Associations Between Cortical Structure and Reading Skills in Beginning Readers

Meaghan Perdue
meaghan.perdue@uconn.edu

Recommended Citation

This work is brought to you for free and open access by the University of Connecticut Graduate School at OpenCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of OpenCommons@UConn. For more information, please contact opencommons@uconn.edu.
Associations Between Cortical Structure and Reading Skills in Beginning Readers

Meaghan V. Perdue

B.A., Emmanuel College, Boston, MA, 2014

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
At the
University of Connecticut
2018
Acknowledgements

I would like to express my gratitude to my advisor, Dr. Nicole Landi, for her support and guidance through this first phase of my graduate education and for continuing to challenge me with innovative interdisciplinary approaches to the study of reading and development. I would also like to thank my committee members, Drs. Ken Pugh and Marie Coppola, for their support of this project; my lab mate, Kayleigh Ryherd, for her positive encouragement and thoughtful feedback; and undergraduate research assistants, Alexa Massad, Katie Foley, and Josh Mednick, for their dedicated time spent preprocessing MRI data.

I am deeply grateful to my husband, Ben Perdue, for his willingness to move and start over in a new state so that I could pursue my education, and for his continued support of my goals in all aspects of my life. Thank you for listening to me talk about research, encouraging me when I feel like giving up, making me laugh, and reminding me to keep my life in balance with the things I love.

This research was supported by NSF Graduate Research Fellowship Program grant DGE-1747453 to Meaghan Perdue, NSF IGERT grant DGE-1144399 to the University of Connecticut, and NIH P01 HD001994-46 to Jay Rueckl, P.I.
Abstract

The acquisition of literacy is a challenging task that engages multiple processing systems to integrate multimodal information. As a result, children exhibit wide variability in reading ability, with the lowest performers classified as having Specific Reading Disability (SRD). Functional and structural magnetic resonance imaging (MRI) research has identified a reading network in the brain consisting primarily of left hemisphere temporo-parietal, occipito-temporal, and inferior frontal regions, but the structural neural substrates that support skilled reading during the formative years of reading acquisition remain underspecified. The present study applied a continuous analytic approach to investigate distinct surface-based properties of cortical structure (cortical thickness and surface area) in association with individual differences in reading-related skills in a sample of children ages 5-8 years. Brain structure was examined in relation to reading-related skills measured at concurrent and lagged follow-up time points in order to assess the stability of these relationships over time. Significant positive correlations with cortical structure were observed in the analysis of reading skills measured 8-12 months after MRI. Cortical surface area in left occipito-temporal, left prefrontal, and right superior temporal regions was associated with sight-word reading performance; Cortical thickness in the right postcentral cortex was associated with phonological awareness performance. These findings indicate that distinct properties of cortical structure are independently related to reading skills and may be salient predictors of reading outcomes over time.
Associations Between Cortical Structure and Reading Skills in Beginning Readers

Introduction

“The whole world opened up to me when I learned to read.” – Mary McLeod Bethune

Literacy is a defining feature of modern society, and the ability to read has become crucial for day-to-day functioning in much of the world. Reading is a complex behavior that requires engagement of multiple cognitive and perceptual systems, from basic visual and auditory processing to working memory and executive function. Because reading evolved relatively recently in human phylogeny, the brain has not developed a specialized “reading apparatus”, but instead recruits multiple neural systems that serve various functions related to reading. Thus, successful acquisition of reading skills relies on adequate development of the various perceptual, cognitive, and neural systems that support reading, as well as responsiveness of these systems to reading instruction. Early language and cognitive development lay the foundation of literacy acquisition long before formal reading instruction begins. During the formative years of reading instruction in alphabetic languages (usually kindergarten through grade 3 in the United States), children progress through phases of learning from establishing the alphabetic principle, the knowledge that letters represent the sounds of language, through applying this rule to sound out words and build a repertoire of automatically recognizable sight-words, and ultimately attaining the ability to fluently read and comprehend text. Given this complex process of learning to read, children exhibit a high degree of variability in their reading outcomes that may affect their abilities at levels of word reading, fluent text reading, and/or reading comprehension. Establishing strong word-level reading skills is critical for subsequent development of reading
fluency and comprehension, which become fundamental skills as schooling transitions from a focus on learning to read to reading to learn.

Children who persistently fail to acquire adequate word-level reading skills despite access to instruction, sufficient motivation, and absence of gross sensory, intellectual, or neurological impairments are diagnosed with Specific Reading Disability (SRD; also called developmental dyslexia), a learning disability characterized by poor word recognition and decoding abilities (Lyon, Shaywitz, & Shaywitz, 2003). SRD affects approximately seven percent of school-age children (Peterson & Pennington, 2015), placing it among the most common learning disabilities. Furthermore, substantial heritability of SRD has been observed through family and twin studies (Grigorenko, 2004), which have led to genetic investigations of this disorder and the identification of candidate genes that confer elevated risk for SRD. Nine of these genes have been replicated in at least one independent sample, giving credence to a potential role in reading (DYX1C1, DCDC2, KIAA0319, C2orf3, MRPL19, ROBO1, GRIN2B, FOXP2 and CNTNAP2; Mascheretti et al., 2017). Broadly, these genes are thought to regulate aspects of neurodevelopment, including processes such as neuronal migration, neurite branching and synapse function (Eicher & Gruen, 2013; Mascheretti et al., 2017). Based on these roles, researchers have proposed that variations on candidate genes contribute to reading disability by setting up atypical neural architecture and functioning early in prenatal development (Bates et al., 2010; Meng et al., 2005).

