 0.70, age at T2 = 12.20, SD = 0.67 and fifteen males (age at T1 = 10.33, SD = 0.95, age at T2 = 12.64, SD = 0.87) remained in the dataset.

**Instruments**

_Cognitive Abilities_

Full scale IQ (FSIQ), Verbal IQ (VIQ) and Performance IQ (PIQ) of the participants (n=132) was estimated using Wechsler Abbreviated Scale of Intelligence (WASI) (Suárez-Pellicioni, et al., 2019). One of the stated purposes of WASI is to enable a fast estimate of general intellectual functioning for research (Pearson, 2022). WASI has been previously used in a wide variety of studies to obtain intellectual functioning in diverse categories of samples, such as schizophrenics, psychopaths and people with brain injuries (eg., Hope, et al., 2015; Sharratt, et al., 2020; Heitger, et al., 2004) and, specifically in children and adolescents (Golding, et al., 2021; Willcock, et al., 2011).

WASI includes four subtests: (1) Matrix Reasoning, (2) Similarities, (3) Block Design and (4) Vocabulary. Pearson (2022) states that WASI I and WASI II has the same structure and format. In WASI, three dimensions of the individual are computed as a composite score: (1) Full Scale Intelligence Quotient (FSIQ), (2) Performance Intelligence Quotient (PIQ) and (3) Verbal intelligence Quotient (VIQ) (Wechsler, 1999). In the data from study 2, the composite PIQ is utilized. PIQ is calculated using the Block Design and Matrix Reasoning, excluding the Vocabulary and Similarities tests (i.e., excluding the more crystallized parts of the FSIQ and thus estimating fluid intelligence.).

**Statistical analysis**

All data from the vol2Brain pipeline was downloaded in .csv format. The data of each brain comes in a separate file. A power query was therefore conducted in Excel (Microsoft Corporation) to create a coherent file including all brains. The resulting Excel file was then
imported to IBM SPSS Statistics (Version 28). All statistical analyses were then performed in SPSS.

Outlier detection was conducted (when it came to cortical thickness of the gyri, the rate of change in cortical thickness) using a z-score of ±3.29 as threshold for potentially excluding participants. The threshold is recommended in relevant literature (Tabachnick & Fidell, 2013). No outliers were detected using this method.

The association between the rate of change in cortical thickness in lingual gyri and change in PIQ was examined using partial (r') correlation, controlling for the effects of grey matter volume. Controlling for grey matter volume is not uncommon in fMRI and MRI studies of different sorts (cf. Jamadar, et al., 2013; Veldsman, et al., 2020; Jamadar, 2020; Lim, et al., 1999). For a more geometrical and mathematical discussion about the relationships between cortical thickness, grey matter volume and surface area, see Winkler and colleagues (Winkler, et al., 2010). Head size and total volume has not been controlled for, due to recommendations against adjusting for those variables in cortical thickness models (Barnes, et al., 2010; Westman, et al., 2013; Schwarz, et al., 2016).

Since normally distributed variables is a basic assumption for correlations (Haylcek & Peterson, 1977), all variables were controlled for non-normal distributions. Including grey matter volume. Due to small sample size for female subjects (n = 21) and for male subjects (n = 15) Shapiro-Wilk test was utilized (for rationale for the choice of Shapiro-Wilk in small samples, see Keskin, 2006; Yap & Sim, 2011). It should be noted that according to several authors correlations are insensitive to violation of normal distributed data (Haylcek & Peterson, 1977; Knief & Forstmeier, 2021). All variables indicated that the variables were normally distributed. See Supplementary material for tests of normality.

Procedure

At first (T1) the parents of the subjects were handed questionnaires, related to development. Parts of the questionnaires were used to exclude participants. For example, and of relevance for the present thesis, subjects were excluded if they (1) had Attention Deficit Hyperactivity Disorder (ADHD), (2) neurological disease, (3) were left-handed or (4) prematurely born. For more additional excluding criteria and details, see Suárez-Pellicioni et al., 2019. Thereafter the subjects completed various standardized tests at the lab, including the WASI test described in detail under the paragraph ‘Cognitive Abilities’. For a more detailed description of tests, see Suárez-Pellicioni et al., (2019). After the standardized tests the subjects were scanned (MRI as well as fMRI protocols were utilized) in a Siemens 3T TIM Trio MRI scanner. The subjects completed four tasks, unrelated to the present thesis inside the scanner (including two arithmetic tasks and two localizertasks).

