# Associations Between Cortical Thickness and Verbal Fluency in Childhood, Adolescence, and Young Adulthood

**James N. Porter**[1,4,5], **Paul F. Collins**[1,4], **Ryan L. Muetzel**[1,3], **Kelvin O. Lim**[2,3,4], and **Monica Luciana**[1,4,5]

[1] Department of Psychology, University of Minnesota

[2] Department of Psychiatry, University of Minnesota

[3] Center for Magnetic Resonance Research, University of Minnesota

[4] Center for Neurobehavioral Development, University of Minnesota

[5] Center for Cognitive Sciences, University of Minnesota

## 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. In particular, changes in cortical thickness may index maturational progressions from an overabundance of neuropil toward efficiently pruned neural networks. Developmental changes in structural MRI measures have rarely been examined in relation to discrete neuropsychological functions. In this study, healthy right-handed adolescents completed MRI scanning and the Controlled Oral Word Association Test (COWAT). 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 perisylvian regions surrounding Wernicke's and Broca's areas. Task performance was also correlated with regions associated with effortful verbal processing, working memory, and performance monitoring. Structure–function associations were not significantly different between older and younger subjects. 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 cortical thicknesses that support verbal fluency are apparent by middle childhood, but with regionally separate developmental trajectories for males and females, consistent with other studies of adolescent development.

## Introduction

There is a nearly 150-year history in the literature of neuropsychology concerned with the neural bases of language. Much of that history has relied upon studies of individuals with brain lesions and disorders, but recent advances in brain imaging techniques have allowed for the assessment of brain–behavior relationships within healthy individuals. Functional imaging is often used to assess the degree of relationship between neuronal activity and

Corresponding Author: James N. Porter, University of Minnesota, Department of Psychology, 75 East River Road, Elliott Hall N218, Minneapolis, MN 55455, +1 612-626-0450 (P), +1 612.626.2079 (F), porterj@umn.edu.

linguistic abilities (Price 2000; Price 2010), with a large number of studies describing the functional neural correlates of verbal fluency (Costafreda et al., 2006; Indefrey & Levelt 2000). On the other hand, structural imaging studies of performance on standardized neuropsychological verbal fluency tasks are less numerous (Grogan et al., 2009; Lee et al., 2007), and studies relating linguistic performance of any kind to brain structure within healthy developmental populations are particularly rare (Sowell et al., 2004).

Since Broca's (1861) and Wernicke's (1874) early investigations of speech-impaired patients, the bulk of brain–behavior literature has been centered on the effects of lesions on the ability to produce and comprehend language, such as damage to the posterior portions of the inferior frontal gyrus (i.e., pars triangularis and pars opercularis) and posterior superior temporal gyrus in the left hemisphere resulting in the verbal dysfluency syndromes of Broca's and Wernicke's aphasias, respectively (Damasio & Geschwind, 1984; Dronkers et al., 2007). Since those early discoveries, the standard picture of language in the brain has evolved as one that is compartmentalized and 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 complex behaviors such as language production.

In-depth structural magnetic resonance imaging (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). Additionally, sMRI studies of living aphasia patients have indicated that direct damage to Broca's and Wernicke's regions was not found to contribute to the observed deficits, but rather that lesions in other nearby structures were more commonly associated with aphasic syndromes (Dronkers et al., 2004). Though the discrete functions once attributed to these regions may not be accurate, their enduring roles as markers for major "hubs" within the broader language network is supported by research that inverts the lesion research paradigm, working from discrete behavioral and cognitive components towards brain-based associations.

As lesion-based brain–behavior research continued, myriad neuropsychological tests were designed to characterize and quantify discrete deficits (e.g., verbal dysfluency; Lezak et al., 2004). Such standardized neuropsychological tasks were originally designed to elicit discrete behavioral components, but reflective of the cognitive complexity of the tasks and the distributed and interconnected nature of neural structure and functioning, it must be asserted that particular neuropsychological tests do not simply 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. Thus, at its core, the COWAT is a test of verbal fluency. However, proficient task performance not only requires ample verbal output but also requires behavioral control abilities that fall under the rubric of "executive function," including rule maintenance, performance monitoring, response search and selection, and control of performance anxiety (Lezak et al., 2004). Left hemisphere inferior frontal gyrus damage is indeed 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 (Lezak et al., 2004). In addition, fMRI studies of language reception and production indicate that specific verbal behaviors can be expected to engage specific combinations of regions, with certain areas of information processing being more extensive, or crucial, to performance of task components than others. Thus, the COWAT can be conceptualized as a measure of language fluency that requires, in addition to basic language skills, a number of executive processes.

In keeping with this view, Indefrey and Levelt (2000) presented a componential model of speech production that posits that almost every spontaneous utterance requires six levels of mental processing: conceptual preparation, grammatical encoding, morphophonological encoding, phonetic encoding, articulation, and performance monitoring. In this formulation, the COWAT requires all six levels of processing. Each of these cognitive components can be isolated by comparing data across different language-based tasks (i.e., subtractive isolation), and in their meta-analysis of 58 functional neuroimaging studies of language production, Indefrey and Levelt (2000) demonstrated that each component can also be associated with unique constellations of brain regions. Conceptual preparation was associated with the middle section of left middle temporal gyrus. Phonological encoding was associated with the posterior sections of inferior frontal gyrus (i.e., Broca's area), the middle section of superior temporal gyrus, left fusiform gyrus, the middle section of the left inferior frontal gyrus (i.e., pars opercularis), and left thalamus. Phonological code retrieval was associated with left thalamus and the posterior sections of left middle temporal gyrus and superior temporal gyrus (i.e., Wernicke's area). Articulation was associated with the anterior section of superior temporal gyrus, bilateral ventral primary and sensory motor areas, right hemisphere supplementary motor area, and the cerebellum. Subsequent meta-analyses of language-based fMRI experiments have broadly replicated these regional functional demarcations within the language network (Binder et al., 2009; Vigneau et al., 2006), and newer dynamic causal modeling techniques have been used to provide evidence for how portions of this network operate in sequence during performance of verbal fluency tasks. Eickhoff et al. (2009) identified a temporal order of activation flowing from left hemisphere par opercularis through insula, caudate, cerebellum, and then to premotor/motor cortex areas corresponding to facial control.

The research discussed above utilized functional imaging methods to identify brain regions recruited during a variety of linguistic tasks, and does not demonstrate specific effects of brain structure upon verbal fluency. Modern lesion studies clearly demonstrate the need to preserve structural integrity within certain brain regions to maintain optimal language capabilities. For example, performances on a variety of verbal fluency tasks are hindered by lesions in bilateral supplementary motor area (Alexander et al., 1989), left hemisphere dorsolateral prefrontal lesions (e.g., BAs 9, 46, & 47; Eskes et al., 2003), and lesions throughout the frontal and temporal lobes (Troyer et al., 1998). However, conceptual replications of such work within non-lesioned brains that investigate regional differences in structural morphometry and how such differences correlate with differences in language abilities, particularly standardized COWAT performance, are limited. A positive correlation between grey matter volume and COWAT scores has been found in within left inferior frontal gyrus in healthy adults (Newman et al., 2007), and grey matter insufficiency in left hemisphere middle frontal, superior frontal, superior temporal, and supplemental motor cortex has been related to decreased COWAT performance in psychiatric populations (Bonilha et al., 2008).

Because the COWAT taps into a number of executive processes that are known to improve throughout childhood and into adulthood, it seems well suited to index these developmental processes. However, examinations of verbal fluency and brain structure within healthy developmental populations are also rare, with no studies to date reporting on structural correlates of standardized COWAT performance. Sowell et al. (2004) investigated the relationship between longitudinal changes in grey matter thickness and changes in vocabulary as measured by the Wechsler Intelligence Scale for Children in children aged 5 to 11. The rate of increase in raw vocabulary scores was positively correlated with the rate of cortical thinning (i.e., greater performance improvements were associated with greater thinning) in left hemisphere regions, including the left lateral dorsal frontal and left lateral parietal regions. Lee et al. (2007) investigated the neural correlates of increases in

vocabulary within an adolescent sample, using a comparison of vocabulary and verbal fluency tests. Their presentation focused on the associations of vocabulary performance beyond those of verbal fluency, but they did not report fluency-related findings on their own. To our knowledge, studies of child and adolescent brain development and its relationship to a single, clinically established measure of verbal fluency have not been published. Additionally, purely behavioral studies of age-related changes in verbal fluency abilities have been generally focused on the changes from adulthood to old age (e.g., Steinberg et al., 2005), as opposed to focusing on developmental changes. 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 et al., 2001) studies reviewed included subjects below the age of 18.

To address these gaps, cross-sectional data on verbal fluency abilities and cortical thickness from subjects in an ongoing longitudinal study were investigated. Verbal fluency was measured by performance on the COWAT, with the expectation that age-related changes in behavioral performance would primarily reflect the maturation of executive skills. Cortical thickness was derived from high-resolution sMRI scans. Expanding upon the findings of Sowell et al. (2004), it was hypothesized that increased performance on the 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. Additionally, as the COWAT is not merely a measure of language abilities, areas previously related to increased effortful processing, rule maintenance, and sustained attention (that is, the task's executive components) were also expected to be related to increased task performance. Given prior reports on variability of structure–function relationships in younger children as compared to adolescents (Karama et al., 2009; Shaw et al., 2006), differential relationships between regional cortical thickness and COWAT performance were expected between younger and older subjects, with the anticipation that younger subjects, who would not be expected to perform as well on the task, would show a weaker structure–function relationship than older subjects, who would be expected to also have superior performance. Differential structure–function relationships were also examined between males and females, as there is evidence that adult males and females exhibit different overall verbal fluency performance abilities (Loonstra et al., 2001; Rodríguez-Aranda & Martinussen 2006). However, as the previous literature on sex differences in COWAT performance deals primarily with aging adult populations, no specific predictions were made as to how the structure–function relationship would differ between the sexes in a developmental sample.

