

Cortical Maturation and Verbal Fluency in
Childhood, Adolescence, and Young Adulthood

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

James Norby Porter

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

Monica M. Luciana, Ph.D.

November 2009

© James N. Porter 2009

Acknowledgements

This project was supported by funding awarded to James N. Porter by the Center for Cognitive Science's Interdisciplinary Training Program in Cognitive Science (NIH 2T32HD007151), funding awarded to Monica M. Luciana by the National Institute on Drug Abuse (R01DA017843), the University of Minnesota's Center for Magnetic Resonance Research (P41 RR008079-13 & P30 NS057091), the Center for Neurobehavioral Development, and the Minnesota Supercomputing Institute for Advanced Computational Research.

Special thanks to Monica M. Luciana, Paul F. Collins, Kelvin O. Lim, Angus W. MacDonald III, Kathleen M. Thomas, Ryan L. Muetzel, and the students and scientists in the Brain and Behavioral Studies Laboratory.

Abstract

Neuroimaging studies of normative human brain development indicate that the brain matures at differing rates across time and brain regions, with some areas maturing into young adulthood. High-resolution structural magnetic resonance imaging (sMRI) allows for a detailed morphometric analysis of these changes. In particular, changes in cortical thickness may index maturational progressions from an overabundance of neuropil toward efficiently pruned neural networks. Changes in sMRI measures have rarely been examined in relation to neuropsychological functions. In this study, healthy right-handed adolescents completed sMRI scanning and the Controlled Oral Word Association Test (COWAT). Age-related associations of task performance and cortical thickness were assessed with cortical-surface-based analyses. Significant correlations between increasing COWAT performances and decreasing cortical thickness were found in left hemisphere language regions, including Wernicke's and Broca's areas. Task performance was also correlated with regions associated with intellectual capacity, effortful verbal and working memory processing, as well as performance monitoring. Structure-function associations were not significantly different between older and younger subjects. However, a main effect of sex was significant in left rostral middle frontal gyrus, with the effect driven primarily by younger males. Decreases in cortical thicknesses in regions that comprise the language network likely reflect maturation toward adult-like cortical organization and processing efficiency. The changes in brain structure that support verbal fluency appear to be reached in the early teens but with separate developmental trajectories for males and females, consistent with other studies of adolescent development.

Table of Contents

Acknowledgements	i
Abstract	ii
List of Tables	iv
List of Figures	v
Introduction	1
Historical Context.....	1
Brain Behavior Research: fMRI.....	3
Brain-Behavior Research: sMRI	5
Developmental Perspectives.....	7
Hypotheses of the Current Study.....	9
Methods & Results	10
Participants	10
Protocol: Psychometric Tasks	11
Protocol: MRI Scanning & Processing	12
Statistical Approach: Behavioral Analyses	14
Statistical Approach: Neuroimaging Analyses.....	17
Discussion	19
Summary of Current Findings.....	19
Developmental Findings.....	21
Sex Differences	22
Limitations and Future Directions.....	23
Tables	26
Figures	28
References	35

List of Tables

Table 1	26
Table 2	27

List of Figures

Figure 1.....	28
Figure 2.....	29
Figure 3.....	30
Figure 4.....	31
Figure 5.....	32
Figure 6.....	33
Figure 7.....	34

Introduction

Historical Context

There is a nearly 150-year history in the literature of neuropsychology concerned with the neural bases of language. Since Paul Broca's 1861 investigations of speech-impaired patients, the bulk of this literature has been centered on the effects of lesions on the ability to produce and comprehend language, such as damage to left hemisphere frontal and temporal areas resulting in Broca's and Wernicke's aphasias, respectively (cf. Damasio & Geschwind, 1984; Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007). Since those early discoveries, the standard picture of language in the brain has been one that was compartmentalized and predominantly left lateralized.

However, the notion of discretely localized, single-function units within the brain has not withstood the test of time; instead, higher-order and expansive networks of brain regions with co-localized functional capabilities support behaviors such as language production. Indeed, even the conception of *production* and *comprehension* as the main components of language is overly restrictive. Likewise, overlooking the role of the brain's right hemisphere sells short the complexity of language, as evidenced by right side brain damage leading to deficits in the tremendously important lingual functions of prosody, inflection, and intonation (Lezak, Howieson, & Loring, 2004). The centrality of Broca's and Wernicke's areas in language production has also been called into question. In-depth structural magnetic resonance imaging (sMRI) studies of living aphasia patients have indicated that these two regions were not found to contribute to the deficit, but that other nearby structures were (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004). Additionally, sMRI reexaminations of the preserved brains of Broca's original patients

have demonstrated that those patients' lesions were far more extensive medially than Broca was able to discern (Dronkers et al., 2007). However, though the discrete functions once attributed to these regions may not be accurate, their enduring role as markers for major hubs within the broader language network is supported by studies in childhood language acquisition and dyslexia (cf. Shaywitz, Lyon, & Shaywitz, 2006) and examinations of the white matter pathways that connect these regions (for review, see Friederici, 2009).

Neuropsychological tasks can elicit discrete behavioral components, but reflective of the distributed and interconnected nature of neural structure and functioning, particular neuropsychological tests do not tap particular task-related "centers," but rather engage a wide range of cortical and subcortical brain regions. For instance, the Controlled Oral Word Association Test (COWAT) is often characterized as a test of frontal cortex functioning. The COWAT requires one to produce as many unique words as possible, all beginning with a target letter, while being explicitly timed. Proficient task performance requires more than simple verbal output, but also rule maintenance, performance monitoring, response selection, and control of performance anxiety. Left hemisphere inferior frontal gyrus damage is strongly associated with decreased COWAT performance, but a variety of conditions (e.g., dementia, temporal lobe epilepsy, Parkinson's disease) can also engender declines in performance (cf. Lezak et al., 2004). In addition, functional magnetic resonance imaging (fMRI) studies of language reception and production indicate that specific behaviors can be expected to engage specific combinations of regions, with certain areas of information processing being more extensive, or crucial, than others.

Brain Behavior Research: fMRI

Much of our knowledge on the localization of behavioral abilities comes from brain lesion and degenerative disease research, which often involves extensive traumatic injuries, widespread deterioration, and postmortem analyses. Recent advances in neuroimaging now allow investigation of relationships between brain and behavior within healthy, living subjects. However, the powerful technique of comparing typical to atypical populations is often necessary, particularly with complex behaviors such as language use.

Studies of children and adolescents with Reading Disorder (RD) highlight the relative importance of regions within the language network for different aspects of language comprehension and production. Shaywitz and colleagues (2002) recruited children and adolescents (ages 7-18), with and without RD, to perform rhyming identifications of real words and pseudowords, as well as semantic categorizations of real words, while undergoing fMRI scanning. Compared to participants with RD, the control participants showed significantly greater levels of task-related blood-oxygen-level-dependent (BOLD) signal change in the left hemisphere language systems, particularly in left inferior frontal and superior/middle temporal cortices. Remediation studies in RD have shown that language task performance increases are accompanied by normalization of such BOLD responses. Richards and Berninger (2008) found significant group differences (RD vs. controls) within seeded connectivity analyses between left inferior frontal gyrus, bilateral middle frontal gyri, right superior frontal gyrus, bilateral supplemental motor area, and left precentral gyrus. The RD children showed increased functional connectivity between left and right inferior frontal gyrus compared to controls,

indicating excessive, compensatory recruitment of non-language-devoted regions during the language task. The RD participants then completed a three-week training program in creative writing and phonological familiarity. After the training sessions, task performance improved for the RD group, and the accompanying patterns of BOLD response connectivity were statistically indistinguishable between the two groups, indicating that with training, aberrant functional connectivity patterns can quite quickly reorganize and conform to normalcy (Richards & Berninger, 2008).

Similar performance-related differences have been found for studies of verbal fluency. Wood, Saling, Abbott, and Jackson (2001) had young adult participants silently perform the COWAT while undergoing fMRI scanning, and then obtained an out-of-scanner COWAT performance measure two weeks later. Comparisons of task-related BOLD response to that during rest showed a heavily left-lateralized pattern of activation in the inferior and middle frontal cortices. Additionally, the number of significantly activated voxels within each region of interest (ROI) was positively predictive of subsequent COWAT production levels, though the relationship was statistically significant only in the left middle frontal cortex.