Around two years (T2) later all subjects were asked to redo testing and to be rescanned. Suárez-Pellicioni et al., (2019) mentions that some subjects were not available for longitudinal assessment due to relocating to a different area, obstructs to new MRI scans and “lack of interest” (p. 3). At T2 some, but not all of the standardized tests were repeated and new structural and fMRI data was obtained. All of the tasks relevant to the present thesis were included in the behavioral assessment of T2.

Ethical Considerations

OpenNeuro has a policy of defacing subjects before uploading to their repository (OpenNeuro, 2022). Defacing is a process in which the faces of the subjects are removed to hinder identification of the subjects (Theyers, et al., 2021). Hence, the subjects are anonymous to the author of the thesis. It should also be noted that study 1 that led to the OpenNeuro dataset
ds002726 was ethically approved by the research ethics committee, the Institutional Review Board of the University of Barcelona (Solé-Casals, et al., 2019) and that study 2 that led to the OpenNeuro dataset ds001486 was approved by the Institutional Review Board at Northwestern University. Furthermore, guardians and parents gave consent to the original data collection and sharing of data in which the children was going to be deidentified (Suárez-Pelllicioni, et al., 2019).
Results

The results from each study will be presented separately in the same order as in the Material and Methods section. To partly foreshadow the results, despite two independent samples the lingual gyri was found to be significantly thinner in the High IQ group compared to the control group. Change-change correlations were present in the participants independent of sex. Specifically, a faster thinning of the lingual gyri was correlated with a positive change in PIQ.

Study 1

In order to test if the High IQ group had thinner cortices in the lingual gyri than the control group an independent t-test was conducted. This test was found to be statistically significant, \( t(27) = -3.02, p = .003; d = 0.21 \). The effect size exceeds Cohen’s threshold for a small effect (Cohen, 1988, 1992). The results indicate that the lingual gyrus is bilaterally thinner in the High IQ group (\( M = 2.95, SD = 0.16 \)) in comparison to the control group (\( M = 3.12, SD = 0.26 \)). As can be seen in Fig. 1, cortical thickness of the lingual gyri in the High IQ group ranged from 2.59 to 3.23. Cortical thickness of the lingual gyrus in the control group ranged from 2.71 to 3.53 (see fig. 1 below). No outliers were detected. For boxplot of cortical thickness, see Figure 1. Control analysis was conducted to rule out a left side driven mean difference in cortical thickness. The analysis indicated that the differences in cortical thickness was bilateral (for statistical results and Boxplots of the control analysis, see Supplementary material ‘Control analysis 1’).

Figure 1

Boxplot of Total cortical Thickness of the Lingual Gyri in mm.

Note. Bilateral cortical thickness of the lingual gyri related to IQ. The boxplot was created in SPSS. End of the bars denote maximum and minimum values. Top of box represents third quartile; bottom represents First quartile and, Median is represented by the black line within each box.
Study 2

To test if the rate of cortical thinning in children was related to change in PIQ between T1 and T2, a partial correlation was performed. When both grey matter volume and sex were controlled for in the relationship between rate of change in cortical thickness and change in PIQ, the following significant negative partial correlation was found $r = -.50$, $p = .001$ (1-tailed, see Fig. 2). This result suggest that a faster cortical thinning is correlated to a positive change in fluid intelligence in children. Control analysis was conducted to rule out that the correlation was driven by a specific sex. The analysis indicated that both sexes had virtually indentical correlation in terms of strength and direction (for statistical results and visual representations of the control analysis, see Supplementary material, ‘Control analysis study 2’).

Figure 2

*Scatter dot of correlation between rate of change in cortical thickness and Performance IQ when Grey Matter is controlled for, in children.*

*Note.* The Partial Correlation is plotted by correlating the residuals of (1) the correlation between Change in Performance IQ and Grey Matter Volume and (2) the correlation between rate of cortical thinning and Grey Matter Volume. The residuals utilized to generate the plot are unstandardized. Also, notice: sex is not controlled for in the scatter dot as in the partial correlation.
Discussion

Main Findings

In the present thesis structural magnetic imaging data was used to investigate the association between cortical thickness in the lingual gyri and general intelligence in childhood. First, using the data from study 1 a t-test was conducted to compare the mean cortical thickness of the lingual gyri in the High IQ group and the control group. The groups had statistically significant differences in cortical thickness. By using partial correlation an association between the rate of cortical thinning in the lingual gyri and PIQ was found. Grey matter volume as well as sex was controlled for in the partial correlation, demonstrating that the association between the rate of cortical thinning and PIQ persist when the grey matter volume and the sex variables are held constant. Thus, the hypotheses of the thesis have been confirmed.