The proposed genetic effects on neurobiology fit well with evidence from functional and structural magnetic resonance imaging (MRI), which suggests that language-related regions adjacent to the sylvian fissure, along with left occipito-temporal regions are altered in individuals with SRD relative to their typically reading peers (Maisog, Einbinder, Flowers, Turkeltaub, & Eden,
ASSOCIATIONS BETWEEN CORTICAL STRUCTURE AND READING

2008; Richlan, Kronbichler, & Wimmer, 2009, 2013). These regions are thought to be involved in aspects of phonological processing and recoding, integration of orthographic, phonologic, and lexico-semantic information, and automatic word recognition (Pugh et al., 2001). Studies in pre-reading children at familial risk of SRD suggest that atypical structure and function observed in posterior brain regions, particularly in the left temporo-parietal and occipito-temporal cortices, predate reading onset, and may thus represent a cause, rather than a consequence, of poor reading (see Vandermosten, Hoeft, & Norton, 2016 for a review and meta-analysis). These findings point toward the potential identification of neural traits that may supplement behavioral measures as predictors of reading outcomes.

In the sections that follow, I will provide an overview of behavioral and neural correlates of reading ability, address limitations of methods commonly applied to examine brain-behavior relationships in reading, and propose alternative methods to build upon the findings in the existing literature.

Behavioral predictors of reading ability

Early identification of children most likely to struggle with reading is key to best serve the educational needs of children as they learn to read. Several early language skills that can be tested prior to reading acquisition have been identified as strong predictors of reading outcomes, namely letter knowledge, phonological awareness, and rapid naming (for a review, see Ozernov-Palchik & Gaab, 2016). Schatschneider and colleagues (2004) found that these three skills measured at kindergarten served as robust predictors of word identification, reading fluency, and reading comprehension skills at grades 1-2, while vocabulary, oral language (syntax), and visual perceptual skills were weak predictors of the same skills.
Letter knowledge (LK) encompasses the skills of letter naming and knowledge of letter sounds. Letter naming, the ability to identify visually presented letters, provides the foundation of orthographic knowledge upon which reading is built. Letter sound knowledge is a somewhat more complex skill that requires an understanding that letters represent sound units independent of their own names, and thus a greater reliance on phonological skills. Both of these letter knowledge sub-skills uniquely predict subsequent word recognition performance (Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012; Leppänen, Aunola, Niemi, & Nurmi, 2008; McBride-Chang, 1999; Schatschneider et al., 2004).

Phonological awareness (PA), the ability to accurately perceive and manipulate the structural units of speech sounds (Melby-Lervåg, Lyster, & Hulme, 2012), is a critical building block for decoding text (orthographic information) and mapping it to its corresponding spoken form (phonological information). Accordingly, this skill is particularly important in the early stages of reading acquisition in which children rely heavily on sounding out words that they do not automatically recognize. Several studies corroborate evidence that PA skills predict later reading ability, especially at the word-reading level (Hulme et al., 2002; Kirby, Parrila, & Pfeiffer, 2003; Muter, Hulme, Snowling, & Stevenson, 2004). Notably, Kirby and colleagues (2003) reported a strong relationship between kindergarten PA skills and word reading outcomes in grades 1-2, but the strength of the relationship diminished in grades 3-5. The diminishing effects likely indicate that as children become more proficient readers, they rely more on rapid recognition of familiar words (sight-word reading) rather than phonological decoding.

Rapid automatized naming (RAN) is the ability to quickly and accurately identify items presented in a visual array (Araújo, 2015). RAN assessments may include objects, colors, letters,
digits, or some combination of these (called rapid alternating stimulus). RAN performance is thought to reflect aspects of attention, processing speed, automatic retrieval of phonological information, and integration of visual and phonological representations, skills that are important for accurate and fluent reading (Wolf & Bowers, 1999). A recent meta-analysis revealed moderate-to-strong correlations between RAN and concurrent reading performance (Araújo, 2015). A longitudinal study showed that kindergarten RAN performance significantly predicted word identification performance in grades 1-5 with increased effects in grades 3-4, reflecting a pattern opposite to the diminishing effects of PA in the older grades (Kirby et al., 2003). Consistent with the interpretation of the diminishing PA effects, the stronger effects of RAN may indicate a greater reliance upon rapid and automatic recognition of sight words as children progress in reading acquisition.