The BOLD response within the left-hemisphere language network has also been shown to be constant when task difficulty is manipulated, with increases in difficulty reflected by recruitment of posterior parietal cortex. Dräger et al. (2004) had young adults perform a word-stem completion task while undergoing fMRI scanning. Each word-stem was chosen for having a discrete number of task-allowable completions, providing easy, moderate, and difficult conditions. Across all three conditions, significant left-lateralized activity was shown in language-related regions including the inferior parietal, middle

frontal, and inferior frontal cortices, as well as in the caudate and cingulate. The contrast of task difficulties showed no increase in the language regions, but revealed activation increases in right posterior parietal cortex (precuneus). Thus, a dissociation was indicated between regions related to verbal retrieval and production compared to those that promote increased attention and effort (Dräger et al., 2004).

Thus, the existing literature paints a picture of language-related tasks engaging a widely distributed, but rather left lateralized, brain network. This network includes frontal and temporal regions surrounding Wernicke's and Broca's areas for phonological discrimination, language production, and word selection, regions in the superior and middle frontal lobes for rule maintenance, cognitive control, and error detection, and regions in the lateral and medial parietal lobes for "extra" processing associated with increased task difficulty and effort (for broader reviews, see Friederici, 2006 & Sakai, 2005).

Brain-Behavior Research: sMRI

The work reviewed above utilized fMRI to identify brain regions recruited during language use. Published research that utilizes sMRI to investigate regional differences in structural morphometry and how such differences correlate with neuropsychological task performance is limited, particularly for developmental populations. Co-occurring differences in brain structure and verbal fluency abilities have been found in clinical research. For example, Bonihla et al. (2008) measured COWAT performance in schizophrenia patients and healthy controls and also independently measured grey matter via voxel-based morphometry (VBM). Patients showed both deficits in COWAT performance and lower grey matter volumes in left middle frontal, superior frontal,

superior temporal, and supplemental motor cortex and in right superior frontal cortex and right operculum. However, the direct correlation between COWAT scores and grey matter volume was not reported, and the method of VBM provides only the estimated proportion of grey matter within voxels, which is not directly relatable to cortical anatomy.

The VBM method approaches the cortex within MRI's inherent rigid Cartesian coordinate system, whereas the fundamental nature of the cortex is that of a highly convoluted sheet, referred to as the cortical ribbon. Anatomically guided sMRI analyses require extraction of the cortical ribbon via boundary estimation and controlled image deformation (Fischl, Sereno, & Dale, 1999a; Fischl et al., 1999b). Such approaches allow for the selection of dependent variables that are anatomically meaningful, including the thickness, surface area, and volume of the cortex. Of these three, cortical thickness represents an ideal measure for studies of human brain and behavior development. Whereas cortical surface area and, by extension, volume are closely related to an individual's overall brain size, height, and weight, cortical thickness is relatively constant across individuals. Thus, for studies of the covariance of changes in behavioral performance and changes in brain structure, cortical thickness offers a measure that contains less variance induced by factors unrelated to cognitive abilities. Additionally, reliable measures of cortical thickness are expected, as it has been known since the work of Brodmann and von Economo that the thickness of mature human cerebral cortex has a restricted range of 1 to 4.5 mm, with inter-individual point variances of less than 1.5 mm (Fischl & Dale, 2000). This consistency is especially important for the detection of regionally specific differences in cortical thickness. Furthermore, current conceptions of

cortical thinning over development predict that cortex thins selectively as neuronal connections are pared down to the most efficient networks, a process that would not be precisely captured by volume or surface area measures.

Developmental Perspectives

Structural MRI studies indicate that grey matter proliferates through early childhood, and as adolescence begins, grey matter decreases in thickness (Shaw et al., 2008; Sowell et al., 2004) and volume (Giedd et al., 1999; Gogtay et al., 2004, Lenroot et al., 2007). The timing of these changes is regionally variable, and there is also evidence of sex effects on the onset and rate of cortical maturation, though the exact estimates vary by methodologies. Grey matter volume studies indicate sex differences such that females reach a peak volume of grey matter approximately 1-2 years earlier than males (Giedd et al., 1999; Lenroot, et al., 2007). The Lenroot et al. (2007) sample ranged in age from 3 to 27 years, and estimates of ages at peak volumes for females versus males were: parietal lobes at 7.5 versus 9 years, frontal lobes at 9.5 versus 10.5 years, and temporal lobes at 10 versus 11 years (peaks for occipital volumes were not reported). In comparison to the studies measuring volume, those measuring grey matter thickness generate sometimes strikingly different trajectories for regional cortical maturation. Grey matter thickness in the occipital lobes may peak in late childhood, as early as 8-9 years old (Shaw et al., 2008; Sowell et al., 2004). The estimates for parietal and frontal lobes roughly align with the volumetric studies, showing thickness peaks in late childhood to the early teens, but the temporal lobes are estimated to not reach peak thickness until the middle 30s (Sowell et al., 2007). Such discrepancies in findings beg the question of which grey matter measure is optimal for understanding structure-function relationships. While that question

remains unanswered, the use of thickness as the dependent measure holds several advantages, as discussed above.

Publications utilizing grey matter thickness measures in anatomically guided approaches to the structure-function question within developmental populations are beginning to emerge. Shaw et al. (2006) and Karama et al. (2009) have reported on the association of cortical thickness and performance on broad measures of intelligence. For individuals in middle adolescence and beyond (i.e., 14 and older), both groups found significant positive correlations between grey matter thickness and intelligence across both hemispheres and all four lobes of the brain, concentrated primarily in association cortex. However, for younger children, the patterns were significantly weaker (Karama et al., 2009) or reversed (Shaw et al., 2006). The wide scope of measures that tap general intelligence, or *g*, may play a part in these disparate findings, and studies investigating more circumscribed behaviors and abilities may provide clearer and more specific findings.

Sowell et al. (2004) investigated the relationship between longitudinal changes in grey matter thickness and changes in a composite measure of verbal abilities (i.e., combinations of Wechsler & Woodcock-Johnson subtests) in children aged 5 to 11. The rate of increase in verbal skills was positively correlated with the rate of cortical thinning (i.e., greater performance improvements were associated with greater thinning) in left hemisphere language regions, including middle temporal, supramarginal, precentral, caudal middle frontal, and rostral superior frontal cortices. To our knowledge, studies of adolescent brain development and its relationship to a single, established measure of verbal fluency have not been published. Additionally, studies of age-related changes in

verbal fluency abilities have been generally focused on the changes from adulthood to old age. Indeed, in two recent meta-analyses of COWAT behavioral studies, only 3 of the 26 (Rodríguez-Aranda, & Martinussen, 2006) and 4 of the 32 (Loonstra, Tarlow, & Sellers, 2001) studies investigated included participants below the age of 18.

Hypotheses of the Current Study

To address this gap, cross-sectional data on verbal fluency abilities and cortical thickness from participants in an ongoing longitudinal study were investigated. Verbal fluency was measured by performance on the COWAT, and cortical thickness was derived from high-resolution sMRI scans. We hypothesized that, independent of age, increased performance on a verbal fluency test would correlate with decreased cortical thickness in regions that have been associated with language, such as middle and superior temporal cortex, the temporal-parietal junction, and inferior and middle frontal cortex. Additional areas that have been related to increased effortful processing, rule maintenance, and sustained attention were also expected to be related to increased task performance. Such a relationship would suggest that individual differences in performance are indicative of a task-required cortical networks' structural maturity. Given prior reports on variability of structure-function relationships in younger children compared to adolescents, differential relationships between regional cortical thickness and COWAT performance were expected between younger and older participants, with the anticipation that younger participants would show a weaker structure-function relationship than older participants. Differential structure-function relationships were also expected between males and females. However, as the previous literature on sex differences in COWAT performance deals primarily with aging adult populations, no

specific predictions were made as to the structure-function relationship would differ in a developmental sample.