## Methods

### Subjects

Subjects aged 9 through 23 years were recruited for a study of normative adolescent brain development. The total sample used in these analyses contained 167 healthy and typically developing children, adolescents, and young adults (74 males, 93 females) aged 9.29 – 23.96 years (Median=16.45, Mean=16.39, SD=3.80) with both usable COWAT data and viable MRI data. Two methods of contacting possible subjects aged 9 to 17 years were used. Parents of children and adolescents were contacted through a subject 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 civil service employees who may have had children in the targeted age range, and interested parents contacted our laboratory. Subjects age 18 and above were solicited

through flyers posted in the University of Minnesota community. Potential subjects 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 subjects and separately with parents and children. Just prior to scanning sessions, all subjects were re-screened for MRI contra-indications in accordance with standard MRI safety guidelines (e.g., Sawyer-Glover & Shellock, 2000).

### Protocol: Psychometric Task

Subjects completed a battery of cognitive and neuropsychological tasks as well as MRI scanning. The COWAT was selected for these analyses because of verbal fluency's historically demonstrated neural correlates, and because performance is known to increase with development (Kavé, 2006; Lezak et al., 2004; Martins et al., 2007). 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 subject 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).

The inclusion of key demographic variables has been shown to improve the statistical accuracy of analyses using psychometric test scores to predict other individual difference measures (e.g., Crawford et al., 1989). Thus, additional demographic variables were obtained through testing and self-report: IQ estimates, sex, age, and education level, all of which have all been clearly shown to be important predictors of COWAT performance (Lezak et al., 2004; Loonstra et al., 2001; Rodríguez-Aranda & Martinussen, 2006; Spreen & Strauss, 1998; Steinberg et al., 2005). Estimates of verbal intelligence (VIQ), performance intelligence (PIQ), and full scale IQ (FIQ) were obtained via the Wechsler Abbreviated Scales of Intelligence (WASI; The Psychological Corporation, 1999). Though the use of IQ as a covariate in developmental studies is sometimes unwarranted (Dennis et al., 2009), it was important to measure and account for subjects' IQs in the present study because associations between verbal fluency and IQ are abundant in the literature, such that low IQ is associated with reduced fluency (Lezak et al., 2004; Loonstra et al., 2001). Additionally, IQ has been shown to be correlated with differences in cortical thickness (Karama et al., 2009; Narr et al., 2007; Shaw et al., 2006), which underscores the need to statistically control for the variance in cortical thickness associated with IQ scores in the imaging analyses. 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. Lower educational attainment is also associated with reduced verbal fluency

(Loonstra et al., 2001); the use of educational levels in adult samples is meant to be a proxy for both socioeconomic status and intellectually enriched home environment. For developmental samples such as this, the inclusion of the subjects' education levels will be spuriously correlated with age (e.g., all 12-year-olds will be in either 6th or 7th grade). Thus, in this study, years of formal education for subjects' parents (which has also been associated with verbal fluency performance (Ardila et al., 2005; Hurks et al., 2006)) were collected instead. 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-Telsa 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 =$1.0 \times 1.0 \times 1.0$, flip angle = 7°, FOV = 256 mm).

Estimates of cortical thickness were obtained by processing the high-resolution anatomical images through the *FreeSurfer* v.4.0.5 image analysis suite, which is documented and freely available for download online (http://surfer.nmr.mgh.harvard.edu/). We followed the standard *FreeSurfer* processing pipeline. All automatically generated surface volumes were inspected for accuracy, and manual corrections using *FreeSurfer* tools were carried out as needed. The technical details of these procedures are described in prior publications (Dale & Sereno, 1993; Dale, Fischl, Sereno, 1999; Fischl et al., 1999; Fischl, Sereno, Dale, 1999; Fischl & Dale, 2000; Fischl, Liu, Dale, 2001; Fischl et al., 2002; Fischl et al., 2004a; Fischl et al., 2004b; Han et al., 2006; Jovicich et al., 2006; Segonne et al., 2004). Briefly, this processing includes motion correction and averaging of multiple volumetric T1 weighted images (when more than one is available), removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Segonne et al., 2004), automated Talairach transformation, segmentation of the subcortical white matter and deep grey matter volumetric structures (Fischl et al., 2002; Fischl et al., 2004a), intensity normalization (Sled, Zijdenbos, Evans, 1998), tessellation of the grey matter white matter boundary, automated topology correction (Fischl, Liu, Dale, 2001; Segonne et al., 2004), and surface deformation following intensity gradients to optimally place the grey/white and grey/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale & Sereno, 1993; Dale, Fischl, Sereno, 1999; Fischl & Dale, 2000). Once the cortical models are complete, a number of deformable procedures can be performed in further data processing and analysis including surface inflation (Dale, Fischl, Sereno, 1999), registration to a spherical atlas which utilized individual cortical folding patterns to match cortical geometry across subjects (Fischl, Sereno, Dale, 1999), parcellation of the cerebral cortex into units based on gyral and sulcal structure (Desikan et al., 2006; Fischl et al., 2004b), and creation of a variety of surface based data maps. This method uses both intensity and continuity information from the entire three-dimensional MR volume in segmentation and deformation procedures to produce representations of cortical thickness, calculated as the closest distance from the grey/white boundary to the grey/CSF boundary at each vertex on the tessellated surface (Fischl & Dale, 2000). The maps are created using spatial intensity gradients across tissue classes and are therefore not simply reliant on absolute signal intensity. The maps produced are not restricted to the voxel resolution of the original data thus are capable of detecting submillimeter differences between groups. Procedures for the measurement of cortical thickness have been validated against histological analysis (Rosas et al., 2002) and manual measurements (Kuperberg et al., 2003; Salat et al., 2004). *FreeSurfer* morphometric procedures have been demonstrated to show

good test-retest reliability across scanner manufacturers and across field strengths (Han et al., 2006). In all, the *FreeSurfer* processing created a thickness map for the brain of each subject 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 COWAT was chosen for this study partly because of its well-documented associations with development and aging (Kavé, 2006; Lezak et al., 2004; Loonstra et al., 2001; Martins et al., 2007; Rodríguez-Aranda & Martinussen, 2006). Assessment of the linearity of the developmental trajectory of COWAT performance in the whole sample was carried out by a series of regression analyses predicting COWAT score from linear, quadratic, and cubic age terms. *F*-tests on the change in $R^2$ between linear versus quadratic and linear versus cubic models were then carried out to determine which model provided the best fit to the data. In order to investigate possible age-related differences in the COWAT-cortical thickness relationship in later imaging analyses, the continuous age variable needed to be discretized. Thus, a series of one-way ANOVAs (IV=age group, DV=COWAT total score) were carried out, iteratively placing subjects into age bins of 10, 5, and 4 equally sized 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.

After appropriate age groupings were determined, it was important to establish which performance-relevant factors representing different domains of influence should be statistically controlled for during the planned neuroimaging analyses. 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 subjects' innate abilities to perform well on cognitive tests. Including all three IQ measures (i.e., VIQ, PIQ, & FIQ) in an expanded ANOVA would violate assumptions of collinearity and independence of observations, so which IQ measure to use was assessed by comparing correlations with COWAT total score, both with their zero-order correlations and partial controlling for the effects of the other IQ measures. Once all necessary covariates were determined, the neuroimaging analyses were carried out.

## Statistical Approach: Neuroimaging Analyses

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, as implemented in the QDEC interface. Surface volumes were smoothed at 10mm FWHM before the analyses. Neuroimaging analyses were planned for the full sample and between the age groups determined by the behavioral analyses. In addition to COWAT total score, the regression matrix included covariates derived from the behavioral analyses described above. In order to ensure that modeling a linear relationship between COWAT total score and cortical thickness was more appropriate than a nonlinear relationship, we evaluated three versions of the final regression model, with linear, quadratic, and cubic COWAT score terms. Thus, the correlation between COWAT total score and cortical thickness was

assessed across and within different age groups by regressing performance scores upon cortical thickness measures at each vertex of the reconstructed cortical surface, while accounting for the effects of key individual differences measures. 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 by running Monte Carlo-style simulations (10,000 iterations) to determine cluster-wise correction rates. All cluster-wise values were calculated to maintain an alpha value of $p \leq 0.05$ for two-tailed tests, using the *FreeSurfer* programs mri_glmfit and mri_surfcluster under the assumption that the program output was *z*-distributed. This thresholding procedure removes reliance upon more stringent parametric assumptions about the data and is similar to that employed by AFNI's AlphaSim program (http://afni.nimh.nih.gov/afni), a widely used cluster-wise thresholding tool for fMRI data. Localization and distinction of anatomical nomenclature was carried out using the Talairach Client v.2.4.2 (http://www.talairach.org/client.html) and standard *FreeSurfer* atlases (Desikan et al., 2006).