**Structural brain associations with reading**

To date, the majority of research examining relationships among brain structure and reading have used a case-control approach to compare groups with and without SRD and have applied voxel-based morphometry methods to characterize cortical structure in terms of gray matter volume. Meta-analyses of neural differences between individuals with and without SRD point to atypical structure and function in regions ranging from frontal, perisylvian and occipito-temporal cortices to sub-cortical regions including thalamus and cerebellum (Eckert, Berninger, Vaden, Gebregziabher, & Tsu, 2016; Linkersdörfer, Lonnenmann, Lindberg, Hasselhorn, & Fiebach, 2012; Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan et al., 2009, 2013). The distribution of reported differences is particularly widespread in studies of brain structure, with a lack of convergence apparent from the distinct regions of gray matter reductions emerging
from three meta-analyses that included overlapping samples (Eckert et al., 2016; Linkersdörfer et al., 2014; Richlan et al., 2013). The inconsistency among reports of neuroanatomical differences in SRD is likely caused by variability in samples and analysis methods (Ramus, Altarelli, Jednoróg, Zhao, & Scotto di Covella, 2018). Indeed, the aforementioned meta-analyses include primarily adult samples, with some representation of adolescents and younger children. Developmental differences among these age groups likely contributes to variation in brain structure associated with reading ability, thus it is important to differentiate brain structure characteristics of adults with (a history of) SRD from those of children with SRD. Studies of children and adolescents with SRD reveal a somewhat more consistent pattern of gray matter reductions in temporo-parietal, occipito-temporal, and inferior temporal cortices, though reports of reductions in frontal cortex and cerebellum remain mixed (Clark et al., 2014; Richlan et al., 2009; Williams, Juranek, Cirino, & Fletcher, 2018).

SRD cannot be reliably diagnosed until children have persistently struggled to read, usually between the ages of 7-9, and studies conducted in children and adults with SRD do not elucidate whether the observed neural differences precede, and potentially cause, reading difficulties, or emerge as a result of impaired reading experience. In an effort to disentangle these possibilities, an additional line of research examines brain anatomy and function in children at familial risk of SRD before the onset of reading instruction. Using this method, researchers aim to identify brain differences that are most likely to play an etiological role in SRD by examining samples of children who are likely to go on to develop reading difficulties due to increased genetic risk relative to peers with no family history of SRD. Accordingly, the majority of research examining brain correlates of SRD during the most formative years of reading acquisition has
focused on these at-risk populations. Among these studies, left temporo-parietal and occipito-temporal regions have emerged as the predominant areas of gray matter reductions associated with SRD risk (Black et al., 2012; Clark et al., 2014; Kraft et al., 2016; Linkersdörfer et al., 2014; Raschle, Chang, & Gaab, 2011), along with frontal and right-hemisphere regions in some reports (Black et al., 2012; Clark et al., 2014). Furthermore, several studies of cortical topological properties and lateralization have revealed additional differences associated with SRD risk (Altarelli et al., 2014; Hosseini et al., 2013; Im, Raschle, Smith, Grant, & Gaab, 2016; Vanderauwera et al., 2018). Children with familial risk of SRD showed reduced structural connectivity in the left hemisphere surface area network and greater connectivity the right hemisphere network when examined using a graph theory approach (Hosseini et al., 2013). Furthermore, Im et al. (2016) reported atypical left-hemisphere sulcal pattern (based on between-subjects similarity measures) in children with SRD and pre/beginning readers with familial risk for SRD relative to typical readers and children without familial risk. Reduced leftward lateralization of planum temporale surface area has also been associated with SRD and family risk of SRD (Altarelli et al., 2014; Vanderauwera et al., 2018). These findings indicate that examination of specific properties of gray matter structure may reveal additional links among reading and brain structure that may not be evident when examining only gray matter volume.

Limitations of voxel-based imaging/analysis methods

As noted above, a closer examination of the literature reveals some limitations in the convergence of findings related to the neuroanatomy of SRD that may be caused by heterogeneity of study samples and variability in analysis methods (Ramus, Altarelli, Jednoróg, Zhao, Scotto, et al., 2018). One contributing factor is the use of volume-based characterization
of cortical structure in the majority of morphometric studies of SRD to date. Volumetric measures have been useful to identify cortical regions that are associated with SRD and/or familial risk thereof, but these measures are limited in their ability to capture specific characteristics of cortical structure. Cortical volume is a composite of cortical thickness and cortical surface area, and these two components are thought to reflect distinct features of the underlying neural architecture (Wierenga, Langen, Oranje, & Durston, 2014). According to the radial unit hypothesis of neurodevelopment (Rakic, 1995), processes of neuronal proliferation and migration form a cortical sheet composed of columns. Proliferation of founder cells determines the number of columns, and thus the surface area, of the cortical sheet, whereas column formation (which involves division of daughter cells and migration of neurons towards the outer layers of the cortex) determines the cortical thickness. These processes occur in two phases, so the duration of each phase during development influences the relative size of surface area and cortical thickness. If cortical thickness and cortical surface area are determined by separate neurodevelopmental processes, distinct genes likely contribute to each component. This hypothesis is supported by studies that used twin heritability (Panizzon et al., 2009) and neuroimaging genetic approaches (Winkler et al., 2010) to show independent genetic associations with cortical thickness and surface area. Furthermore, longitudinal and cross-sectional studies that examine the developmental trajectories of cortical surface area, cortical thickness, and cortical volume indicate differential patterns of maturation overall, as well as regionally heterogeneous patterns among these indices from infancy through early adulthood (Amlien et al., 2016; Lyall et al., 2015; Wierenga et al., 2014). The differential patterns of cortical thickness and cortical surface area development likely reflect distinct developmental processes,
therefore caution should be taken in interpreting cortical volume indices alone as they may mask developmental effects by conflating cortical thickness and surface area into one metric (Wierenga et al., 2014).