Methods & Results

Participants

Participants from age 9 through age 23 were recruited for a study of normative adolescent brain development. Two methods of contacting possible participants aged 9 to 17 years were used. Parents of children and adolescents were contacted through a participant database maintained by the University of Minnesota's Institute for Child Development (ICD). The Institute contacts parents listed in public birth records, and those who indicate a willingness to be contacted in the future for study participation are added to the database. Families with children within the targeted age range were chosen via random selection from the database and contacted about possible participation. Additionally, informational postcards were mailed to University of Minnesota employees who may have had children in the targeted age range, and interested parents contacted our laboratory. Participants age 18 and above were solicited through flyers posted in the University of Minnesota community. Potential participants were screened via telephone interview for exclusion criteria, including a) major physical or genetic abnormalities, b) mental retardation, c) head injuries that resulted in loss of consciousness, d) current or past neurological, psychological, or psychiatric illnesses, e) learning disabilities, f) tobacco, alcohol, or drug abuse, g) current or past use of psychoactive medications, h) being a non-native English speaker, i) abnormal or uncorrected vision and hearing, and j) MRI contraindications (e.g., severe claustrophobia, braces, pregnancy, medical implants). The University of Minnesota's Institutional Review Board approved the protocol. Adults

and parents of minors all provided informed consent, and all minor children assented to participate in the study.

Upon enrollment in the study, right-handedness was confirmed through the Edinburgh Handedness Inventory (Oldfield, 1971). Additionally, to ensure that the sample included only typically developing, healthy subjects, comprehensive mental health interviews using the Kiddie-SADS-Present and Lifetime Version (K-SADS-PL; Kaufman et al., 1997) were conducted with adult participants and separately with parents and children. Just prior to scanning sessions, all participants were re-screened for MRI contra-indications in accordance with standard MRI safety guidelines (e.g., Sawyer-Glover & Shellock, 2000).

Protocol: Psychometric Tasks

Participants completed a battery of cognitive and neuropsychological tasks as well as MRI scanning. The measures reported here were selected for these analyses because they have historically demonstrated neural correlates and performance is known to increase with age (cf. Lezak et al., 2004).

The COWAT was administered according to standardized rules (Lezak et al., 2004; Spreen & Strauss, 1998). Instructions were to generate as many words as possible beginning with a target letter within 60 seconds. Three letters were used: F, A, and S. Correct responses could not include proper nouns, numbers, or a single word root with multiple endings (e.g., speak, speaks, speaking, spoke). A COWAT total score for each participant was calculated, representing the total number of words generated across all three trials after deductions for rule violations, set-loss errors (i.e., words not beginning with target letters), and perseverations (i.e., saying the same word more than once).

Estimates of verbal intelligence (VIQ), performance intelligence (PIQ), and full scale IQ (FSIQ) were obtained via the Wechsler Abbreviated Scales of Intelligence (WASI; The Psychological Corporation, 1999). The WASI consists of four subtests, two of each in the verbal and performance domains. VIQ estimates are derived from a combination of scores from the Vocabulary and Similarities subtests. The subtests that contribute to PIQ are Block Design and Matrix Reasoning.

Additional demographic variables were obtained through self-report, including sex, age, and parental education level. Age was calculated from date of birth to date of participation. As a proxy for both socioeconomic status and intellectually enriched home environment, years of formal education were reported for both the mother and father of each participant. Ranges, means, and standard deviations are presented for demographic and psychometric measures in Table 1.

Protocol: MRI Scanning & Processing

All MRI images were acquired on a 3-Tesla Siemens Trio scanner (Siemens Medical Systems, Erlangen, Germany) at the University of Minnesota Center for Magnetic Resonance Research. Three-dimensional brain images were obtained with a coronal T1-weighted Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence (TR = 2530 msec, TE = 3.65 msec, TI = 1100 msec, 240 slices, voxel size = 1.0 x 1.0 x 1.0, flip angle = 78, FOV = 256 mm).

These high-resolution anatomical images were processed through *FreeSurfer* v.4.0.5 (Fischl et al., 2004b; <http://surfer.nmr.mgh.harvard.edu/>) to obtain estimates of cortical thickness. The technical details of these procedures are described elsewhere (Dale, Fischl, & Sereno, 1999; Dale & Sereno, 1993; Fischl & Dale, 2000; Fischl, Liu, &

Dale, 2001; Fischl et al., 2002; Fischl et al., 2004a; Fischl et al., 1999a; Fischl et al., 1999b; Fischl et al., 2004b; Han et al., 2006; Jovicich et al., 2006; Segonne et al., 2004). Image preprocessing in the *FreeSurfer* pipeline included motion correction using the FLIRT algorithm (Jenkinson & Smith, 2001; Jenkinson, Bannister, Brady, & Smith, 2002), scanner-based noise reduction via N3 intensity non-uniform correction (Sled, Zijdenbos, & Evans, 1998), translation of individual brains into Talairach space (Talairach & Tournoux, 1988), and removal of non-brain tissue from the image using a hybrid watershed/surface deformation procedure (Segonne et al., 2004).

Closer approximation of the inherent geometry of the cortical ribbon was then obtained through conversion of the static, stereotaxic, voxel-based images to a deformable, surface-based 3-D representation. This conversion included fitting tessellated “wireframes” of triangles to the pial and grey/white surface boundaries, creating what has been called a "double-hull surface pair" out of the two boundaries (Rehm, Ju, Stern, Schaper, & Rottenberg, 2004). Correction of physiological and geometric impossibilities on the surfaces (Fischl et al., 2001; Segonne, Pacheco, & Fischl, 2007), and optimized classification of the tessellation cells as grey matter, white matter, or cerebrospinal fluid (Dale et al., 1999; Dale & Sereno, 1993; Fischl & Dale, 2000) was then carried out, resulting in a wireframe model of the cortical ribbon that was deformed, or adjusted in size and shape, while retaining all prior location and geometric information at each vertex in the tessellation (Fischl et al., 1999a). This procedure allowed for highly accurate inter-subject registration based on gyral and sulcal topology (Fischl et al., 1999b), topological parcellations and labeling of sulci and gyri (Desikan et al., 2006; Fischl et al., 2004b), and statistical analysis of morphometric properties such as cortical thickness. Thickness

of the grey matter was calculated as the closest distance from the grey/white boundary to the pial boundary at each vertex (Fischl & Dale, 2000). In all, the *FreeSurfer* processing created a thickness map for the brain of each participant that was accurately aligned to a standard topological model, and these maps were then entered into a series of general linear model analyses investigating the relationships between cortical thickness and psychometric measures.

Statistical Approach: Behavioral Analyses

The total sample used in these analyses contained 168 healthy and typically developing children, adolescents, and young adults (75 males, 93 females) aged 9.29 – 23.96 years (Median=16.45, Mean=16.37, SD=3.80) with both usable COWAT data and viable MRI data. Aside from two adult males who did not report parental education levels, all participants had full demographic data available. Though the number of females in the sample was nominally greater, the sample did not significantly over represent females ($\chi^2(1)=1.93, n.s.$). Across the whole sample, there were no significant differences between females and males on COWAT total score ($t(166)=1.49, n.s.$), age ($t(166)=0.82, n.s.$), Average Parental Education ($t(164)=-0.43, n.s.$), FSIQ ($t(166)=-0.02, n.s.$), VIQ ($t(166)=0.41, n.s.$), or PIQ ($t(166)=-0.45, n.s.$).

Assessment of the developmental trajectory of COWAT performance in the whole sample was carried out by a series of one-way ANOVAs (IV=age group, DV=COWAT total score) iteratively placing participants into age bins of 10, 5, and 4 equal groups (i.e., the deciles, quintiles, and quartiles). The percentile strategy was chosen to reduce the effect of score fluctuations within short age intervals and to maximize power in later MRI analyses, by making comparisons between behaviorally distinct groups of equivalent

sizes. All ANOVAs indicated significant COWAT performance differences among the age groupings, reflective of the overall improvement in performance from the younger to older ages. Homogeneous subsets were evaluated at each step by pairwise comparisons using Tukey's HSD at $\alpha=0.05$ in order to determine behaviorally distinct groups, as indicated when no one age group was included in two or more subsets.