## Results

Aside from two adult males and one adult female who did not report parental education levels, all subjects had full demographic data available. Though the number of females in the sample was nominally greater, the sample did not significantly over represent females ($\chi^2(1)=2.16$, *n.s.*). Across the whole sample, there were no significant differences between females and males on COWAT total score ($t(165)=1.43$, *n.s.*), age ($t(165)=0.74$, *n.s.*), Average Parental Education ($t(162)=-0.04$, *n.s.*), FIQ ($t(165)= -0.39$, *n.s.*), VIQ ($t(165)=0.13$, *n.s.*), or PIQ ($t(165)= -0.78$, *n.s.*).

### Behavioral Analyses

Statistically significant fits to the data were obtained when regressing age terms upon COWAT score in linear ($R^2=0.396$), quadratic ($R^2=0.398$), and cubic ($R^2=0.407$) models. Tests of the change in $R^2$ were not statistically significant between the linear and quadratic models ($F(1,164)=0.86$, *n.s.*) or between the linear and cubic models ($F(1,163)=1.49$, *n.s.*). A linear developmental trajectory across the age range of the sample provided the best fit to the COWAT performance data.

Analyses to discretize the age variable for later neuroimaging analyses were conducted, first by binning subjects in age deciles. This approach 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 subjects 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., subjects 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 statistically significantly different from each other, indicating that the most robust COWAT performance difference was reflected by a median split in the age range. Thus, among the age quartile groupings, 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 statistically leveling off across ages thereafter (see Figure 1 for graphical and statistical summaries).

A follow-up 4×2 ANOVA assessing the importance of additional variables was carried out. In order to determine which of the WASI measures to use in this ANOVA, the associations of each with COWAT total score were examined. VIQ ($r$=0.203, $p$<0.05; see Figure 1, panel E) and FIQ ($r$=0.162, $p$<0.05) were related to COWAT total score, but PIQ was not ($r$=0.059, $n.s.$). Though both partial correlations were not statistically significant, when controlling for the variance related to FIQ, VIQ retains a larger association to COWAT total score ($r_{FIQ}$=0.14, $n.s.$) than does FIQ when controlling for VIQ ($r_{VIQ}$=−0.003, $n.s.$). Therefore, VIQ was chosen as the measure of intellectual capacity to account for in further analyses.

The aforementioned 4×2 ANOVA was run with quartile placement and sex entered as fixed factors, VIQ, and average parental education entered as continuous covariates, and COWAT total score as the dependent variable. There were significant main effects of age group ($F(3,154)$=39.35, $p$<0.05, partial $\eta^2$=0.434) and VIQ ($F(1,154)$=16.75, $p$<0.05, partial $\eta^2$=0.098), but there were not main effects for average parental education ($F(1,154)$=1.87, $n.s.$) or sex ($F(1,154)$=0.53, $n.s.$). There was, however, a significant interaction effect of age group × sex ($F(3,154)$=3.33, $p$<0.05, partial $\eta^2$=0.061). Follow-up Bonferroni-corrected $t$-tests within each quartile showed the interaction effect was driven by high-scoring females ($M(SD)$=37.62(8.82)) compared to lower-scoring males ($M(SD)$=28.76(7.79)) in the second quartile ($t(40)$=3.48, $p$<.0125), indicating that there may be separate verbal fluency developmental trajectories for the males and females in this age range (see Figure 1, panels A–D).

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. The behavioral analyses also indicated that sex (differentially across age groups) and VIQ were statistically important factors with respect to COWAT performance. Following the behavioral analysis results, neuroimaging analyses in the full sample and within the younger than median and older than median age groups were carried out with the goal of determining whether brain–behavior associations changed with performance maturation.

### Cortical Thickness–COWAT Main Effect

The regression analyses evaluated the effect of COWAT total score upon cortical thickness, while controlling for the effects of age group, sex, and VIQ. To test whether a linear relationship between COWAT total score and cortical thickness was more appropriate than a nonlinear relationship, we also evaluated models with quadratic and cubic COWAT score terms. Neither of the polynomial models provided a statistically significant increase in model fit across the entire cortex[1]. The full sample analysis yielded a distributed pattern of clusters showing negative linear associations between thickness and COWAT total score

---

[1]Thus, at each vertex the omnibus linear regression model reported on here was:

$$Thickness = \beta_0 + \beta_1 COWAT$$
$$+ \beta_2 AgeGroup$$
$$+ \beta_3 Sex$$
$$+ \beta_4 VIQ$$
$$+ \beta_5 (COWAT \times AgeGroup)$$
$$+ \beta_6 (COWAT \times Sex)$$
$$+ \beta_7 (COWAT \times AgeGroup \times Sex)$$

across both hemispheres. For a full map of the significant clusters and a listing of regions (in both Talairach atlas nomenclature and by Brodmann area (BA; Brodmann, 1909) covered by the clusters, see Figure 2 and Table 2. Clusters covering the posterior and anterior portions of the left hemisphere language network (Costafreda et al., 2006;Indefrey & Levelt, 2000;Vigneau et al., 2006) were identified. Bilateral prefrontal areas associated with online performance monitoring, such as anterior cingulate cortex and lateral prefrontal cortex (D'Esposito et al., 1999;MacDonald et al., 2000;Miller & Cohen, 2001), and bilateral medial parietal areas associated with increased effort and concentration during word finding, such as the precuneus (Dräger et al., 2004), were also identified. In all identified clusters, decreased cortical thickness was associated with better COWAT performance after controlling for age group, VIQ, and sex. No regions showed a relationship of greater cortical thickness with increased COWAT performance.

### Age Group Comparisons

Within the omnibus regression, the 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 cluster-wise correction. There was one cluster identified in the right hemisphere superior frontal gyrus, but its presence is likely due to chance (cluster-wise $p$=0.281; see Table 2). Relaxing the thresholding criterion to $p \leq 0.05$, uncorrected for multiple comparisons, showed modest indications of stronger COWAT–cortical thickness correlations for the older subjects within the more anterior and dorsal portions of the full sample left hemisphere map. Stronger COWAT–cortical thickness correlations in the younger subjects were seen within the more posterior portions of the full sample left hemisphere map. Performing *post hoc* FWE-corrected regression analyses separately for the younger and older groups showed a similar pattern, as well as a posterior to anterior shift with age along the right hemisphere medial wall (see Figure 3). Additional *post hoc* analyses comparing quartile 1 to quartiles 3 and 4 combined as well as quartile 2 to the older subjects showed similar subthreshold trends, with indications that the thickness–COWAT relationship may be slightly stronger in frontal regions for the older compared to younger subjects.

### Sex Comparisons

There were FWE-corrected clusters in three regions that showed significant differences between the correlation strengths for females and males: left anterior middle frontal gyrus, left superior frontal gyrus, and left precuneus. Graphical investigation of the effect indicated that males and females have strikingly opposing patterns for their COWAT performance–cortical thickness relationships within the younger ages for the lateral frontal regions (see Figure 4). The precuneus difference was driven by older males having less of a thickness–COWAT correlation compared to the other groups. No right hemisphere regions survived cluster-wise correction for this comparison, nor did any regions in a general comparison of average cortical thickness between males and females, independent of COWAT performance.

## Discussion

The primary aim of this study was to examine MRI-based structural brain correlates of verbal fluency, as measured by standardized clinical administration of the COWAT, in a sample of children, adolescents, and young adults. To accomplish this aim, COWAT performance was measured outside of the scanner and associated with cortical thickness using *FreeSurfer*'s analytic suite. Performance patterns from the behavioral assessment were used to inform the analysis of the neuroimaging data.

### Behavioral Finding

The developmental trajectory of COWAT performance for all subjects was best summarized by a linear regression across the sample's full age range. The strongest predictors of COWAT performance were age and VIQ. To further explore this finding, the sample was classified by age into discrete groups. When COWAT performance was examined across groups, there was a statistically reliable performance difference between subjects above and below the median age of the sample. There was a smaller, but statistically reliable, effect suggesting that the developmental trajectory may have been influenced by a sex × age interaction, such that females in this sample approached the peak performance range earlier than males did. Overall, these findings suggest that behavioral maturation of COWAT performance is achieved around the age of 16, consistent with other large-scale studies of cognitive development that emphasize executive functions in samples of similar ages (Steinberg et al., 2009). This age demarcation was used to examine correlates of COWAT performance before versus after this point of performance maturation.

### Imaging: Developmental Findings and Implications

On the whole, structural brain development studies indicate that grey matter proliferates through early childhood, and as adolescence begins, grey matter decreases in thickness (Shaw et al., 2006; Sowell et al., 2004) and volume (Giedd et al., 1999; Gogtay et al., 2004; Lenroot et al., 2007), reflecting shifts from overabundant neuropil to mature, efficient cortical networks. 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 subjects. A primary hypothesis was that older subjects would show a cortical thickness–performance association primarily driven by brain regions that participate in executive functions. 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 subjects when the effects of sex and VIQ were controlled and statistical test results were corrected for family-wise error inflation. Rather, as will be discussed below, the cortical thickness–behavior associations that were observed are equivalent for each age group and more strongly reflect areas that participate in basic language skills as opposed to areas that participate in executive function.

Therefore, the strength of the relationship between thinner cortex and improved COWAT performance in regions that comprise what is commonly considered to be the language network is equivalent across this sample's age range, regardless of whether performance has matured or if it is still improving. In this respect, the imaging findings do not provide a direct correspondence to the behavioral findings. However, this pattern was the result of analyses conducted to satisfy the need for a discrete grouping variable in the QDEC analyses, and the regression analyses of age polynomials predicting COWAT score did not indicate a nonlinear trajectory for COWAT scores for the full sample across the age range. The 4×2 ANOVA in the behavioral analyses demonstrated the presence of separate developmental COWAT trajectories for males and females, which was supported by the imaging analyses (discussed below).