Neuroanatomical studies applying surface-based analysis methods have shown differential effects of structural properties in relation to reading and SRD. As described above, the majority of research examining gray matter structure in SRD has applied voxel-based volumetric approaches, and points to a trend of gray matter reductions in the reading network of affected children (Eckert et al., 2016; Linkersdörfer et al., 2012; Richlan et al., 2013). In contrast, studies that have applied surface-based approaches to analyze cortical thickness and surface area independently reveal that gray matter reductions may not adequately characterize the neuroanatomy of SRD. Several reports have identified a heightened association between gray matter volume and surface area relative to cortical thickness (Frye et al., 2010; Winkler et al., 2010; Yang, Qiu, Wang, Liu, & Zuo, 2016), indicating that the existing findings from volumetric studies are largely driven by surface area characteristics, and as a result, relationships among cortical thickness and reading may be underrepresented in the literature. Frye and colleagues (2010) suggested that cortical surface area, rather than thickness, drives the reading-related effects observed in gray matter volume. Their findings showed correspondence between volume and surface area in adults with and without a history of SRD, while group differences in cortical thickness showed independent trends. In addition, Black and colleagues (2012) identified associations between history of maternal reading disability and gray matter volume and surface area, but not cortical thickness, in frontal and temporoparietal regions.
Several other studies have revealed effects specific to cortical thickness, including some evidence linking decreased, rather than increased, thickness to better performance in certain reading skills. When examining cortical thickness in children with and without SRD, Clark and colleagues (2014) found reductions in the SRD group in left orbitofrontal cortex and left superior/middle temporal cortex at age 11, and in the left temporal, frontal, middle cingulate, occipito-temporal, and right frontal regions at the pre-reading stage. In a study of children representing a broader age range (7-16 years), Ma and colleagues (2015) reported increased cortical thickness in the left fusiform gyrus and right superior temporal gyrus in children with SRD, and no group differences in surface area or gray matter volume. A recent study of relationships among individual differences in reading-related skills and cortical thickness and gray matter volume revealed distinct brain-behavior relationships by structural metric and reading skill in young adults (Johns et al., 2018). Word-level reading (indexed by a word and pseudoword decoding composite score) positively correlated with gray matter volume in widespread bilateral regions, and with cortical thickness in only right hemisphere regions. In contrast, reading comprehension and print exposure negatively correlated with cortical thickness in widespread bilateral regions; print exposure negatively correlated with gray matter volume only in the right superior frontal cortex, and reading comprehension did not correlate with gray matter volume. These findings point toward some degree of specificity among features of cortical structure and domains of reading ability that warrants further investigation.

Given the independence of cortical surface area and thickness characteristics, surface-based analysis techniques may provide insight to brain-reading relationships beyond what has been revealed through voxel-based morphometry. Several software toolboxes are available to
characterize cortical structure according to surface-based features measured from structural MRI (e.g. Freesurfer: Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999; and Caret: Van Essen et al., 2001) and can be applied to examine relationships between cortical structure and reading. Using these tools, aspects of cortical structure including surface area, cortical thickness, gyrification, and curvature can be analyzed independently across the whole brain or in cortical regions of interest defined by standard atlases or manual drawing. The application of such techniques may help to disambiguate extant reports of gray matter volume associations with reading and provide a more precise characterization of relationships among cortical structure and reading ability. Such research has the potential to inform future efforts to identify brain-based predictors of reading outcomes and draw links among genes and reading skills via intermediate phenotypes at the neural level.

Limitations of the case/control approach

Research in children with or at familial risk of SRD provides an informative basis for identifying brain-behavior associations with reading ability, but several limitations must be addressed before conclusions can be drawn with regard to these relationships. One problem that is germane to case/control studies in research on SRD and familial risk of SRD is the use of arbitrary cut-points to classify groups with SRD based on reading achievement scores. A second issue that emerges from this approach is that group comparisons of children with and without a family risk of SRD provide information about neurobiological characteristics that are related to the family risk, but not necessarily to subsequent reading outcomes. An associated limitation is that this line of research cannot account for characteristics of children who go on to develop SRD despite no familial risk. These concerns will be detailed below, and I will propose the application
of continuous analytic methods to examine individual differences among beginning readers to complement ongoing familial risk studies and address gaps in the information provided by case/control research.

Comparisons of groups with and without SRD provide information about the underlying neural characteristics of reading difficulties, but it remains unknown whether the same neuroanatomical features associated with SRD/SRD-risk are related to individual differences in reading and reading-related skills over the formative years of reading acquisition. Indeed, reading ability exists on a continuum, with SRD representing the lower end, and there has been no reliable consensus to show the existence of a natural breaking point in the distribution by which to classify a disability (Fletcher, 2009; Francis et al., 2005). IQ-discrepancy and low-achievement cut-point based methods of group classification that are frequently used in research on SRD have been shown to be unreliable because they utilize an arbitrary split of a normal distribution of achievement scores that does not adequately represent the characteristics of the resulting groups (Francis et al., 2005). The division of a continuous distribution in this way typically results in non-normal distribution and non-homogeneous variance in the resulting sub-groups, which may then be poorly suited to group comparison statistics such as ANOVA and t-tests (Francis et al., 2005). Furthermore, this approach washes out potentially relevant information about individual differences that is represented in the whole group distribution. Data derived from a continuous distribution of achievement scores is thus best treated as such, using appropriate statistical models to examine patterns of covariance among the variables.