Analysis by binning participants in age deciles showed no coherent group structure on COWAT performance. The homogeneous subsets analysis generated five groups, with each group sharing two to four of the age bins with its neighbor. The decile groupings did reveal an overall developmental increase in performance, and the plot of the group means revealed a distinct mean score difference between groups below the 50th and above the 60th percentiles. A performance "changeover point" near the median age was still apparent when participants were placed into age quintile bins. However, demarcation of homogeneous subsets was still unclear, with the groups clustering into three overlapping subsets. With quartile binning, it became clear that the two oldest groups (i.e., participants 16.45-23.96 years old) were not significantly different from each other in terms of COWAT performance, but that they were significantly different from each of the youngest two groups (i.e., 9.29-16.45 years old). Additionally, the two younger groups were not significantly different from each other, indicating that the developmental trajectory reflected a median split in the age range. Thus, within this sample, overall performance abilities on the COWAT appeared to develop until a point near the sample's median age of 16.45 years (16 years, 5 months), with performance leveling off across ages thereafter (see Figure 1 for graphical summaries, and Table 2 for the homogeneous subset analyses).

However, there is evidence that adult males and females exhibit different overall verbal fluency performance abilities (Loonstra et al., 2001; Rodríguez-Aranda & Martinussen, 2006). To investigate the possibility of different developmental trajectories for verbal fluency in males and females, we wished to examine the effect of sex within the quartile split structure upon COWAT total score with an expanded ANOVA analysis. Additionally, for the planned neuroimaging analyses it was important to determine which performance-relevant factors representing different domains of influence should be statistically controlled. To that end, sex and age group were included in the set of categorical predictors, and additional continuous covariates of intellectual capacity (IQ) and average parental education were included in the ANOVA model. Average parental education was chosen as a marker for enriched learning environment and increased socioeconomic status, and intellectual capacity (i.e., IQ) was chosen to account for the participants' innate abilities to perform well on cognitive tests. Including all three IQ measures (i.e., VIQ, PIQ, & FSIQ) in an expanded ANOVA would violate assumptions of collinearity and independence of observations, so which IQ measure to use was assessed by comparing their partial correlations with COWAT total score, controlling for the effects of quartile age group and sex. VIQ ($r_{Age-Sex}=.29, p<0.05$) was chosen, as it was nominally more strongly correlated than FIQ ($r_{Age-Sex}=.27, p<0.05$); PIQ ($r_{Age-Sex}=.17, n.s.$) was not significantly correlated with COWAT total score.

Thus, a 4x2 ANOVA was run, with quartile placement and sex entered as fixed factors and VIQ and average parental education entered as continuous covariates. There were significant main effects of age group ($F(3,156)=3.59, p<0.05$, partial eta squared=0.43) and VIQ ($F(1,156)=14.6, p<0.05$, partial eta squared=0.08), but there were

not main effects for average parental education ($F(1,156)=1.53, n.s.$) or sex ($F(1,156)=0.34, n.s.$). There was, however, a significant interaction effect of age group*sex ($F(3,156)=3.59, p<0.05$, partial eta squared=0.06). Follow-up Bonferroni-corrected t -tests within each quartile showed the interaction effect was driven by high-scoring females ($M(SD)=37.00(9.09)$) compared to lower-scoring males ($M(SD)=28.76(7.79)$) in the second quartile ($t(41)=3.81, p<.0125$), indicating that there may possibly be separate verbal fluency developmental trajectories for the males and females in this sample (see Figure 2).

Statistical Approach: Neuroimaging Analyses

To summarize, the behavioral analyses showed that, regardless of sex, there was a clear demarcation in COWAT total score above and below the median age in this sample. Neuroimaging analyses were planned for the full sample, between the younger than and older than median age groups, as well as additional follow-up investigations of the quartile age groups.

Assessments of the covariation of COWAT total scores and cortical thickness were performed with a general linear model approach, using *FreeSurfer's* `mri_glmfit` program. In addition to COWAT total score, based on the behavioral analyses described above, the variables included in the regression matrix for the analyses included age group, sex, and VIQ. Thus, the correlation between COWAT total score and cortical thickness was assessed across and within different age groups by regressing performance upon cortical thickness at each vertex of the reconstructed cortical surface, while accounting for the effects of individual mean cortical thickness, sex, and verbal intelligence. In order to maintain proper statistical control over family wise error in the

face of thousands of statistical tests across the cortical surface vertices, the results of all analyses were thresholded via the False Discovery Rate method (FDR; Genovese, Lazar, & Nichols, 2002) at an alpha value of $p \leq 0.05$.

The full sample analysis yielded a distributed pattern of clusters across both hemispheres, including middle temporal cortex, supramarginal cortex, pars opercularis, rostral middle frontal cortex, superior frontal cortex, precentral cortex, lateral occipital cortex, inferior parietal cortex, superior parietal cortex, precuneus, posterior cingulate, and anterior cingulate. For a full map of the significant clusters and a listing of regions covered by them, see Figure 3.

A comparison of the above and below median age groups yielded no regions wherein the between-groups difference in the strength of COWAT-cortical thickness correlation could pass FDR correction. Relaxing the thresholding criteria to $p \leq 0.01$, uncorrected for multiple comparisons, showed a few key areas of difference in this contrast (see Figure 4), with regions within left hemisphere precentral, rostral middle frontal, and superior temporal cortices showing stronger COWAT-cortical thickness negative correlations in the older participants. Regions in the left hemisphere pars opercularis, rostral middle frontal cortex, inferior temporal cortex, posterior cingulate, fusiform, and precuneus showed a stronger negative correlation in the younger participants. In the right hemisphere, younger participants showed stronger negative relationships in supramarginal cortex, superior temporal cortex, and cingulate isthmus; Older participants showed stronger negative relationships in superior frontal cortex. However, at such a lax statistical threshold, great caution is warranted when interpreting the differential importance of these regions to COWAT performance across development.

There was a region that showed significant difference between the correlation strengths for females and males, located in the left rostral middle frontal cortex (inferior to the rostral middle frontal region discussed above). Graphical investigation of the effect indicated that male and females have a strikingly opposite pattern of the COWAT performance-cortical thickness relationship, particularly within the younger ages (see Figure 5). No right hemisphere regions survived FDR correction for this comparison, nor did any regions in a general comparison of average cortical thickness between males and females, independent of COWAT performance.

Statistical power was considerably reduced for comparisons of the effects within the smaller quartile groups, and no comparisons passed the FDR thresholding procedures. Relaxing the thresholds considerably to $p \leq 0.01$, uncorrected for multiple comparisons, revealed a pattern of clustering that resembled the whole sample main effect of increased COWAT performance, and reiterated that the comparison of that effect between females and males was driven by differences among the participants within the younger age range. See Figures 6 and 7 for details on regional variations by quartile age groups and sex.

Discussion

Summary of Current Findings

The hypothesis that increased performance on a verbal fluency test would correlate with decreased cortical thickness in regions that have been associated with language was supported. The full sample results included clusters showing a correlation between increased performance and thinner cortex in areas commonly associated with language functions (Sakai, 2005; Shaywitz, et al., 2006) as well as areas associated with

increased intellectual capacity (Karama et al., 2009; Shaw et al., 2006), online performance monitoring (MacDonald, Cohen, Stenger, & Carter, 2000), and with increased effort and concentration (Dräger et al., 2004). Language-related regions included bilateral superior and middle temporal cortex, supramarginal cortex, pars opercularis, rostral middle frontal cortex, lateral occipital cortex, and inferior parietal cortex. The correlation values surrounding Wernicke's area were stronger in the left hemisphere than in its right hemisphere homologue. With respect to Broca's area, the left hemisphere pars opercularis was identified, and on the right hemisphere a region spanning pars opercularis and pars triangularis was identified. Quite interestingly, the full sample analysis even revealed a cluster in the left postcentral gyrus (i.e., somatosensory cortex) that would correspond to the region that receives sensory information from the lips, tongue, and throat.

Areas that have been related to increased effortful processing, rule maintenance, sustained attention, and performance monitoring were identified in both hemispheres. Lateral surface clusters in this group were more abundant in the right than left hemisphere, and included middle and superior frontal cortex, superior parietal cortex, and inferior parietal cortex. Bilateral medial surface clusters were predominately situated among the precuneus, posterior cingulate, and anterior cingulate; right hemisphere medial orbitofrontal and superior frontal cortex clusters were also found.