Two questions that arise from this pattern of increasing performance with decreasing cortical thickness concern the nature of the mechanism that leads to cortical thinning in support of task performance as well as why there are not strong differences in the brain–behavior associations before versus after the point of performance maturation. With respect to the first issue, it is broadly accepted that during adolescence, there is maturational thinning across the cortex (Giedd et al., 1999; Gogtay et al., 2004; Sowell et al., 2004), a

phenomenon that is typically attributed to synaptic pruning. In the paradigm of "use it or lose it," pruning results from the loss of non-preferred cortical connections in favor of retaining those functional connections that support necessary and frequently used skills. Another interpretation may be that the observed reduction in measured grey matter instead reflects an increase in myelination such that white matter growth encroaches upon what was previously classified as grey matter (Paus, 2005). Thus, cortical thinning during this active phase of brain development could reflect either synaptic pruning or increased myelination.

The biological mechanism through which a thinner cortex confers better task performance is not clear, although several speculations can be offered. If cortical thinning reflects synaptic pruning, then pruning may occur relatively early for language-based abilities, as opposed to higher order functions. The functional benefit would be a tightly honed neural system that is impervious to "neural interference," a context that would benefit COWAT performance by facilitating access to one's store of verbal information as well as the two main strategies of enhanced fluency: 1) focus within a given semantic category and 2) switching between categories (Troyer et al., 1997). In contrast, within the age range of this study, reports of cortical thickness associations with IQ have shown a pattern opposite to our findings, with thicker cortex associated with better performance (Karama et al., 2009; Narr et al., 2007; Shaw et al., 2006). However, these studies used an overall measure of general intelligence (e.g., Full-Scale IQ, estimates of $g$), and do not explore the individual subtests that contribute to this measure. To the extent that VIQ taps crystallized intelligence and PIQ taps fluid intelligence (e.g., Benson et al., 2010), it is possible that tasks most closely related to VIQ, such as COWAT performance in this study, will be highly related to brain-based endophenotypic measures that represent experience-based acquisition and solidification of knowledge and skills. This study provides initial evidence on this point.

A thicker cortex may enhance general intellectual functions by encouraging activity in associative networks both within and across regions in a manner that would be detrimental to performance on a task such as the COWAT. Whether or not a thicker versus thinner cortex then confers enhanced abilities to perform tests of fluid intelligence is a hypothesis to be pursued in future work. Interpretations are not straightforward, given that the relationships between cortical thickness and some aspects cognitive abilities (i.e., general intelligence) change between early and late childhood (Shaw et al., 2006), the functional implications of thinning cortex become detrimental with increasing age (Dickerson et al., 2008), and because the mechanisms of brain maturation include more than just pruning within grey matter (Ashtari et al., 2007; Paus et al., 2001). Insofar as myelination represents increases in neuronal signaling speed and efficiency, decreases in tissue classified as grey matter reflect a process that also results in functional networks that better support necessary skills.

Given our findings, it would be expected that studies of children and adolescents with language disorders would show a blunted pattern of thinning in language network regions, given their lower levels of linguistic achievement. Experiments using fMRI have shown that "compensatory" processes in regions outside the left hemisphere language network are evident in children with Reading Disorder (Shaywitz et al., 2002) and that such differences are reduced with treatment (Richards & Berninger, 2008). We would then expect untreated subjects to show patterns of thinning associated with language tasks to appear in those right hemisphere regions that compensate for reduced involvement of left hemisphere language regions, whereas, over time, these associations would be attenuated in subjects who have received successful interventions.

Because COWAT performance is highly correlated with age, the expectation would be that task-specific regional thinning occurs as a developmental process in tandem with increased

age and relevant life experience. Our findings suggest, in contrast, that the network that supports stronger task performance has "thinned out" at a relatively early age (perhaps by 9–12 years), which may represent the end of the "sensitive period" for language acquisition in child development (Sakai, 2005). This pattern is surprising given the more extended pace at which performance develops but nevertheless, it coheres with other studies of verbal skill development in even younger children (Sowell et al., 2004). However, in contrast to Sowell's study, this report focused on a specific aspect of language ability (fluency) independent of general verbal ability, which was the process examined in that report. Moreover, the current study focuses on a broader age range that extends into adolescence and young adulthood. Thus, our report and that of Sowell et al. cohere in suggesting that gray matter changes that support COWAT performance are particularly prominent in areas that reflect basic language processing and that this aspect of the brain structure/behavior association is established in childhood, despite the fact that the task requires a number of executive skills that mature later.

The fact that the age group comparisons did not yield associations between cortical thickness in frontal regions and behavior, especially in the older group, was surprising to us. Although the omnibus imaging analysis indicated that there was no reliable difference between the strength of the COWAT performance–cortical thickness correlation between the older and younger subjects, some age related distinctions were suggested by *post-hoc* analyses. Looking at the thickness–COWAT correlation within the median split subgroups in isolation showed that the younger group had a nominally stronger relationship in left supramarginal and angular gyri (see Figure 3, panel A), and the older group had nominally stronger relationships in the left frontal regions of the language network (see Figure 3, panel B). Additionally, the parietal to frontal shift with age is illustrated if clusterwise thresholding is performed on the difference between the unthresholded $p$-value maps from each age group (see Figure 3, panel C). These shifts in the localization of the cluster maps may indicate regionally differential COWAT–cortical thickness relationships for younger and older subjects. It could be that more mature cortex in phonological retrieval and concept preparation regions (i.e., left temporal cortex; Indefrey & Levelt, 2000) is an indicator of increased exposure to or achievement with linguistic information for younger children, but that the older subjects have relatively comparable levels of maturation in these areas, and their performances are differentiated by greater mental agility (not necessarily linguistic in nature) that is indexed by the correlations across frontal and parietal association cortex. Lesions in left dorsolateral and superior prefrontal lesions (e.g., BAs 9, 46, & 47) impair the ability to engage phonemic cueing and effectively switch among phonemic clustering strategies (Eskes et al., 2003;Troyer et al., 1998), whereas temporal lobe lesions impair the use of semantic clustering strategies (Troyer et al., 1998). Thus, a posterior to anterior shift in the strength of the COWAT–thickness relationship could also reflect the cognitive heterogeneity of the COWAT, such that younger children and adolescents rely more upon verbal store and semantic clustering and older adolescents and adults engage in more flexible strategies of clustering and switching (Troyer et al., 1997). The lack of clear developmental effects could also be explained if the suggested switch from posterior to anterior reliance occurs somewhat later than the top of our age range, and the present sample is too young to robustly display the effect. However, these possible explanations must be tempered by the fact that the omnibus regression's between-groups comparison is presented at a lax statistical threshold (see Figure 3, panel D), the separately run FWE-corrected regression analyses do not directly compare the two groups to each other, and the FWE-corrected comparison of $p$-values is not as statistically sound as performing the comparison in the omnibus regression. Great caution is warranted when interpreting the differential importance of these regions to COWAT performance across development[2].

Prior investigations of developmental changes in the functional language network in children indicates that the presence of left-lateralized selective response to language is in place even in early infancy (Dehaene-Lambertz et al., 2002; Peña et al., 2003), and that, while the fMRI signal may be less focal in younger subjects, the regions of activation in children as young as 8 years old are the same as those in adults (Gaillard et al., 2000). Given the centrality of language to human development and the evidence for very early ontological differentiation of the language network, it is possible that the basic cortical structural organization of the language network is in place and concomitant maturational pruning processes have begun before the youngest age in this sample. Alternatively, as discussed further below, it may be that cortical thickness is not the optimal aspect of brain structure through which to discern *developmental differences* in structure–function relations, particularly as they apply to executive functions, with respect to verbal fluency.

Given the lack of reliable differences in the cortical thickness–verbal fluency association between the older and younger cohorts in this sample, analyses of the full sample may be most informative regarding how cortical thickness is related to verbal fluency. Indeed, these findings are striking in relation to the regional specificity of the COWAT–cortical thickness associations.

## Imaging: Main Effect

In general, 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 (Costafreda et al., 2006; Eskes et al., 2003; Indefrey & Levelt, 2000; Troyer et al., 1998; Vigneau et al., 2006). Language-related regions included bilateral superior and middle temporal gyrus, left inferior parietal cortex (including supramarginal gyrus and angular gyrus), left pars opercularis, bilateral pars triangularis, bilateral anterior middle frontal gyrus, bilateral fusiform gyrus, and bilateral lateral occipital gyrus.

The strongest correlations between COWAT performances and cortical thicknesses were found in the temporal lobes and temporal-parietal junction (see Table 2). This pattern may indicate that greater crystallized intelligence, as represented by more mature word representation storage regions (Frith et al., 1991), is of greater benefit to COWAT performance than is reaching a similar maturational extent in regions engaged in cognitive control (e.g., lateral prefrontal cortex). That is to say, given a large store of verbal knowledge, cognitive flexibility is perhaps less important to COWAT performance. Indeed, it has been suggested that the main cognitive components of COWAT performance are general word knowledge and auditory attention, with lesser contributions from executive functioning abilities (Ruff et al., 1997). However, there is also strong behavioral evidence for the centrality of employing clustering and switching behaviors to enhance fluency performance (Troyer et al., 1997), which have been shown to be dependent upon posterior temporal (clustering) and lateral prefrontal (switching) regions overlapping with regions indentified in these analyses (Troyer et al., 1998). Future investigations of the structural correlates of the discrete cognitive components of verbal fluency are warranted, via contrast and comparison of findings reported for a variety of linguistic tasks (e.g., Indefrey & Levelt, 2000;Troyer et al., 1997).