Concerning examination of familial risk, a neuroanatomical feature may be related to risk of SRD, but not differentiate between at-risk children who will and will not go on to develop SRD.
In this case, it is unclear whether the neural substrate directly impacts reading ability. This is important to consider as approximately 40-60% of children with a familial risk of SRD go on to develop typical reading skills (Gallagher, Frith, & Snowling, 2000; Pennington & Lefly, 2001; Snowling, Gallagher, & Frith, 2003), in accordance with a continuous model of genetic liability for complex traits (Bishop, 2015; Snowling et al., 2003). If neural features are to serve as predictors of reading outcomes, they should distinguish between at-risk children who will and will not go on to develop SRD, rather than duplicate risk information that is known based on family history. Some studies address this problem by including brain-behavior correlations to link findings to early language and reading skills (e.g., Raschle, Chang, & Gaab, 2011). Ultimately, longitudinal follow-up of at-risk children is required to differentiate between those neural characteristics that are linked with a genetic predisposition for SRD and those that are related to the subsequent emergence of SRD in at-risk prereaders.

Few studies to date have used this approach, and the findings suggest that specific measures of brain structure in prereaders may be differentially related to familial risk and subsequent development of SRD. Wang and colleagues (2017) combined cross-sectional and longitudinal methods to examine development of white matter tracts across pre-reading, early reading, and fluent reading stages in children with and without familial risk of SRD, and reported mixed findings related to SRD risk and reading outcomes. The family risk group showed reduced white matter integrity in portions of the left arcuate fasciculus and superior longitudinal fasciculus across stages. Longitudinal analysis revealed a slower rate of white matter development in the left arcuate fasciculus in children who subsequently went on to become poor readers than those who went on to become good readers, while no group differences in rate of
development of this region were present when comparing family risk groups. Finally, the rate of development in the right superior longitudinal fasciculus was examined within the family risk group as a potential protective factor, and those at-risk children who went on to become good readers showed a higher rate of white matter development in this region relative to those who went on to become poor readers. Though further research is warranted, these findings suggest that some aspects of white matter structure and development may be related to both familial risk and subsequent reading outcomes.

Using a similar approach, Vanderauwera and colleagues (2018) recently reported that reduced leftward lateralization of surface area in the planum temporale (a region located in the superior temporal lobe that is functionally implicated in higher-order auditory and language processing) was associated with familial risk of SRD in prereaders and adolescents regardless of whether the children were subsequently characterized as having SRD. Because the planum temporale lateralization patterns did not differentiate between children with SRD and typical readers, the authors suggest that planum temporale asymmetry may share an underlying genetic cause with SRD, but likely does not mediate the development of reading difficulties. Together, these studies provide preliminary evidence that aspects of neural structure observed at the pre-reading stage may be differentially related to a genetic predisposition for SRD and later reading ability, and the field will benefit from further longitudinal follow-up of at-risk prereaders. Additionally, neuroimaging genetic approaches that target specific gene-brain-behavior relationships will help to inform a more comprehensive model of genetic risk and mediating factors at the neural level.
Selection of pre-readers at familial risk of developing SRD provides the advantage of increasing the chances that the study includes children who will go on to develop SRD, but this approach is limited in its ability to provide information about characteristics of children who go on to develop reading difficulties despite no known family history. According to a meta-analysis of large-scale twin studies, genetic factors are estimated to account for approximately 51-64 percent of variation in reading ability, depending on the inclusion of samples selected for poor reading ability (Grigorenko, 2004). While this high degree of heritability motivates the examination of SRD based on familial risk, there remains substantial variation that cannot be explained by hereditary influences, with reading difficulties present an estimated 6-13% of children without a familial risk (Gallagher et al., 2000; Pennington & Lefly, 2001; Snowling et al., 2003). It is therefore important to identify traits at the pre-reading level that are related to reading outcomes regardless of familial risk status. This can be achieved by analyzing individual differences across a continuum of reading ability that is unselected for any specific risk factors.

Investigation of individual differences in brain structure and reading outcomes in prereaders not selected for SRD risk using continuous analytic methods may complement case/control research by addressing the above limitations. First, continuous analytic approaches can appropriately characterize the distribution of reading skills without applying arbitrary cut-points to form artificial groups that are poorly suited for group comparisons. Second, brain/behavior associations observed in a sample unselected for risk are expected to reveal factors linked to reading outcomes, rather than a genetic predisposition for SRD. Finally, a sample that is not selected for SRD risk is more likely to represent the distribution of reading ability in
the population and has the potential to reveal brain/behavior associations that may be present regardless of familial risk status.

**Linking brain and behavior to predict reading outcomes**

Pre-literacy skills including LK, PA, and RAN provide substantial prediction of reading outcomes, with overall prediction rates ranging from 73-84%, but a trade-off between sensitivity and specificity of these models leads to problems with over- and under-identification of children at risk of reading difficulties (Ozernov-Palchik & Gaab, 2016). Adding predictors at the genetic and neural levels may serve to build more robust models than those based on behavioral measures alone. Factors such as having a family member with SRD may be included to account for genetic risk factors, and neural measures may serve as intermediate phenotypes that mediate genetic and behavioral risk factors. In one case, the addition of functional neural measures (from EEG and fMRI) significantly increased the variance of reading outcomes explained over behavioral predictors alone (Bach, Richardson, Brandeis, Martin, & Brem, 2013), showing the promise of functional neural measures to contribute to prediction of reading outcomes. However, the potential of structural brain characteristics to contribute to improved sensitivity and specificity of models predicting reading outcomes remains underexplored. Applying structural, rather than functional, neuroimaging in these models may provide practical advantages of ease of acquisition (children are not required to learn a task) and comparability across studies (structural imaging is not susceptible to task differences). Assessment of brain structure as a predictor of reading outcomes depends on first establishing an understanding of the stability of brain/behavior relationships over time, as it is unknown whether characteristics of brain structure associated with concurrent reading skills will maintain their association with reading outcomes measured at
a later follow-up time point. Further research is needed to build a stronger understanding of the neurobiological underpinnings of reading during the pivotal stages of reading acquisition, and their relationships to subsequent reading performance.