That the clusters found in these analyses are not restricted to the regions generally associated with language production likely reflects that the COWAT is not a "pure" index of verbal production, but also of broader executive functioning. The task requires one to not just use verbal skills, but also general executive functioning skills such as rule

maintenance, impulse inhibition, and manipulation of working memory. Additionally, COWAT performance has been shown to be highly reflective of underlying intellectual capacity, as indexed not simply by one's age, but by one's years of formal education beyond high school (Loonstra et al., 2001; Rodríguez-Aranda & Martinussen, 2006). Thus, in this study the broader executive functioning components of the COWAT are reflected in the identification of regions including superior parietal cortex, dorsolateral prefrontal cortex, precuneus, cingulate, and medial orbitofrontal cortex. The negative correlational relationship between verbal fluency abilities and regional cortical thickness suggests that increasing COWAT performance is indicative of an increasingly mature cortical network underlying language- and test performance-related abilities.

Developmental Findings

Given prior reports on variability of structure-function relationships in younger children compared to adolescents (Karama et al., 2009; Shaw et al., 2006), as well as the behavioral differences observed in this sample, differential relationships between regional cortical thickness and COWAT performance were expected between younger and older participants. This hypothesis was not fully supported by the data from this sample. There was no overall difference between the strength of the COWAT performance-cortical thickness correlation between the older and younger participants when the effect of sex was controlled. Looking at the median and quartile subgroups in isolation undercut the analyses in terms of power to detect an effect after correction for multiple comparisons, so finer analyses of differential effects within each subgroup was not possible. Looking across the quartile analyses, there did appear to be a sense of broadening of the cortical

network associated with COWAT performance as age increased, but only if statistical thresholds were dropped to untenable levels (e.g., $p < .05$, uncorrected).

Sex Differences

There was a significant difference between the correlation strength between females and males across the sample that was robust to correction for multiple comparisons. This difference was primarily driven by males in the younger half of sample, and, at first glance, presents an intriguing correspondence between the behavioral analyses showing a distinct performance difference between males and females between the ages of 13.55 and 16.45 years. However, the low-powered quartile analyses would seem to indicate that the sex effect is coming from differences in the COWAT performance and grey matter thickness correlation within the youngest participants (ages 9.29 to 13.55). This may indicate a difference in the manner in which males and females of that age approach the task, rather than their overall output scores. However, this finding may also be driven by sex differences in the timing of brain maturation (Lenroot et al., 2007; Sowell et al., 2007) that is related to a more general change in executive functioning. Indeed, Lenroot et al. (2007) estimated that females begin the process of maturational thinning in the frontal lobes at approximately 9.5 years of age, whereas males were estimated to begin the process around 10.5 years. If this were accurate, it would indicate that the females in the first age quartile of this sample had an extra year, on average, of maturation in a region that subserves major components of COWAT performance. That the males in that age range show a strong *positive* correlation between COWAT performance and grey matter thickness in the rostral middle frontal cortex supports this postulate. This evidence may indicate that the source of such reversals in

correlation direction is due not just to differences in age, as seen in the general intelligence studies of Shaw et al. (2006), but due to differences in the timing of pubertal development. Unfortunately, cross-sectional studies such as this one are unlikely to be able to tease apart the differential effects of general aging versus pubertal timing, as the two measures are highly correlated. Proper investigation of changes in pubertal status versus changes in age requires longitudinal methods that allow for modeling of both between- and within-subjects variances. The data for these cross-sectional analyses come from a larger longitudinal study, which has recently completed data collection for a second time point. Longitudinal analyses of COWAT performance and grey matter thickness across the two time points are planned.

Limitations and Future Directions

As this study is correlational in nature, it is not possible to directly infer that thinner cortex in the regions identified will necessarily lead to improved verbal fluency. The overlap in regions identified in these analyses with those identified as correlates of broader intellectual capacity (e.g., Karama et al., 2009; Shaw et al., 2006) could suggest an alternative explanation, namely that the findings in this study are driven by possible *g* saturation of the COWAT. However, these analyses included VIQ as a covariate, in order to explicitly partial out the effect of intellectual capacity. One may argue that there is an effect of *g* above and beyond the VIQ measure that we have not accounted for, but given the correlation in this sample between VIQ and FSIQ (which theoretically should provide a stronger index of *g* saturation), such an explanation is unlikely.

An additional limitation is that the word “network” may not accurately describe the pattern of clusters identified here, as no functional imaging of the brain was

performed. It is possible that the basic structural organization of the language network is in place by the early teens, but that further maturational differences will be better indexed by functional differences within the network. Since collection of COWAT data using standard administration practices is unlikely during fMRI scanning, additional structure-function methods should be pursued. Future studies utilizing techniques such as probabilistic tractography via diffusion tensor imaging could shed light upon the strength of anatomical connectivity between the language network regions, by identifying white matter pathways that connect the clusters to one another. Clarification of the importance of these regions to language processing could then be gained through examining the correlations between white matter integrity measures and verbal fluency. Such analyses could also shed light upon the nature of the relationship of regions across the hemispheres. A related limitation to this study is that it has recruited only right-handed participants, and the data cannot speak to possible differences in the brain's language network between left- and right-handed individuals.

A further limitation inherent in the methodology of cortical surface reconstruction via *FreeSurfer* is simply that it focuses solely on the cortex. In doing so, we implicitly disregard the influence of other brain structures that are likely to contribute to effective verbal fluency. Indeed, central roles in language processing for subcortical structures have been proposed (Lieberman, 2002), as have important roles for regions of the cerebellum (Ackerman, Mathiak, & Riecker, 2007). It should be clearly stated that the analyses described indicate that the regions identified may be necessary for the performance of verbal fluency tasks, but by no means do these analyses indicate that they are fully sufficient for doing so.

In sum, significant correlations between improved COWAT performances and decreased cortical thickness were found in regions corresponding to left hemisphere language regions, including the traditional language hubs of Wernicke's and Broca's areas, as well as temporal cortex related to auditory processing, and somatosensory cortex related to the operation of the lips, tongue, and mouth. Significant right hemisphere regions included frontal and parietal areas associated with effortful verbal processing, increased working memory load, and performance monitoring. These associations were found when age, VIQ, and sex were controlled. The structure-function associations were not stronger in older compared to younger subjects, but did differ between the sexes, with the effect driven primarily by differences between younger males and females. Thus, cortical thickness decreases in regions that comprise these language and executive functioning networks likely begin maturation toward adult-like cortical organization and processing efficiency by late childhood. The changes in brain structure that support the Controlled Oral Word Association Task appear to be reached by the early to mid teens, but regional the timing of maturational completion is different between males and females, consistent with other studies of frontal lobe development.

Tables

Table 1. Measures of interest.

Measure	Full Sample	Age < Median	Age > Median
Sex			
Female	93	44	49
Male	75	41	35
Age (years)			
Range	9.29 – 23.96	9.29 – 16.45	16.49 – 23.96
Mean (SD)	16.40(3.81)	13.21(2.22)	19.62(1.88)
COWAT total words			
Range	8 – 88	8 – 58	21 – 88
Mean (SD)	37.46(12.02)	30.46(9.22)	44.55(10.28)
Avg. Parental Education (years)			
Range	10.5 – 23.0	11.5 – 20.5	10.5 – 23.0
Mean (SD)	16.00(2.03)	16.03(1.81)	15.96(2.26)
FSIQ			
Range	75 – 148	75 – 148	90 – 134
Mean (SD)	116.03(10.21)	116.15(12.00)	115.90(8.08)
VIQ			
Range	80 – 146	80 – 146	89 – 136
Mean (SD)	114.62(10.23)	114.81(11.23)	114.42(9.17)
PIQ			
Range	74 – 149	74 – 149	94 – 129
Mean (SD)	113.70(11.47)	113.74(13.91)	113.65(8.38)

Table 2. Homogeneous subsets analyses for the decile, quintile, and quartile age groups. Statistically reliable group differences (via Tukey's HSD at $\alpha=0.05$) are only obtained for the median split comparison within the quartile analysis. Group means are displayed in the cells.