To the best of the author’s knowledge no previous studies have investigated the association between intelligence and structural properties of the lingual gyri. However, as mentioned in the introduction, several studies have linked cortical thickness (both thinner and thicker) to intelligence. The present thesis therefore (1) adds evidence that cortical thickness and change in cortical thickness in bilateral lingual gyri has associations to intelligence.

A speculative interpretation of the results of this thesis in relation to the other studies linking abilities - such as math, divergent thinking and forming metaphors - to the lingual gyri is that the lingual gyri serve as an integration hub for both visual and verbal stimuli. Thus, assisting in sense making of patterns, whether mathematical or linguistic. Somewhat in line with a study of 10 adults (Ghosh, et al., 2010, s. 185) in which the authors pointed out that “the lingual gyrus activity may be linked to content-specific retrieval 'orientation' relevant to recall, association, and integration of word-related information”.

An unexpected finding in Study 1 was that the standard deviation was nearly twice as large in the control group compared to the High IQ group when it came to the thickness of the lingual gyri. A speculative interpretation is that there is an optimal range of thickness in a specific age to obtain higher IQ. However, different interpretations are possible. It is, for example, conceivable that some parts of the control group have different spatial restrictions in their brains compared to the High IQ group: if rate of change is the most important feature to obtain high IQ, then a group with lower IQ are less restricted in two extreme ends: Firstly, they can have thick cortices that does not change and remain thick. Secondly, they might have a thin cortex to begin with, that does not have enough space for a pronounced cortical thinning. Thus, in that line of thought greater variance is expected in group with lower IQ than a group with higher IQ.

The statistically significant difference in the right lingual gyrus that is part of the control analysis is of specific interest. For the left lingual gyrus, it can be argued that the lateralized difference might be an artifact of the already known differences in the left hemisphere (cf. Schnack, et.al., 2015; Solé-Casals, et. al., 2019). Thus, the difference in the right lingual gyrus can be interpreted as an indication of a causal relationship between the structural properties of the lingual gyri and intelligence. It should be noted, however, that a possible causal relationship is not further examined within the context of the present thesis.

The results might have implications within several areas. Regarding educational psychology, the results suggests that structural changes in the brain might explain more of the variance than previously thought when it comes to cognitive performance. Thus, lack of progress in some children might have more to do with brain development than absence of motivation, inadequate teaching methods or other factors. On the other hand, fast progress in some children might have more to do with brain development than motivation, teaching
methods or other factors. It could, for example, be argued that optimal education needs to be adapted to individuals brain development rather than other factors such as age.

When it comes to clinical psychology, although speculative the results might nuance the view of mild cognitive impairment/disability. If cortical thickness and rate of change of cortical thickness in the lingual gyri (and elsewhere) is normally distributed, then it could be that a substantial portions of mild cognitive impairment in children has more to do with normal distribution of those variables than brain pathology. This interpretation is in line with fairly recent research on the genetics of intellectual disabilities in which most cases of mild cognitive disability seem to correspond to the lower end of the normal distribution (Reichenberg, et al., 2016). A similar discussion has for quite some time been around regarding dyslexia (cf. Shaywitz, et al., 1992; Wagner, et al., 2020). Thus, this thesis might add to more precise distinctions between quantitative and pathological/qualitative cognitive disabilities (for a thorough discussion on how disorders in several cases might be quantitative traits rather than qualitative conditions, see Plomin, et al., 2009).

Regarding study 2, it is not possible to draw causal conclusions from the change-change correlation of Performance IQ and cortical thinning (i.e., does changes in Performance IQ cause the cortices of the lingual gyri to get thinner or does the Performance IQ go up because of the cortical thinning). It is possible that a third factor is at play, that causes both the cortices in the lingual gyri to become thinner and the performance IQ to go up. However, it could be argued that the change-change correlation (at group level) points towards biological and structural foundations of Performance IQ (at group level). This is somewhat in line with previous studies of how structural changes in the brain is related to change in cognitive performance (Ramsden, et al., 2011). Ramsden and colleagues (2011) found that changes in grey matter density in cerebellum (anteriorly) correlated with changes in Performance IQ. It is also in line with the previously mentioned study by Schnack et. al., (2015) in which cortical thinning in the left hemisphere is correlated with higher IQ.