**Aims**

The present study investigates the association between cortical structure (indexed by cortical thickness and surface area) and reading-related skills in children with brain imaging and behavioral testing measures at concurrent time points and children with behavioral testing measures conducted at a lagged follow-up timepoint 8-12 months after imaging. I aim to (1) investigate whether cortical thickness and surface area in brain regions previously shown to have gray matter differences associated with SRD exhibit relationships with individual differences in reading-related skills in children unselected for SRD/SRD-risk; (2) explore relationships between reading skills and cortical structure across the whole brain; (3) examine stability of brain/behavior relationships when behavioral testing is conducted at a concurrent vs. lagged time point. A regions-of-interest approach is applied to examine brain regions that have previously been associated with SRD and/or familial risk (Eckert et al., 2016; Richlan et al., 2013; Vandermosten et al., 2016). I predict that cortical structure in bilateral temporo-parietal and left occipito-temporal regions will be related to individual differences in reading-related skills. Due to the distinct developmental trajectories of cortical thickness and surface area, I predict that differential effects will be observed for these two measures. A follow-up exploratory analysis is conducted using whole-brain correlation analyses of cortical thickness and surface area with reading skills in order to identify whether any regions that have not been consistently linked to familial risk for SRD are associated with individual differences in reading ability. I aim to
contribute to the existing literature on brain-reading relationships in beginning readers by applying continuous analytic methods to complement findings from previous case-control studies of children with or at familial risk of SRD. Furthermore, I will use surface-based measures of cortical structure in an effort to disentangle effects of cortical thickness and cortical surface area, which may be influenced by distinct developmental processes. Finally, analysis of brain imaging and behavioral measures acquired at concurrent and later follow-up time points will allow for assessment of the stability of brain-reading relationships over time and may inform future research practices for using metrics of brain structure as predictors of subsequent reading outcomes.

**Methods**

**Participants**

The present study included 26 participants drawn from a larger study that examines behavioral and neural characteristics of reading acquisition from kindergarten to grade 3. Participants were classified into two overlapping groups: the Concurrent Group included children who completed a behavioral testing session and a structural MRI scan at a concurrent time point (0-3 months between visits; n = 15, age in months at MRI $M = 77.53, SD = 6.31$; age in months at behavioral testing $M = 76.27, SD = 6.64$; 9 females, 6 males); the Follow-up Group included children who completed an initial structural MRI scan followed by a behavioral testing session 8-12 months later (n = 19, age in months at MRI $M = 79.58, SD = 4.31$; age in months at follow-up $M = 89.05, SD = 4.31$; 11 females, 8 males). Eight subjects were included in both Concurrent and Follow-up groups. All participants were native speakers of American English, reported no history of neurological or psychiatric disorder, had normal or corrected-to-normal vision, normal
hearing, and met a minimum standard score of 85 on an age-appropriate test of full-scale IQ. Three participants had a family member with a history of SRD according to parent report. The study protocol was approved by the Yale University Human Investigation Committee.

**Behavioral assessment**

Participants completed a battery of standardized assessments of reading and language at a concurrent (within 0-3 months) and/or follow-up (8-12 months later) time point in reference to their MRI scan. Assessments that reflect early word reading and strong predictors of word reading were selected for the present analysis, including:


*Comprehensive Test of Phonological Processing: Elision and Blending Words* subtests to test phonological awareness (R. K. Wagner, Torgesen, & Rashotte, 1999). The Elision sub-test involves segmenting and removing phonological units from spoken words to form other words (e.g., say “toothbrush”; now say “toothbrush” without saying “tooth”). The Blending Words subtest requires the examinee to form a word from sound units presented serially (e.g., “What word do these sounds make: /k/ ... /a/ ... /t/?”). Scores from these two subtests were averaged to obtain a composite phonological awareness (PA) score for each participant.

*Comprehensive Test of Phonological Processing: Rapid Digit Naming* subtest to test rapid automatized naming (RAN) of digits (R. K. Wagner et al., 1999). A card displaying an array of
numbers is presented and the examinee must name the numbers in sequential order as quickly as possible.

Full-scale IQ was assessed to determine study eligibility using the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011) for children who initially entered the study at age 6 or above, or the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III; Wechsler, 2002) for children who initially entered the study younger than 6 years of age.