Decile	N	1	2	3	4	5
1	17	24.59				
3	18	29.17	29.17			
2	17	30.41	30.41			
4	16	33.06	33.06	33.06		
5	17		35.29	35.29	35.29	
7	17			41.71	41.71	41.71
6	17			42.24	42.24	42.24
8	17			43.35	43.35	43.35
9	16				45.00	45.00
10	16					49.81

Quintile	N	1	2	3
1	34	27.50		
2	34	31.00		
3	34		38.76	
4	34		42.53	42.53
5	32			47.41

Quartile	N	1	2
1	42	27.88	
2	43	32.98	
3	42		42.64
4	41		46.10

Figures

Figure 1. Age-related trajectories of COWAT performance improvements with varied age bins. Results of one-way ANOVAs (IV=age group, DV=COWAT total score) are inset. Left to right: decile, quintile, and quartile groupings.

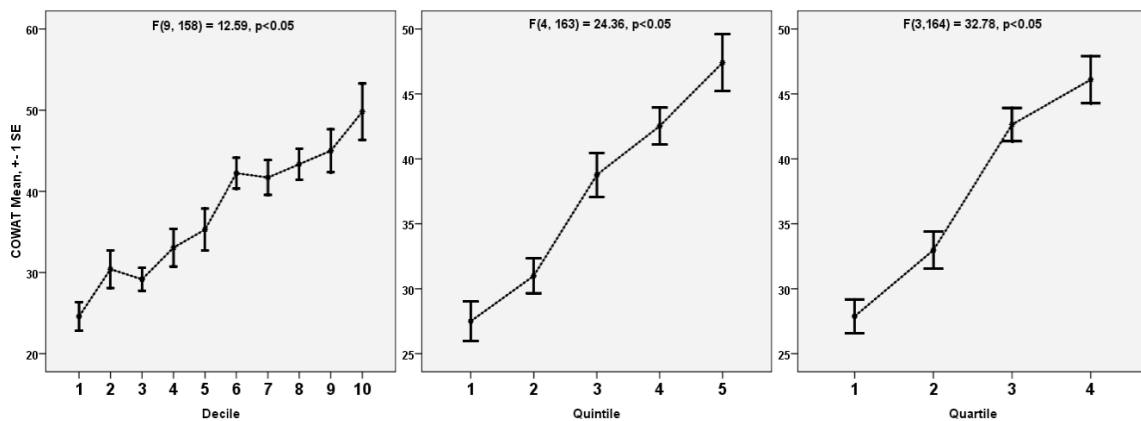


Figure 2. Age-related trajectories for COWAT performance improvements, accounting for quartile age group and sex. Within quartiles, the difference between males and females is only statistically significant in the 2nd quartile (ages 13.58-16.45).

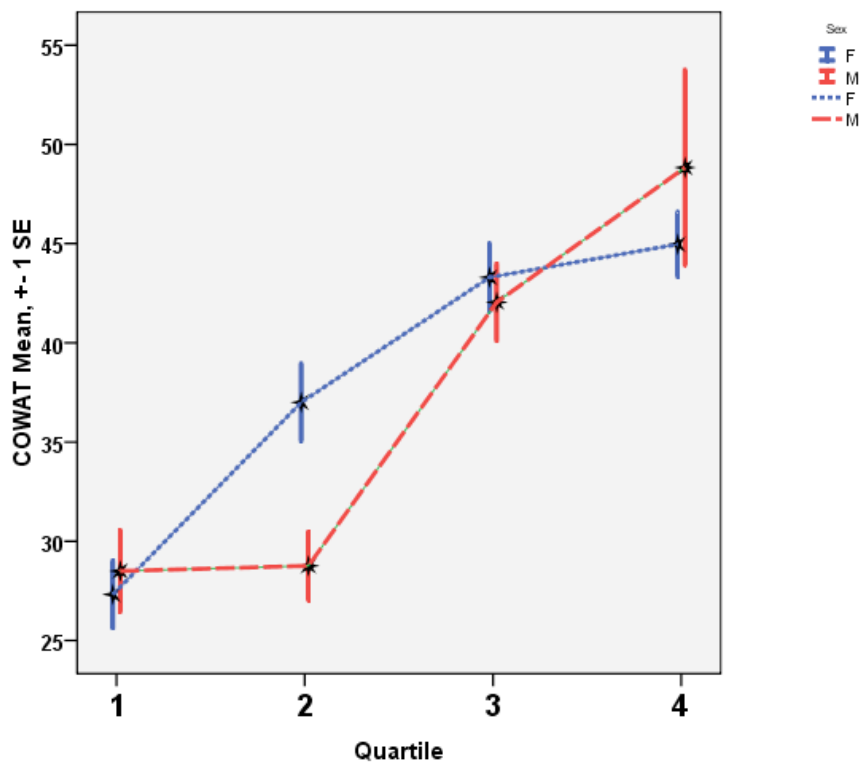


Figure 3. Areas of significant correlation between COWAT total score and grey matter thickness, controlling for the effects of age group, sex, and VIQ in the full sample. Statistics are FDR corrected at $\alpha = p \leq 0.05$. Regions found in the left hemisphere: are superior frontal, rostral middle frontal (2), pars opercularis, insula, postcentral (2), supramarginal (3), superior parietal, inferior parietal (3), lateral occipital, middle temporal, superior temporal, lingual, precuneus, paracentral, posterior cingulate, caudal anterior cingulate, and rostral anterior cingulate. Regions found in the right hemisphere are: pars orbitalis (2), pars triangularis, lateral orbitofrontal, rostral middle frontal, caudal middle frontal, superior frontal, precentral (2), postcentral, supramarginal (2), superior parietal (2), inferior parietal (2), superior temporal, middle temporal, lateral occipital, lingual, fusiform, precuneus, cingulate isthmus, posterior cingulate, medial superior frontal, and medial orbitofrontal.

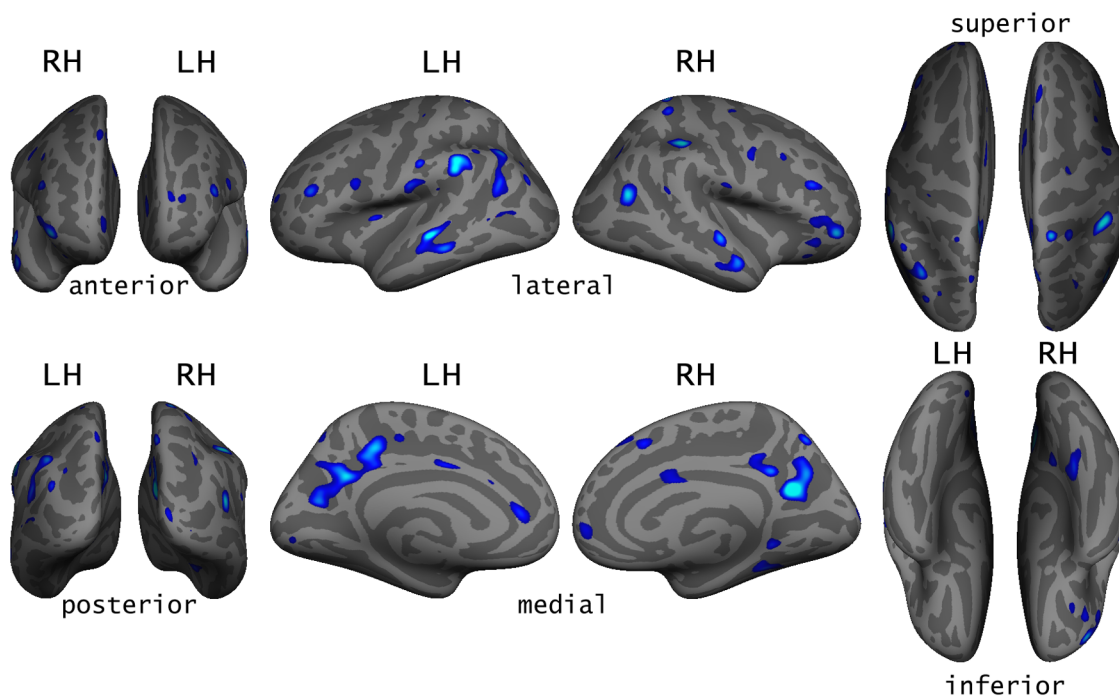


Figure 4. Regions where the strength of the correlation between COWAT performance and grey matter thickness differs between participants above and below the median age of the sample. Red areas indicate a stronger negative correlation within the older participants; blue areas indicate a stronger negative correlation in the younger participants. Regions in the left hemisphere are: rostral middle frontal, precentral (2), postcentral, superior temporal, lateral occipital (2), fusiform, and isthmus cingulate. Regions in the right hemisphere are: superior frontal, supramarginal, superior temporal (2), and isthmus cingulate. Note: Statistics in this image are thresholded at $p < 0.01$, uncorrected for multiple comparisons.