[2]Following the suggestion that age 12 represents the end of a sensitive period of language acquisition, *post hoc* investigations of the difference in the COWAT–thickness relationship between the 9.5 to 12 year-olds and the rest of the sample did show a statistically significant effect in the left hemisphere superior and rostral middle frontal gyrus, such that the association in this region was stronger for the older groups than the younger group (data not shown). The size and location of this cluster aligned with the trend-level developmental findings already discussed in this manuscript, as illustrated in Figure 3, panels B and C.

Another pattern that we observed is that the correlation values surrounding Wernicke's area were stronger than those surrounding Broca's area. This pattern would be expected in neural correlates of a task that relies more heavily on phonological code retrieval than upon phonological encoding (Indefrey & Levelt, 2000). Additionally, the presence of clusters across the fusiform gyri would be predicted for a task relying on phonological retrieval (Price & Friston, 1997). The cluster map that covered Broca's area included portions of both pars opercularis and pars triangularis. A recent meta-analysis of the specific roles of these sections of the left inferior frontal gyrus in verbal fluency indicates that BA 44 (i.e., pars opercularis) may subserve phonological fluency, whereas the BA 45 (i.e., pars triangularis) may be more involved in semantic fluency (Costafreda et al., 2006). The current results provide somewhat muted support of this distinction. Both areas are included in the coverage of the identified cluster, but the shape of the cluster is concentrated around two foci, with pars opercularis strongly represented and pars triangularis contacted by an offshoot of the focus over middle frontal gyrus (BA46). In the distinction drawn through meta-analysis by Costafreda et al. (2006) and reinforced by fMRI experimentation by Heim et al. (2008), this pattern of results would be predicted, as the COWAT is a phonologically based verbal fluency task. 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. The involvement of this region aligns with the Indefrey and Levelt (2000) model of language production, in the service of response articulation.

Areas that have been related to increased effortful processing, rule maintenance, sustained attention, performance monitoring, and phonemic clustering and switching strategies were seen in both hemispheres. Lateral surface clusters included bilateral middle and superior frontal gyrus and left superior parietal gyrus. Bilateral medial surface clusters were predominately situated among the precuneus, posterior cingulate, and anterior cingulate; right hemisphere medial orbitofrontal and superior frontal gyrus clusters were also found. The negative correlational relationship between verbal fluency abilities and regional cortical thickness suggests that increasing COWAT performance is associated with a cortical network underlying language-related and, to a lesser extent, executive functioning abilities.

### Imaging: 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 finding may be driven by sex differences in the regional timing of brain maturation (Lenroot et al., 2007; Sowell et al., 2004) 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 this sample had an extra year, on average, of maturation in a region that subserves major components of COWAT performance and may account for the females' earlier rise in performance trajectory. This finding could reflect girls' increased use of phonemic/semantic cluster switching strategies (Koren et al., 2005), and may be a developmental source of an often-described "female advantage" in verbal fluency tasks (Sincoff and Sternberg, 1998). However, functional scanning studies of verbal fluency (as opposed to other language tasks) generally do not show large sex differences in activations (cf. Plante et al., 2006), so replication of this finding would underscore the need for structural investigations of COWAT performance.

## 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; Narr et al., 2007; 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 (i.e., the observed relation is driven by a general intelligence factor). However, these analyses included VIQ as a covariate, in order to explicitly partial out the effect of intellectual capacity.

A limitation inherent in the methodology of cortical surface reconstruction via *FreeSurfer* is 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, though, as shown by reexamination of Broca's seminal aphasia patients, dysfluent speech is related to damage beyond the depths of the cortical ribbon (Dronkers et al., 2007). In a recent study using voxel-based morphometry (a whole-brain capable methodology) that similarly correlated verbal fluency scores with estimates of grey matter density, Grogan et al. (2009) reported positive associations between behavior and grey matter density in the bilateral areas of cerebellum, caudate, and pre-supplementary motor areas (medial BA 6). We echo those findings[3] with a cluster in right hemisphere superior frontal cortex that crosses BA 6, but the current analyses are unable to speak to the subcortical and cerebellar findings. Indeed, for both verbal fluency and executive functioning, central roles for subcortical structures have been proposed (Eickhoff et al., 2009; Indefrey & Levelt, 2000; Lieberman, 2002), as have important roles for regions of the cerebellum (Ackermann et al., 2007; Indefrey & Levelt, 2000; Lieberman, 2002) and the corpus callosum (Narberhaus et al., 2008; Wozniak et al., 2007). It should be clearly stated that the analyses described indicate that the cortical regions identified may be necessary for the performance of verbal fluency tasks, but by no means do these analyses indicate that they are sufficient for doing so.

The absence of clear age effects in these neuroimaging analyses begs the questions of whether or not *maturational patterns* within the language network across adolescence may be better indexed by other measures. Behaviorally, scoring COWAT performance using more expansive qualitative methods may be fruitful. For instance, age-related differences have been demonstrated in metrics of rate of speech (Martins et al., 2007) and for the size of and frequency of switching among phonemic and semantic clusters (Sauzéon et al., 2004; Troyer et al., 1997). Task-related effects have been shown for differences in output during the initial 15 seconds versus the latter 45 seconds of phonemic and semantic fluency trials (Hurks et al., 2006). Future investigations of the structural correlates of such measures could isolate regions within the language network that differentially assist with the various cognitive components of verbal fluency and may be helpful in differentiating components of the task that are more or less "executive" in nature.

It should also be emphasized that the measurement of cortical thickness does not directly reflect functional activation during task performance. Though applicable motion correction and vocal monitoring techniques are emerging and such fMRI studies indicate activation of regions also found in the present study (cf. Basho et al., 2007; Birn et al., 2010; Heim et al., 2008), collection of COWAT data using fully standardized administration practices during

---

[3]Voxel-based morphometry (VBM) uses grey matter "density," or an estimate of the proportion of a voxel that is taken up by grey matter, as the dependent variable. Variance in the density measure could be due changes in the folding patterns of the cortex, independent of changes in cortical thickness or absolute amount of tissue present in the cortical ribbon for a particular anatomical region. Thus, verbal fluency's positive associations with VBM density and negative associations with thickness are not necessarily at odds.

fMRI scanning has not yet been reported. Functional differences within the network, such as a refinement of localization in the fMRI BOLD signal (Durston et al., 2006; Gaillard et al., 2000), may be another avenue for exploring developmental changes in verbal fluency.

Additional structure–function methods should be pursued in order to better understand the neural substrates of this standardized and widely used neuropsychological test. The flexibility and fluidity needed to perform may be better captured by multimodal techniques assessing white and grey matter development concurrently (e.g., diffusion tensor imaging (DTI) along with sMRI) in order to describe both the cortical maturation and fiber tract integrity of the language network. Future studies utilizing techniques such as probabilistic tractography via DTI 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 and how these measures interact with cortical thickness. Such analyses could also shed light upon the nature of the relationship of regions across the hemispheres and could bring interpretive clarity to the negative COWAT–cortical thickness correlations observed here. Early reports with small sample sizes for DTI data have successfully demonstrated such a relationship between areas involved in vocabulary acquisition (Lee et al., 2007).

## Conclusion

In sum, these analyses supported the utility of examining specific structure–function correlations using metrics, such as cortical thickness, in relation to behavioral tasks that have well-established associations with brain functioning. This study presents evidence of brain–behavior relationships that have specific clinical relevance, as the COWAT is a test used widely in neuropsychological evaluations. Differences in COWAT performances in clinical populations can be compared to these results in service of generating hypotheses about underlying structural brain differences.

Behaviorally, a significant difference in COWAT performance was found between the younger and older subjects in this sample. In a linearly increasing fashion, older subjects achieved higher scores than younger subjects. After cluster-wise correction, 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. However, across the full sample, significant correlations between improved COWAT performance and decreased cortical thickness were found in regions corresponding to left hemisphere language regions, including areas surrounding the traditional language hubs of Wernicke's and Broca's areas, as well as temporal cortex related to auditory processing, somatosensory cortex related to the operation of the lips, tongue, and mouth, and frontal and parietal regions related to effortful processing and performance monitoring. 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. Thus, a major implication of these findings is that cortical thickness, in regions that comprise language and executive functioning networks, likely begins maturation toward adult-like cortical organization and processing efficiency (as reflected by thinning) by middle childhood (i.e., by the lower bound of our sample, age 9). The changes in cortical thickness that support the Controlled Oral Word Association Task are apparent by middle childhood, but regional timing of maturational completion is different between males and females, consistent with other studies of frontal lobe development. This study provides an extension of prior work indicating that increases in verbal abilities are related to regionally specific decreases in cortical thickness. Given that other reports demonstrate relationships between global

intellectual functioning and increasing cortical thickness in adolescent development, it is clear that additional imaging methodologies, such as DTI and fMRI, will be needed to fully characterize the complex relationship between increases in cognitive and intellectual faculties an developmental changes in indexes of brain maturation.