Brain image acquisition and preprocessing

High-resolution T1-weighted 3D MPRAGE anatomical images were acquired using a Siemens 3T Trio magnetic resonance imaging (MRI) system (TE = 2.77 ms, TR = 2530 ms; FOV = 256 x 256 voxel matrix; voxel size = 1.0 x 1.0 x 1.0 mm). Preprocessing and analysis of anatomical images was conducted using FreeSurfer v. 5.3 software (Dale et al., 1999). The automated pipeline for surface-based cortical reconstruction and volumetric segmentation was applied to each participant individually. The reconstruction pipeline includes skull stripping, volumetric labeling, intensity normalization, white matter segmentation, surface atlas registration, surface extraction, and gyral labeling. Output volumes and surfaces from the automated reconstruction pipeline were inspected for accuracy of skull stripping and segmentation of gray matter, white matter, and cerebrospinal fluid. Visual inspection revealed exclusion of gray matter from the segmentations in many subjects, so expert options were applied to correct the problem by adjustment of the intensity thresholds used to classify gray matter. Additional manual edits were made to correct for local skull-stripping and intensity normalization errors as appropriate according to the Freesurfer Troubleshooting Tutorial (https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/TroubleshootingData).
Regions of Interest (ROI) Identification

A priori regions of interest were identified based on several meta-analyses of gray matter structure associated with SRD. A meta-analysis of structural and functional MRI research in pre-readers at risk of SRD (Vandermosten et al., 2016) was selected as the primary source of ROIs because the ages included aligned most closely with the ages of the children in the present study. This study showed SRD-risk related effects in the left temporo-parietal (TP) cortex, the left fusiform gyrus/occipito-temporal (OT) cortex, the right parietal lobe (PL), and left cerebellum. The left cerebellum ROI was omitted from the present analysis because acquisition and processing parameters were not optimized to characterize cerebellar structure. An additional left superior temporal sulcus (STS) ROI was included in the present study because this region was identified in 2 of 3 meta-analyses of gray matter volume associated with SRD in adults and adolescents (Eckert et al., 2016; Richlan et al., 2013).

For the present study, anatomical regions corresponding to the selected ROIs were selected from the Destrieux atlas provided with Freesurfer (Destrieux, Fischl, Dale, & Halgren, 2010). The Freesurfer reconstruction pipeline automatically parcellates the cortex into anatomical regions based on sulco-gyrual structures and outputs cortical surface area, thickness, and volume measures of each region for each subject. The atlas-based parcellations included in the present analysis are listed in table 1. Total surface area and mean cortical thickness of each ROI was extracted for each subject for inclusion in correlation analyses. For the left STS, surface area was summed and cortical thickness was averaged across the 3 parcellations for each subject.
Table 1: ROIs and corresponding Destrieux Atlas regions

<table>
<thead>
<tr>
<th>ROI</th>
<th>Anatomical Regions(s)</th>
<th>Destrieux Atlas Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left TP</td>
<td>inferior parietal/supramarginal gyrus</td>
<td>G_pariet_in-Supramar</td>
</tr>
<tr>
<td>Left OT</td>
<td>occipito-temporal/lateral fusiform gyrus</td>
<td>G_oc-temp_lat-fusifor</td>
</tr>
<tr>
<td>Right PL</td>
<td>inferior parietal/supramarginal gyrus</td>
<td>G_pariet_in-Supramar</td>
</tr>
<tr>
<td>Left STS</td>
<td>superior temporal gyrus/planum temporale</td>
<td>G_temp_sup-Plan_tempo</td>
</tr>
<tr>
<td></td>
<td>transverse temporal sulcus</td>
<td>S_temporal_transverse</td>
</tr>
<tr>
<td></td>
<td>superior temporal gyrus/transverse temporal</td>
<td>G_temp_sup-G-T_transv</td>
</tr>
</tbody>
</table>

Data preparation

Age at MRI, raw behavioral testing scores, and ROI metrics were mean-centered and scaled to improve comparability among variables that are measured on different scales using the Caret package (Kuhn et al., 2016) in R (version 3.5.1; R Core Team, 2016). To test normality of the distribution for each variable, distributions were visually inspected using density plots and Shapiro-Wilk normality tests were conducted (Shapiro & Wilk, 1965) using the Stats package. Skewness was quantified using the e1071 package (Dimitriadou, Hornik, Leisch, Meyer, & Weingessel, 2005) in R. Variables that showed non-normal distribution in density plots and significant Shapiro-Wilk tests at $p < .1$ were transformed to correct for skewness. The following variables exhibited a non-normal distribution: CTOPP RD score and age at MRI in the concurrent analysis group; CTOPP PA composite score, age at MRI, and left TP cortical thickness in the follow-up group. Box-Cox transforms using the equation $y(\lambda) = (y^\lambda - 1)$, given $\lambda \neq 0$; if $\lambda = 0$, then $\log(y)$ were applied to these variables using the Caret package in R. Normality tests on the transformed data revealed that the following distributions were no longer significantly non-normal: transformed CTOPP RD score in the concurrent analysis group, transformed age at MRI in the follow-up group, transformed left TP cortical thickness. Transformed age at MRI in the concurrent...
group and transformed CTOPP PA composite score in the follow-up group remained significantly non-normal. Tukey’s Ladder of Powers transforms (Tukey, 1977) were applied to these variables using the Rcompanion package (Mangiafico, 2017) in place of Box-Cox transforms. The Tukey’s Ladder of Powers transformed distribution of age at MRI in the concurrent group was not significantly skewed. However, the Tukey-transformed CTOPP PA composite score in the follow-up group remained non-normally distributed. The Box-Cox transformed CTOPP PA composite data \((W = 0.899, p = 0.048; \text{skewness} = -0.611)\) was used for subsequent analysis because it showed reduced skewness relative to the untransformed data \((W = .872, p = .015; \text{skewness} = -0.781)\) and the Tukey-transformed data \((W = 0.893, p = 0.0367; \text{skewness} = -0.874)\). Shapiro-Wilk statistics and density plots for all non-normally distributed variables before and after transformation are provided in Supplementary Information (tables S1 and S2; figures S1-S5).