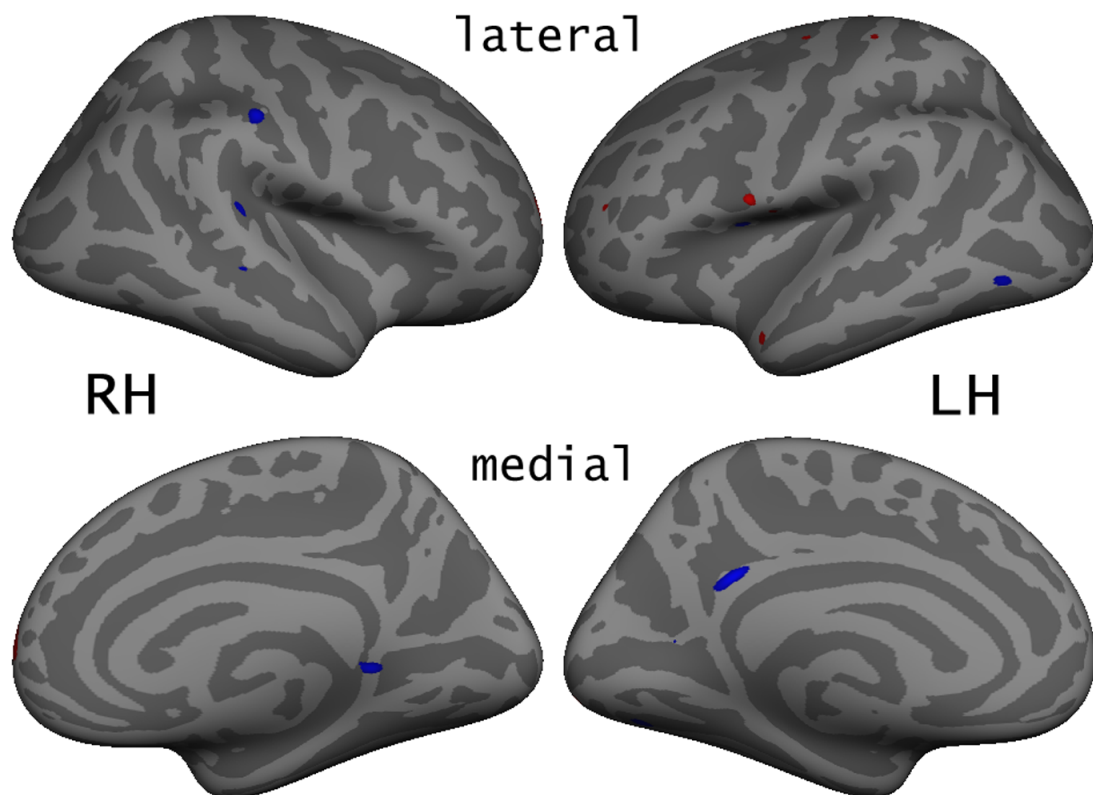


Figure 5. The region in left rostral middle frontal cortex where the direction and magnitude of the correlation between COWAT performance and grey matter thickness differs between males and females. The scatterplot shows that females display the expected negative correlation in both age groups, whereas males show a positive correlation, markedly so for the younger males in the sample.

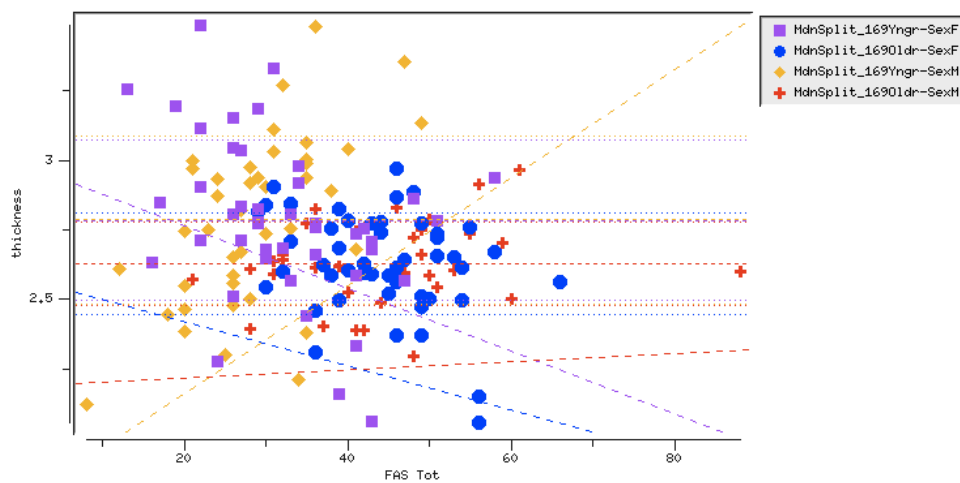


Figure 6. Regions showing correlation between COWAT performance and grey matter thickness, by age quartile. Blue indicates a negative correlation; red indicates a positive correlation. Note: Statistics in this image are thresholded at $p < 0.01$, uncorrected for multiple comparisons.