Although speculative, the results can be interpreted as further evidence that intelligence is associated to genetic factors. Previous studies point out that cortical thickness as well as surface area of the cortex have moderate to strong associations to genetic factors (Schmitt, et al., 2008; Strike, et al., 2019). If cortical thickness has a strong to moderate genetic factor and intelligence is correlated to cortical thickness, then it is conceivable that cortical thickness mediates intelligence (at least to some extent). In line with this interpretation, studies in behavioral genetics and studies focused on DNA have shown that genetic factors have strong correlations to intelligence (Deary, et al., 2022).

**Methodological Issues**

For quite some time, criticism of certain aspects of neuroimaging studies have surfaced (Poldrack, et al., 2017). One such example is power calculations. For example, less than 7 percent of the 1038 most cited structural and functional magnetic resonance imaging studies in 2017 and 2018 had power calculations (Szucs & Ioannidis, 2020).

Small sample sizes are generally viewed as a limitation in both psychology and cognitive neuroscience (Szucs, et al., 2017; Poldrack, et al., 2017). For a broader discussion regarding the problems of small sample sizes in neuroscience, see Button et al., (2013). In the context of the present thesis the small sample sizes could, potentially be problematic in several ways. One such problem could be unbalanced initial conditions when it comes to cortical thickness in the lingual gyri. If normality in the population is assumed when it comes to cortical thickness, a large sample size would likely lead to group differences in thickness between groups with higher IQ than lower IQ if the cortical thinning is indeed faster in the High IQ groups. However, with smaller sample sizes the risk of random effects is greater (i.e., a few
individuals could potentially skew the data in such a way that the group differences are non-representative). Therefore, the small sample size is an important limitation in study 1.

Even though small sample size is important in general (see above), it is somewhat less problematic in studies with longitudinal design. The statistical power is generally higher than in cross-sectional studies since the design is able to pick up within-person changes at different points in time (Guo, et al., 2013; Goulet & Cousineau, 2019; Schober & Vetter, 2018). It can be concluded that sample size is a limitation in both studies, but less so in study 2 due to the longitudinality.

A limitation of study 1 is the lack of individual scores for the different tests. Such information would enable more intricate statistical analysis such as controlling for more variables and/or focusing on specific subtests or composite scores, such as PIQ.

A limitation in study 2 is the within-group variation of age at T1 (range: 8.47 – 11.58, SD = 0.70) as well as the within-group variation at T2 (range: 10.91 – 13.63, SD = 0.67). However, as mentioned under the paragraph ‘Participants study 2’ the measurements is within age limits in which the global rate of change in cortical thickness is significantly associated with performance IQ. Still, a more rigorous age range at T1 and T2 would be preferable. Especially so because of the non-linear relationship between age and rate of cortical thinning in both Schnack et. al. (2015) and Shaw et. al., (2006).

Methodological Strengths

There are several strengths of the present thesis. In both studies, sex is constrained to males in study 1 whereas in study 2 sex was controlled for in the partial correlation. Furthermore, control analysis was performed to exclude the possibility that a certain sex was the driving force of the correlation. Furthermore, handedness was held constant in both studies (only righthanded were tested and scanned).

Another strength was the small deviations of FSIQ in study 1. Generally, creating groups from a continuous variable can be problematic in many ways (Cohen, 1983; MacCallum, et al., 2002; Good & Hardin, 2006, pp. 28-29). In the context of study 1, a potential problem could be that some persons in the extreme ends of the distributions in each group potentially belong to the other group. Such mislabelling could be due to randomness of the instrument itself, due to a subject underperforming or due to randomly guessing correctly in multiple choice question (for more in detail descriptions about how guessing might affect results in matrices test, see Kunda, et al., 2016; Antoniou, et al., 2022). However, the means of the groups are more than a standard deviation apart (more than 15 points). Furthermore, the within group standard deviations from the mean is quite low (less than 5 points for both groups). As a result, the risk that a subject should belong to the other group is therefore reduced by the small within-group deviations and therefore can be considered a strength of study 1.