## Acknowledgments

## References

Ackermann H, Mathiak K, Riecker A. The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. Cerebellum. 2007; 6(3):202–213. [PubMed: 17786816]

Alexander MP, Benson DF, Stuss DT. Frontal lobes and language. Brain and Language. 1989; 37(4): 656–691. [PubMed: 2479448]

Ardila A, Rosselli M, Matute E, Guajardo S. The influence of the parents' educational level on the development of executive functions. Developmental Neuropsychology. 2005; 28(1):539–560. [PubMed: 15992255]

Ashtari M, Cervellione KL, Hasan KM, Wu J, McIlree C, Kester H, Ardekani BA, Roofeh D, Szeszko PR, Kumra S. White matter development during late adolescence in healthy males: a cross-sectional diffusion tensor imaging study. NeuroImage. 2007; 35:501–535. [PubMed: 17258911]

Basho S, Palmer ED, Rubio MA, Wulfeck B, Muller RA. Effects of generation mode in fMRI adaptations of semantic fluency: Paced production and overt speech. Neuropsychologia. 2007; 45(8):1697–1706. [PubMed: 17292926]

Benson N, Hulac DM, Kranzler JH. Independent examination of the Wechsler Adult Intelligence Scale —Fourth Edition (WAIS–IV): what does the WAIS–IV measure? Psychological Assessment. 2010; 22:121–130. [PubMed: 20230158]

Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebral Cortex. 2009; 19(12):2767–2796. [PubMed: 19329570]

Birn RM, Kenworthy L, Case L, Caravella R, Jones TB, Bandettini PA, Martin A. Neural systems supporting lexical search guided by letter and semantic category cues: A self-paced overt response fMRI study of verbal fluency. NeuroImage. 2010; 49(1):1099–1107. [PubMed: 19632335]

Bonilha L, Molnar C, Horner MD, Anderson B, Forster L, George MS, Nahas Z. Neurocognitive deficits and prefrontal cortical atrophy in patients with schizophrenia. Schizophrenia Research. 2008; 101(1–3):142–151. [PubMed: 18164594]

Broca PP. Loss of speech, chronic softening and partial destruction of the anterior left lobe of the brain. Bulletin De La Société Anthropologique. 1861; 2:235–238.

Brodmann, K. Verleichende lokalisations lehre der grosshirnrinde in ihren prinzipien dar gestellt auf grund des zellenbanes. J.A. Barth; Leipzig: 1909.

Costafreda SG, Fu CH, Lee L, Everitt B, Brammer MJ, David AS. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. Human Brain Mapping. 2006; 27(10):799–810. [PubMed: 16511886]

Crawford JR, Stewart LE, Parker DM, Besson AO, Cochrane RHB. Estimation of premorbid intelligence: Combining psychometric and demographic approaches improves predictive accuracy. Personality and Individual Differences. 1989; 10(7):793–796.

Dale AM, Sereno MI. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. Journal of Cognitive Neuroscience. 1993; 5:162–176.

Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. segmentation and surface reconstruction. NeuroImage. 1999; 9(2):179–194. [PubMed: 9931268]

Damasio AR, Geschwind N. The neural basis of language. Annual Review of Neuroscience. 1984; 7:127–147.

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. Functional neuroimaging of speech perception in infants. Science. 2002; 298(5600):2013–2015. [PubMed: 12471265]

Dennis M, Francis DJ, Cirino PT, Schachar R, Barnes MA, Fletcher JM. Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. Journal of the International Neuropsychological Society. 2009; 15(3):331–343. [PubMed: 19402919]

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006; 31(3):968–980. [PubMed: 16530430]

D'Esposito M, Postle BR, Ballard D, Lease J. Maintenance versus manipulation of information held in working memory: An event-related fMRI study. Brain and Cognition. 1999; 41(1):66–86. [PubMed: 10536086]

Dickerson BC, Fenstermacher E, Salat DH, Wolk DA, Maguire RP, Desikan R, Pacheco J, Quinn BT, Van der KA, Greve DN, Blacker D, Albert MS, Killiany RJ, Fischl B. Detection of cortical thickness correlates of cognitive performance: reliability across MRI scan sessions, scanners, and field strengths. NeuroImage. 2008; 39:10–18. [PubMed: 17942325]

Dräger B, Jansen A, Bruchmann S, Forster AF, Pleger B, Zwitserlood P, Knecht S. How does the brain accommodate to increased task difficulty in word finding? A functional MRI study. NeuroImage. 2004; 23(3):1152–1160. [PubMed: 15528114]

Dronkers NF, Plaisant O, Iba-Zizen MT, Cabanis EA. Paul broca's historic cases: High resolution MR imaging of the brains of leborgne and lelong. Brain. 2007; 130(Pt 5):1432–1441. [PubMed: 17405763]

Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition. 2004; 92(1–2):145–177. [PubMed: 15037129]

Durston S, Davidson MC, Tottenham N, Galvan A, Spicer J, Fossella JA, Casey BJ. A shift from diffuse to focal cortical activity with development. Developmental Science. 2006; 9(1):1–8. [PubMed: 16445387]

Eickhoff SB, Heim S, Zilles K, Amunts K. A systems perspective on the effective connectivity of overt speech production. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences. 2009; 367(1896):2399–2421.

Eskes GA, Szostak C, Stuss DT. Role of the frontal lobes in implicit and explicit retrieval tasks. Cortex. 2003; 39(4–5):847–869. [PubMed: 14584556]

Fischl B, Dale AM. 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. 2000; 97(20):11050–11055. [PubMed: 10984517]

Fischl B, Liu A, Dale AM. Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Transactions on Medical Imaging. 2001; 20(1):70–80. [PubMed: 11293693]

Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage. 1999; 9(2):195–207. [PubMed: 9931269]

Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. Human Brain Mapping. 1999; 8(4):272–284. [PubMed: 10619420]

Fischl B, Salat DH, van der Kouwe AJ, Makris N, Segonne F, Quinn BT, Dale AM. Sequence-independent segmentation of magnetic resonance images. NeuroImage. 2004a; 23(Suppl 1):S69–84. [PubMed: 15501102]

Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, et al. Automatically parcellating the human cerebral cortex. Cerebral Cortex. 2004b; 14(1):11–22. [PubMed: 14654453]

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, et al. Whole brain segmentation: Automated labeling of

neuroanatomical structures in the human brain. Neuron. 2002; 33(3):341–355. [PubMed: 11832223]

Frith CD, Friston KJ, Liddle PF, Frackowiak RS. A PET study of word finding. Neuropsychologia. 1991; 29:1137–1148. [PubMed: 1791928]

Gaillard WD, Hertz-Pannier L, Mott SH, Barnett AS, LeBihan D, Theodore WH. Functional anatomy of cognitive development: FMRI of verbal fluency in children and adults. Neurology. 2000; 54(1): 180–185. [PubMed: 10636145]

Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, Paus T, Evans AC, Rapoport JL. Brain development during childhood and adolescence: A longitudinal MRI study. Nature Neuroscience. 1999; 2(10):861–863.

Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, Nugent TF 3rd, Herman DH, Clasen LS, Toga AW, et al. Dynamic mapping of human cortical development during childhood through early adulthood. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(21):8174–8179. [PubMed: 15148381]

Grogan A, Green DW, Ali N, Crinion JT, Price CJ. Structural correlates of semantic and phonemic fluency ability in first and second languages. Cerebral Cortex. 2009; 19(11):2690–2698. [PubMed: 19293396]

Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, Busa E, Pacheco J, Albert M, Killiany R, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. NeuroImage. 2006; 32(1):180–194. [PubMed: 16651008]

Heim S, Eickhoff SB, Amunts K. Specialisation in broca's region for semantic, phonological, and syntactic fluency? NeuroImage. 2008; 40(3):1362–1368. [PubMed: 18296070]

Hurks PPM, Vles JSH, Hendriksen JGM, Kalff AC, Feron FJM, Kroes M, van Zeben TMCB, Steyaert J, Jolles J. Semantic category and fluency versus initial letter fluency over 60 seconds as a measure of automatic and controlled processing in healthy school-aged children. Journal of Clinical and Experimental Neuropsychology. 2006; 28(5):684–695. [PubMed: 16723317]

Indefrey, P.; Levelt, WJM. The neural correlates of language production. In: Gazzaniga, M., editor. The new cognitive neurosciences. MIT Press; Cambridge, MA: 2000.

Jovicich J, Czanner S, Greve D, Haley E, van der Kouwe A, Gollub R, Kennedy D, Schmitt F, Brown G, Macfall J, et al. Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. NeuroImage. 2006; 30(2):436–443. [PubMed: 16300968]

Karama S, Ad-Dab'bagh Y, Haier RJ, Deary IJ, Lyttelton OC, Lepage C, Evans AC. the Brain Development Cooperative Group. Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. Intelligence. 2009; 37(2): 145–155. [PubMed: 20161325]

Kavé G. The development of naming and word fluency: Evidence from Hebrew-speaking children between ages 8 and 17. Developmental Neuropsychology. 2006; 29(3):493–508. [PubMed: 16671864]

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N. 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. 1997; 36(7):980–988. [PubMed: 9204677]

Koren R, Kofman O, Berger A. Analysis of word clustering in verbal fluency of school-aged children. Archives of Clinical Neuropsychology. 2005; 20:1087–1104. [PubMed: 16125896]

Kuperberg GR, Broome MR, McGuire PK, David AS, Eddy M, Ozawa F, Goff D, West WC, Williams SC, van der Kouwe AJ, et al. Regionally localized thinning of the cerebral cortex in schizophrenia. Archives of General Psychiatry. 2003; 60(9):878–888. [PubMed: 12963669]

Lee H, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, Glensman J, Pitcher K, Crinion J, Mechelli A, Frackowiak RS, et al. Anatomical traces of vocabulary acquisition in the adolescent brain. The Journal of Neuroscience. 2007; 27(5):1184–1189. [PubMed: 17267574]

Lenroot RK, Gogtay N, Greenstein DK, Wells EM, Wallace GL, Clasen LS, Blumenthal JD, Lerch J, Zijdenbos AP, Evans AC, et al. Sexual dimorphism of brain developmental trajectories during childhood and adolescence. NeuroImage. 2007; 36(4):1065–1073. [PubMed: 17513132]

Lezak, MD.; Howieson, DB.; Loring, DW. Neuropsychological assessment. 4. Oxford University Press; New York: 2004.