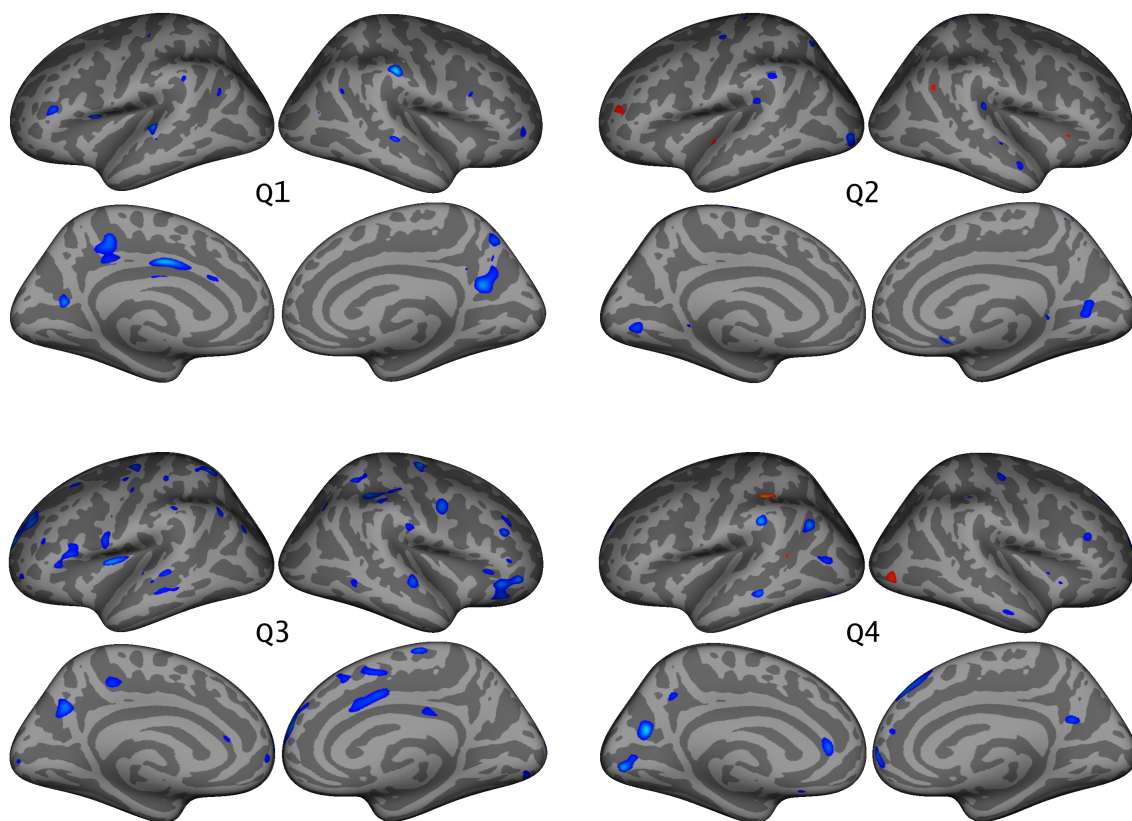
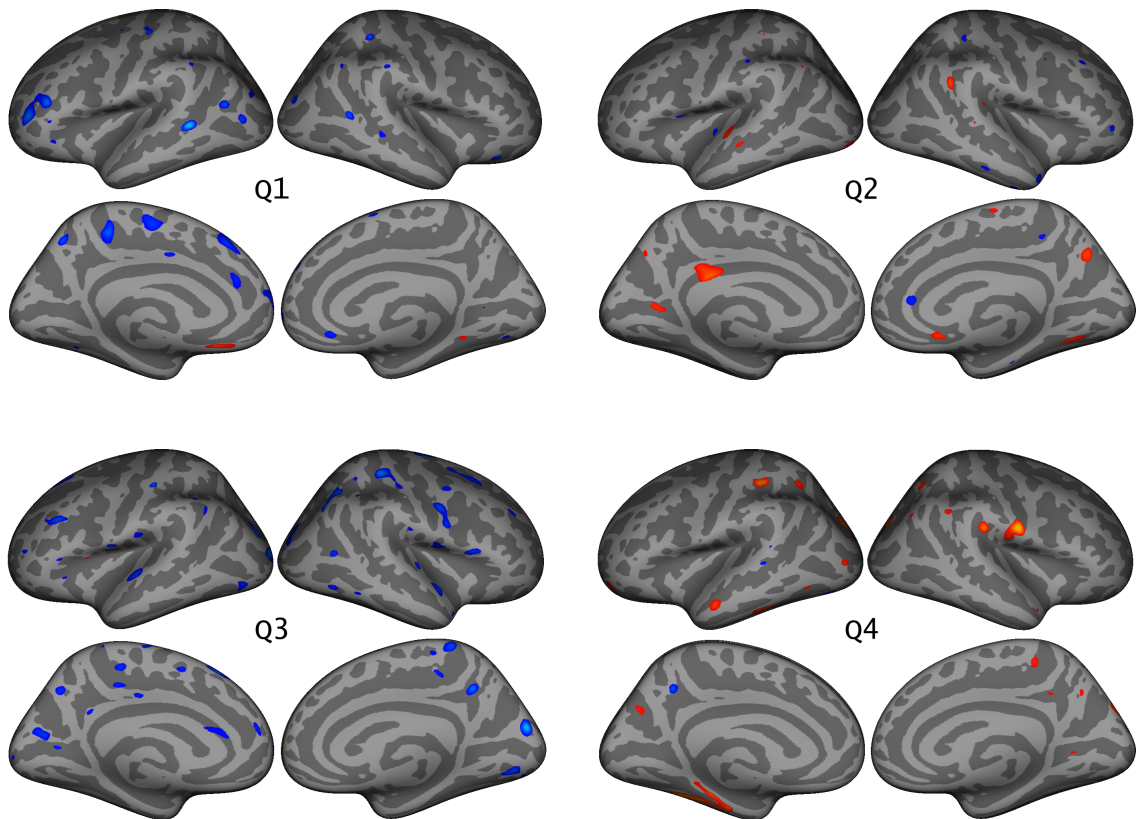


Figure 7. Regions showing a sex difference in the strength or direction of the correlation between COWAT performance and grey matter thickness, by age quartile. Blue indicates a stronger negative correlation for females; red indicates a stronger negative correlation for males. Note: Statistics in this image are thresholded at $p < 0.01$, uncorrected for multiple comparisons.



References

- Ackerman, H., Mathiak, K., & Piecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. *The Cerebellum, 6*, 202-213.
- Bonilha, L., Molnar, C., Horner, M. D., Anderson, B., Forster, L., George, M. S., et al. (2008). Neurocognitive deficits and prefrontal cortical atrophy in patients with schizophrenia. *Schizophrenia Research, 101*, 142-151.
- Dale, A. M., Fischl, B., & Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*, 179-194.
- Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience, 5*, 162-176.
- Damasio, A. R., & Geschwind, N. (1984). The neural basis of language. *Annual Review of Neuroscience, 7*, 127-147.
- Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31*, 968-980.
- Dräger, B., Jansen, A., Bruchmann, S., Förster, A. F., Pleger, B., Zwitserlood, P., et al. (2004). How does the brain accommodate to increased task difficulty in word finding? A functional MRI study. *NeuroImage, 23*, 1152-1160.
- Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain, 130*, 1432-1441.

- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Jr., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*, 145-177.
- Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11050-11055.
- Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Transactions on Medical Imaging*, *20*, 70-80.
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341-355.
- Fischl, B., Salat, D. H., van der Kouwe, A. J., Makris, N., Segonne, F., Quinn, B. T., et al. (2004a). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23*, S69-S84.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999a). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*, 195-207.
- Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999b). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*, 272-284.

- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., et al. (2004b). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*, 11-22.
- Friederici, A. D. (2006). The neural basis of language development and its impairment. *Neuron*, *52*, 941-952.
- Friederici, A. D. (2009). Pathways to language: Fiber tracts in the human brain. *Trends in Cognitive Sciences*, *13*, 175-181.
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870-878.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, *2*, 861-863.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8174-8179.
- Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., et al. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*, 180-194.
- Jenkinson, M., & Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*, 143-156.

- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. M. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*, 825-841.
- Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., et al. (2006). Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*, 436-443.
- Karama, S., Ad-Dab'bagh, Y., Haier, R. J., Deary, I. J., Lyttelton, O. C., Lepage, C., et al. (2009). Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. *Intelligence*, *37*, 145-155.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 980-988.
- Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., et al. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, *36*, 1065-1073.
- Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment*. Oxford University Press, New York.
- Lieberman, P. (2002). On the nature and evolution of the neural bases of human language. *Yearbook of Physical Anthropology*, *45*, 36-62.
- Loonstra, A. S., Tarlow, A. R., & Sellars, A. H. (2001). COWAT metanorms across age, education, and gender. *Applied Neuropsychology*, *8*, 161-166.

- MacDonald, A. W. 3rd, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835-1838.
- Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97-113.
- Rehm, K., Ju, L., Stern, J., Schaper, K., & Rottenberg, D. (2004). ParcelMan and Pigment: Interactive tools for parcellation of cortical grey matter. *NeuroImage*, *22*, A287.
- Richards, T. L., & Berninger, V. W. (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. *Journal of Neurolinguistics*, *21*, 294-304.
- Rodríguez-Aranda, C., & Martinussen, M. (2006). Age-related differences in performance of phonemic verbal fluency measured by Controlled Oral Word Association Task (COWAT): A meta-analytic study. *Developmental Neuropsychology*, *30*, 607-717.
- Sakai, K. L. (2005). Language acquisition and brain development. *Science*, *310*, 815-819.
- Sawyer-Glover, A. M., & Shellock, F. G. (2000). Pre-MRI procedure screening: Recommendations and safety considerations for biomedical implants and devices. *Journal of Magnetic Resonance Imaging*, *12*, 92-106.
- Segonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*, 1060-1075.

- Segonne, F., Pacheco, J., & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Transactions on Medical Imaging*, *26*, 518-529.
- Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., et al. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, *440*, 676-679.
- Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., et al. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience*, *28*, 3586-3594.
- Shaywitz, B. A., Lyon, G. R., & Shaywitz, S. E. (2006). The role of functional magnetic resonance imaging in understanding reading and dyslexia. *Developmental Neuropsychology*, *30*, 613-632.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*, 101-110.
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, *17*, 87-97.
- Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience*, *24*, 8223-8231.

- Sowell, E. R., Peterson, B. S., Kan, E., Woods, R. P., Yoshii, J., Bansal, R., et al. (2007). Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. *Cerebral Cortex*, *17*, 1550-1560.
- Spren, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.
- Talairach, J., & Tournoux, P. (1988). *Co-Planar Stereotactic Atlas of the Human Brain*. Stuttgart/New York: Georg Thieme Verlag/Thieme Medical Publishers.
- The Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI) manual*. San Antonio, TX: The Psychological Corporation.
- Wood, A. G., Saling, M. M., Abbott, D. F., & Jackson, G. D. (2001). A neurocognitive account of frontal lobe involvement in orthographic lexical retrieval: An fMRI study. *NeuroImage*, *14*, 162-169.