The above-mentioned strength is also improved by the additional criteria for belonging to the High IQ group, specifically being above the 90th percentile in spatial, numerical, abstract reasoning, verbal reasoning, and memory. The criteria therefore exclude uneven cognitive profiles. Such profiles may possibly reflect differences on an anatomical and/or neurological level. The criteria might therefore reduce the risk of atypical brains in the High IQ group (i.e., the MRI dataset might better reflect a typical brain at the High IQ level).

A final strength of study 1 was the high IQ of the control group which lowers the probability that a neurological disorder was behind individual IQ differences. For example, an individual could have had a much higher IQ if it wasn’t for a degenerative disorder. The individual brain would then be structurally more like the high IQ group and therefore compromise the group comparison (i.e., an individual could have a cortical thinning process that is optimal for high IQ, but neurological disorders that effect IQ negatively).
One of the strengths of study 2 was the exclusion criteria: amongst them ADHD, neurological disease, and being prematurely born. The criteria reduce the risk that structurally atypical brains and/or non typically developing brains were included in the dataset. It can be argued that it is particularly important when including participants with lower IQ scores since lower IQ is linked to higher probability of neurological disorder and other conditions that might alter the brain on a structural level (cf. Melby et al., 2020; Pyhala et al., 2017). In the original dataset of study 2 the IQ ranged from 81 to 144 at T1 and 79 to 144 at T2 (Suárez-Pellicioni, 2019, p. 3). All in all, the exclusion criteria can be considered an important strength of the study.

Conclusions and Future Directions

Future studies would benefit from larger sample sizes to establish that the cortical thickness (and the rate of change) of the lingual gyri are good biomarkers for intelligence. An alternative would be a larger number of studies with smaller sample sizes. Even though more costly, studies with measurements at additional time points would have great potential to unravel more intricate trajectories of development. A perhaps less costly solution would be to conduct a so-called Integrative Data Analysis, in which many individuals from different datasets are combined into one dataset (Curran, et al., 2009).

Even though it might prove costly, a combination of DNA sampling, longitudinal MRI and cognitive testing might provide answers to which genes affect the rate of cortical thinning (both in the lingual gyri and elsewhere in the brain) as well as to what extent those genes explains variance in intelligence.

All in all, the results in the current thesis linked cortical thickness in early adolescence to FSIQ and rate of cortical thinning to PIQ. This might have implications relevant to clinical and educational psychology as well as cognitive neuroscience. Further research with larger sample sizes would be preferable. Possible combined with biological samples of the subjects to find genetic correlates to intelligence, cortical thickness and thinning.
References


Student. (1908). The probable error of a mean. *Biometrika, 1–25*


Supplementary Material

Figure S1

The Lingual Gyri colored in red, Sagittal View.

Note. The image is an altered version of an image acquired from Wikimedia Commons (https://commons.wikimedia.org/wiki/File:Gray727.svg). The image is free to use for any purpose.

Figure S2a-c

The Left Lingual Gyrus colored in semi-transparent purple

Note. Fig. 2a: Coronal View (frontal view, i.e., subjects left lingual gyrus on right hand side for an observer). Fig. 2b: Sagittal View. Fig. 2c: Axial view. The images were created by overlaying the AAL atlas onto a SPM152 brain volume. The result is therefore an estimation of the position of left lingual gyrus. The AAL atlas is creative commons and free to use or modify. SPM152 is a standard brain template within MRICroGL.
Study 1: Control Analysis (left and right lingual gyrus and IQ)

To test if the High IQ group had thinner cortices in the left lingual gyrus than the control group an independent samples t-test was conducted. This test was found to be statistically significant, $t(27) = -2.72, p < .006; d = 0.22$. The effect size exceeds Cohen’s threshold for a small effect (Cohen, 1988, 1992). The results indicate that the left lingual gyrus is thinner in the High IQ group ($M = 2.88$, $SD = 0.20$) in comparison to the control group ($M = 3.05$, $SD = 0.25$). Cortical thickness of the left lingual gyrus in the High IQ group ranged from 2.61 to 3.23. Cortical thickness of the lingual gyrus in the control group ranged from 2.60 to 3.47. No outliers were detected, using the predetermined threshold of 3.29. For boxplot of cortical thickness of the left lingual gyrus, see Figure S3.