Lieberman P. On the nature and evolution of the neural bases of human language. American Journal of Physical Anthropology Suppl. 2002; 35:36–62.

Loonstra AS, Tarlow AR, Sellers AH. COWAT metanorms across age, education, and gender. Applied Neuropsychology. 2001; 8(3):161–166. [PubMed: 11686651]

MacDonald AW 3rd, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science. 2000; 288(5472):1835–1838. [PubMed: 10846167]

Martins IP, Vieira R, Loureiro C, Santos ME. Speech rate and fluency in children and adolescents. Child Neuropsychology. 2007; 13:319–332. [PubMed: 17564849]

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annual Review of Neuroscience. 2001; 24:167–202.

Narberhaus A, Segarra D, Caldú X, Giménez M, Pueyo R, Botet F, Junqué C. Corpus callosum and prefrontal functions in adolescents with history of very preterm birth. Neuropsychologia. 2008; 46:111–116. [PubMed: 17897687]

Narr KL, Woods RP, Thompson PM, Szeszko P, Robinson D, Dimtcheva T, Gurbani M, Toga AW, Bilder RM. Relationships between IQ and regional cortical gray matter thickness in healthy adults. Cerebral Cortex. 2007; 17(9):2163–2171. [PubMed: 17118969]

Newman LM, Trivedi MA, Bendlin BB, Ries ML, Johnson SC. The relationship between gray matter morphometry and neuropsychological performance in a large sample of cognitively healthy adults. Brain Imaging and Behavior. 2007; 1(1–2):3–10. [PubMed: 18836573]

Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia. 1971; 9(1):97–113. [PubMed: 5146491]

Paus T. Mapping brain maturation and cognitive development during adolescence. Trends in Cognitive Sciences. 2005; 9:60–68. [PubMed: 15668098]

Paus T, Collins DL, Evans AC, Leonard G, Pike B, Zijdenbos A. Maturation of white matter in the human brain: a review of magnetic resonance studies. Brain Research Bulletin. 2001; 54:255–266. [PubMed: 11287130]

Peña M, Maki A, Kovacic D, Dehaene-Lambertz G, Koizumi H, Bouquet F, Mehler J. Sounds and silence: An optical topography study of language recognition at birth. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100(20):11702–11705. [PubMed: 14500906]

Plante E, Schmithorst VJ, Holland SK, Byars AW. Sex differences in the activation of language cortex during childhood. Neuropsychologia. 2006; 44(7):1210–1221. [PubMed: 16303148]

Price CJ, Friston KJ. Cognitive conjunction: A new approach to brain activation experiments. NeuroImage. 1997; 5:261–270. [PubMed: 9345555]

Price CJ. The anatomy of language: A review of 100 fMRI studies published in 2009. Annals of the New York Academy of Sciences. 2010; 1191(1):62–88. [PubMed: 20392276]

Price CJ. The anatomy of language: Contributions from functional neuroimaging. Journal of Anatomy. 2000; 197(Pt 3):335–359. [PubMed: 11117622]

Richards TL, Berninger VW. Abnormal fMRI connectivity in children with dyslexia during a phoneme task: before but not after treatment 1. Journal of Neurolinguistics. 2008; 21:294–304. [PubMed: 19079567]

Rodríguez-Aranda C, Martinussen M. Age-related differences in performance of phonemic verbal fluency measured by controlled oral word association task (COWAT): A meta-analytic study. Developmental Neuropsychology. 2006; 30(2):697–717. [PubMed: 16995832]

Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH, van der Kouwe A, Jenkins BG, Dale AM, Fischl B. Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology. 2002; 58(5):695–701. [PubMed: 11889230]

Ruff RM, Light RH, Parker SB, Levin HS. The psychological construct of word fluency. Brain and Language. 1997; 57(3):394–405. [PubMed: 9126423]

Sakai KL. Language acquisition and brain development. Science. 2005; 310(5749):815–819. [PubMed: 16272114]

Salat DH, Buckner RL, Snyder AZ, Greve DN, Desikan RS, Busa E, Morris JC, Dale AM, Fischl B. Thinning of the cerebral cortex in aging. Cerebral Cortex. 2004; 14(7):721–730. [PubMed: 15054051]

Sauzéon H, Lestage P, Raboutet C, N'Kaoua B, Claverie B. Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. Brain and Language. 2004; 89:192–202. [PubMed: 15010250]

Sawyer-Glover AM, Shellock FG. Pre-MRI procedure screening: Recommendations and safety considerations for biomedical implants and devices. Journal of Magnetic Resonance Imaging. 2000; 12(1):92–106. [PubMed: 10931569]

Segonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, Fischl B. A hybrid approach to the skull stripping problem in MRI. NeuroImage. 2004; 22(3):1060–1075. [PubMed: 15219578]

Shaw P, Greenstein D, Lerch J, Clasen L, Lenroot R, Gogtay N, Evans A, Rapoport J, Giedd J. Intellectual ability and cortical development in children and adolescents. Nature. 2006; 440(7084): 676–679. [PubMed: 16572172]

Shaywitz B, Shaywitz S, Pugh K, Mencl E, Fulbright R, Skudlarski P, Constable T, Marchione K, Fletcher J, Lyon R, et al. Disruption of posterior brain systems for reading in children with developmental dyslexia. Biological Psychiatry. 2002; 52:101–110. [PubMed: 12114001]

Sincoff JB, Sternberg RJ. Development of verbal fluency abilities and strategies in elementary-school-aged children. Developmental Psychology. 1998; 24(5):646–653.

Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Transactions on Medical Imaging. 1998; 17(1):87–97. [PubMed: 9617910]

Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, Toga AW. Longitudinal mapping of cortical thickness and brain growth in normal children. The Journal of Neuroscience. 2004; 24(38):8223–8231. [PubMed: 15385605]

Spreen, O.; Strauss, E. A compendium of neuropsychological tests: Administration, norms, and commentary. 2. Oxford University Press; New York: 1998.

Steinberg L, Graham S, O'Brien L, Woolard J, Cauffman E, Banich M. Age differences in future orientation and delay discounting. Child Development. 2009; 80:28–44. [PubMed: 19236391]

Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. Mayo's older americans normative studies: Age- and IQ-adjusted norms for the trail-making test, the stroop test, and MAE controlled oral word association test. The Clinical Neuropsychologist. 2005; 19(3–4):329–377. [PubMed: 16120535]

The Psychological Corporation. Wechsler abbreviated scale of intelligence (WASI) manual. The Psychological Corporation; San Antonio, TX: 1999.

Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology. 1997; 11(1):138–146. [PubMed: 9055277]

Troyer AK, Moscovitch M, Winocur G, Alexander MP, Stuss D. Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. Neuropsychologia. 1998; 36(6): 499–504. [PubMed: 9705059]

Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, Mazoyer B, Tzourio-Mazoyer N. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. NeuroImage. 2006; 30(4):1414–1432. [PubMed: 16413796]

Wernicke, C. Der aphasische symptomencomplex: Eine psychologische studie auf anatomischer basis. Cohn & Weigart; Breslau: 1874.