To test if the High group had thinner cortices in the right lingual gyrus than the control group an independent samples t-test was conducted. This test was found to be statistically significant, $t(27) = -2.95, p < .003; d = 0.23$. The effect size exceeds Cohen’s threshold for a small effect (Cohen, 1988, 1992). The results indicate that the left lingual gyrus is thinner in the High IQ group ($M = 3.02$, $SD = 0.17$) in comparison to the control group ($M = 3.20$, $SD = 0.29$). Cortical thickness of the left lingual gyrus in the High IQ group ranged from 2.57 to 3.27. Cortical thickness of the lingual gyrus in the control group ranged from 2.82 to 3.68. No outliers were detected, using the predetermined threshold of 3.29. For boxplot of cortical thickness of the right lingual gyrus, see Figure S4.

Figure S3a-b Left and Right Lingual Gyrus

Boxplot of Left and Right Lingual Gyrus Thickness in mm in both groups.

Fig. S3a  Fig. S3b

Note. Fig. S3a represents the group differences of the High IQ group and the control group, in the left lingual gyrus. Fig. S3b represents the group differences of the High IQ group and the Control group, in the right lingual gyrus. The Boxplots was created in SPSS.
Study 2: Control Analysis (sex and change in cortical thickness)

To be able to compare sexes when it comes to correlation between Performance IQ and rate of cortical thinning, partial correlations were performed for both sexes.

For females, when grey matter volume was controlled for in the relationship between cortical thickness rate and change in PIQ, the following negative partial correlation was found $r = -.53$, $p = .008$ (1-tailed). The result suggest that a faster cortical thinning is correlated to a positive change in fluid intelligence in girls. See Fig. S4a.

For males, when grey matter volume was controlled for in the relationship between cortical thickness rate and change in PIQ, the following negative partial correlation was found $r = -.50$, $p = .036$ (1-tailed). This result suggest that a faster cortical thinning is correlated to a positive change in fluid intelligence in boys. See Fig. S4b.

Figure S4a-b

The partial correlation between rate of change in cortical thickness and Performance IQ when Grey Matter is controlled for, in girls and boys respectively.

Note. Fig S4a represents a visualization of the partial correlation between rate of change in cortical thickness and Performance IQ in girls. Fig S4b represents a visualization of the partial correlation between rate of change in cortical thickness and Performance IQ in boys. The Partial Correlation is plotted by using the residuals of (1) the correlation between Change in Performance IQ and Grey Matter Volume and (2) the correlation between rate of cortical thinning and Grey Matter Volume. The residuals utilized to generate the plot are unstandardized. Both plots were created in SPSS.

Tests of Normality, Female Subjects.


Since the change of cortical thickness is dependent on cortical thickness at T1 and T2, the cortical thickness at T1 and T2 were controlled for normality. Shapiro-Wilk test was utilized. Cortical thickness in the lingual gyri at T1: $W(21) = .96, p = .51$, Cortical thickness in the lingual gyri at T2: $W(21) = .98, p = .91$.

Since the change of Performance IQ is dependent on Performance IQ of T1 and T2, the PIQ of T1 and T2 were controlled for normality. Shapiro-Wilk test was utilized. PIQ at T1: $W(21) = .95, p = .38$. PIQ at T2: $W(21) = .98, p = .91$. The Shapiro-Wilk tests showed that no variable in the dataset departed significantly from normality. Furthermore, Q-Q plots for all variables were inspected. No obvious deviations from normality were identified.
Tests of Normality, Male Subjects.

Change in Performance IQ: $W(15) = .96, p = .67$, Rate of cortical change in Lingual Gyri: $W(15) = .96, p = .67$, Total Grey Matter volume: $W(15) = .91, p = .15$.

Since the change of cortical thickness is dependent on cortical thickness at T1 and T2, the cortical thickness at T1 and T2 were controlled for normality. Shapiro-Wilk test was utilized. Cortical thickness in the lingual gyri at T1: $W(15) = .94, p = .43$, Cortical thickness in the lingual gyri at T2: $W(15) = .96, p = .77$.

Since the change of Performance IQ is dependent on Performance IQ of T1 and T2, the PIQ of T1 and T2 were controlled for normality. Shapiro-Wilk test was utilized. PIQ at T1: $W(15) = .95, p = .57$, PIQ at T2: $W(15) = .94, p = .36$. The Shapiro-Wilk tests showed that no variable in the dataset departed significantly from normality. Furthermore, Q-Q plots for all variables were inspected. No obvious deviations from normality were identified.