Wozniak JR, Krach L, Ward E, Mueller BA, Muetzel R, Schnoebelen S, Kiragu A, Lim KO. Neurocognitive and neuroimaging correlates of pediatric traumatic brain injury: A diffusion tensor imaging (DTI) study. Archives of Clinical Neuropsychology. 2007; 22:555–568. [PubMed: 17446039]
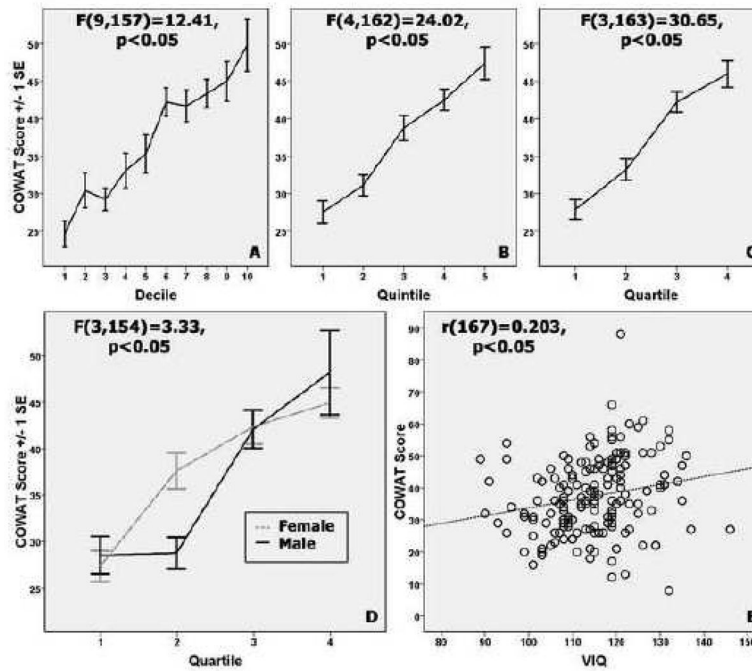
**Figure 1. Age-related trajectories of COWAT performance improvements with varied age bins**
Top panels: Mean COWAT scores across (A) decile, (B) quintile, and (C) quartile age groupings. Statistical results of the age terms in the one-way ANOVAs (IV=age group, DV=COWAT total score) are inset for each plot. Panel D: Age-related trajectories for COWAT performance improvements, accounting for quartile age group and sex (dashed line=females, solid line=males). The difference between males (M(SD)=28.76(7.79)) and females (M(SD)=37.00(9.08)) is only statistically significant ($t(41)=-3.184$, $p<0.05$) within the 2nd quartile (ages 13.58–16.45). Panel E: Correlation between VIQ and COWAT score.
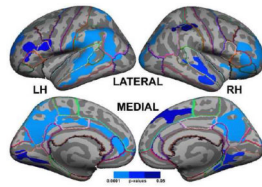
**Figure 2. Areas of significant correlation between COWAT total score and grey matter thickness, controlling for the effects of age group, sex, and VIQ in the full sample**
Statistics are FWE corrected at $p \leq 0.05$. Clusters cover the following cortical areas in the left hemisphere: superior frontal, anterior middle frontal, pars opercularis, pars triangularis, postcentral, supramarginal, superior parietal, inferior parietal, lateral occipital, middle temporal, superior temporal, lingual, fusiform, precuneus, paracentral, isthmus of the cingulate, posterior cingulate, and anterior cingulate. Regions found in the right hemisphere are: pars orbitalis, pars triangularis, lateral orbitofrontal, anterior middle frontal, superior frontal, inferior parietal, superior temporal, supramarginal, postcentral, middle temporal, lateral occipital, lingual, fusiform, precuneus, isthmus of the cingulate, posterior cingulate, and anterior cingulate. Colored lines across the cortex represent anatomical boundaries as defined by Desikan et al. (2006). Focal coordinates for each cluster are presented in Table 2.
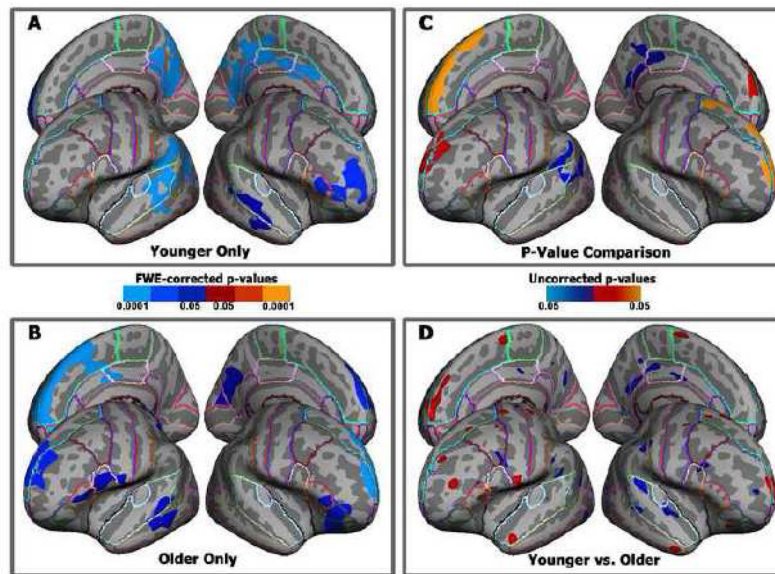
**Figure 3. Comparisons of age group differences in the strength of the COWAT–cortical thickness correlation**

Panels A & B: Illustrations of the COWAT–cortical thickness relationship for the younger age group (9.29–16.45 years, Panel A) and for the older age group (16.46–23.96 years, Panel B). Statistics are FWE corrected at p = 0.05. Blue indicates that the direction of the correlation is negative, i.e., higher COWAT scores predict thinner cortex. Panel C: Illustration of the differences in the COWAT–thickness correlations made by thresholding the differences between the unthresholded *p*-value maps that went into creating Panels A & B. Statistics are FWE corrected at p = 0.05. Red–yellow indicates stronger COWAT–thickness relationships for the older age group; Blue–cyan indicates areas where the COWAT–thickness relationship is stronger for the younger age group. Panel D: Direct comparison of younger versus older subjects taken from the omnibus regression analysis. Statistics are shown at p ≤ 0.05, not corrected for multiple comparisons. Red–yellow indicates stronger COWAT–thickness relationships for the older age group; Blue–cyan indicates areas where the COWAT–thickness relationship is stronger for the younger age group.
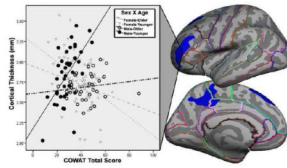
**Figure 4. Sex differences in the COWAT–thickness relationship**
Left hemisphere brain images showing clusters covering areas in rostral middle frontal
gyrus, superior frontal gyrus, medial paracentral gyrus, precuneus, and posterior cingulate
where the direction and magnitude of the correlation between COWAT performance and
grey matter thickness differs between males and females. The scatterplot shows the mean
cortical thickness in this cluster plotted against COWAT total score. 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.

**Table 1**

Demographic and psychometric measures of interest

| Measure | Full Sample | Age < Median | Age > Median |
|---|---|---|---|
| Sex | | | |
|    Female | 93 | 44 | 49 |
|    Male | 74 | 40 | 34 |
| Age (years) | | | |
|    Range | 9.29 – 23.96 | 9.29 – 16.45 | 16.49 – 23.96 |
|    Mean (SD) | 16.39(3.80) | 13.22(2.23) | 19.60(1.88) |
| COWAT total words | | | |
|    Range | 8 – 88 | 8 – 58 | 21 – 88 |
|    Mean (SD) | 37.38(11.93) | 30.49(9.28) | 44.35(10.18) |
| Avg. Parental Education (years) | | | |
|    Range | 10.5 – 21.0 | 11.25 – 20.25 | 10.5 – 21.0 |
|    Mean (SD) | 16.04(2.03) | 16.09(1.74) | 15.98(2.31) |
| FIQ | | | |
|    Range | 89 – 148 | 89 – 148 | 90 – 134 |
|    Mean (SD) | 116.24(9.76) | 116.64(11.19) | 115.83(8.11) |
| VIQ | | | |
|    Range | 89 – 146 | 90 – 146 | 89 – 136 |
|    Mean (SD) | 114.84(9.93) | 115.23(10.63) | 114.45(9.23) |
| PIQ | | | |
|    Range | 83 – 149 | 83 – 149 | 94 – 129 |
|    Mean (SD) | 113.86(11.07) | 114.21(13.29) | 113.49(8.31) |

**Table 2**

Results of the full-sample sMRI analysis. Talairach coordinates, anatomical name, and Brodmann area (BA) numbers are reported for the maximum *t*-value contained within each cluster. Some clusters may span more than one anatomical region or BA; only the localizations of peak statistics are reported here. Talairach Client *v.2.4.2* (http://www.talairach.org/client.html) and standard *FreeSurfer* atlases (Desikan et al., 2006) were used to determine anatomical nomenclature.

| Contrast | Hemi | Max *t* | Talairach X,Y,Z | *p*-value | Size (mm²) | *FreeSurfer* Atlas Label | BA |
|---|---|---|---|---|---|---|---|
| | LH | −5.815 | −50.3, −26.5, −4.1 | 0.0001 | 7377.17 | Superior Temporal | 21 |
| | | −4.976 | −13, −40.4, 42.5 | 0.0001 | 5775.97 | Precuneus | 31 |
| | | −3.869 | −40.7, 34.7, 19.2 | 0.0037 | 1306.3 | Rostral Middle Frontal | 46 |
| | | −3.118 | −15.6, 59.4, 12.7 | 0.0003 | 1779.02 | Superior Frontal | 10 |
| | | −3.071 | −4.9, −88.6, −8.7 | 0.0128 | 1091.83 | Lingual | 18 |
| | | −2.662 | −29.5, −71, −12.5 | 0.0077 | 1179.5 | Fusiform | 19 |
| COWAT Total Score | RH | −6.617 | 7.6, −60.6, 22.1 | 0.0001 | 2640.78 | Precuneus | 31 |
| | | −5.163 | 41.1, 45.1, −5.6 | 0.0001 | 4675.66 | Pars Orbitalis | 10 |
| | | −5.082 | 42.5, −35.9, 39.5 | 0.0453 | 933.13 | Supramarginal | 40 |
| | | −4.501 | 57.4, −13.8, −4 | 0.001 | 1608.6 | Superior Temporal | 21 |
| | | −3.656 | 33.5, −44.2, −15.1 | 0.0013 | 1534.36 | Fusiform | 37 |
| | | −3.431 | 8.3, 17.4, 47.7 | 0.0119 | 1149.69 | Superior Frontal | 6 |
| Female vs. Male | LH | −6.027 | −38.4, 47, 9.5 | 0.0034 | 1321.64 | Rostral Middle Frontal | 10 |
| | | −3.583 | −7.9, 30.3, 54.4 | 0.0216 | 1011.07 | Superior Frontal | 6 |
| | | −3.286 | −12.2, −50.8, 41 | 0.0169 | 1050.93 | Precuneus | 7 |
| Younger vs. Older | RH | 2.719 | 14.5, 61.2, 7.8 | 0.281§ | 616.8 | Superior Frontal | 10 |

§Note: the cluster reported for the age group contrast does not pass correction for multiple comparisons.