**ROI analysis**

Semi-partial correlations controlling for effects of age and sex on cortical structure measures were conducted to test relationships between each ROI metric (cortical thickness and surface area) and each reading related measure (LW, WA, PA, and RAN). Mean-centered, scaled, and transformed (as appropriate) variables were entered into analysis. Semi-partial correlations were run by first fitting a linear model for each ROI metric (e.g. left TP surface area) with sex and age at MRI to calculate residual variance not explained by sex and age. Residuals for each ROI were then entered into Pearson correlation tests with each centered, scaled, and transformed behavioral measure. Correlations were run separately within each analysis group (Concurrent Group and Follow-up Group). Analysis was conducted using the stats package in R (R Core Team, 2016).
Exploratory whole-brain analysis

Exploratory analysis was conducted in the Freesurfer neuroimaging analysis suite (Dale et al., 1999; Fischl et al., 1999) to examine relationships between cortical structure and reading-related skills across the whole brain. Generalized linear models (GLMs) were built using Freesurfer’s MRI_glmfit function. Centered, scaled and transformed variables were entered into analysis, with sex and age at MRI as variables of no interest and preprocessed behavioral test scores as predictor variables. Independent GLMs were built for each reading-related measure (LW, WA, PA and RAN) for each structural metric, and were run separately for right and left hemispheres. Independent models were run within each analysis group (concurrent testing group and lagged follow-up testing group). Cluster-wise correction for multiple comparisons was conducted using Monte-Carlo Simulation in Freesurfer with a vertex-wise cluster forming threshold of $p < .05$ over 10,000 iterations, and results were evaluated at a corrected threshold of $p < .05$.

Results

Behavioral Characteristics

Descriptive statistics of behavioral testing scores for each analysis group are reported in Table 2.
Table 2: Behavioral characteristics by analysis group. Raw and standardized scores presented for comparability with published literature. Centered and scaled raw scores were entered into analyses. WJIII: *Woodcock Johnson III Tests of Achievement*; CTOPP: *Comprehensive Test of Phonological Processing*. CTOPP PA Composite Raw score was calculated by averaging raw scores from CTOPP Elision Blending Words subtests. CTOPP PA Composite Scaled score was calculated by averaging scaled scores from CTOPP Elision Blending Words subtests.

<table>
<thead>
<tr>
<th>Analysis Group</th>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concurrent</td>
<td>Age at behavioral testing (months)</td>
<td>76.267</td>
<td>6.638</td>
</tr>
<tr>
<td>N=15</td>
<td>Age at MRI (months)</td>
<td>77.533</td>
<td>6.312</td>
</tr>
<tr>
<td></td>
<td>Months between testing and MRI</td>
<td>1.267</td>
<td>1.710</td>
</tr>
<tr>
<td></td>
<td>IQ Standard Score</td>
<td>105.933</td>
<td>13.525</td>
</tr>
<tr>
<td></td>
<td>WJIII LW Raw Score</td>
<td>30.867</td>
<td>8.782</td>
</tr>
<tr>
<td></td>
<td>WJIII LW Standard Score</td>
<td>114.600</td>
<td>14.510</td>
</tr>
<tr>
<td></td>
<td>WJIII WA Raw Score</td>
<td>10.467</td>
<td>4.897</td>
</tr>
<tr>
<td></td>
<td>WJIII WA Standard Score</td>
<td>115.267</td>
<td>8.233</td>
</tr>
<tr>
<td></td>
<td>CTOPP Elision Raw Score</td>
<td>15.667</td>
<td>4.370</td>
</tr>
<tr>
<td></td>
<td>CTOPP Elision Scaled Score</td>
<td>10.333</td>
<td>1.799</td>
</tr>
<tr>
<td></td>
<td>CTOPP Blending Words Raw Score</td>
<td>18.667</td>
<td>5.010</td>
</tr>
<tr>
<td></td>
<td>CTOPP Blending Words Scaled Score</td>
<td>10.533</td>
<td>2.356</td>
</tr>
<tr>
<td></td>
<td>CTOPP PA Composite Raw Score</td>
<td>17.167</td>
<td>4.082</td>
</tr>
<tr>
<td></td>
<td>CTOPP PA Composite Scaled Score</td>
<td>10.433</td>
<td>1.646</td>
</tr>
<tr>
<td></td>
<td>CTOPP Rapid Naming Digits Raw Score</td>
<td>26.667</td>
<td>5.010</td>
</tr>
<tr>
<td></td>
<td>CTOPP Rapid Naming Digits Scaled Score</td>
<td>10.667</td>
<td>1.543</td>
</tr>
<tr>
<td>Follow-up</td>
<td>Age at behavioral testing (months)</td>
<td>79.579</td>
<td>4.312</td>
</tr>
<tr>
<td>N=19</td>
<td>Age at MRI (months)</td>
<td>89.053</td>
<td>4.314</td>
</tr>
<tr>
<td></td>
<td>Months between testing and MRI</td>
<td>9.474</td>
<td>1.429</td>
</tr>
<tr>
<td></td>
<td>IQ Standard Score</td>
<td>108.632</td>
<td>10.101</td>
</tr>
<tr>
<td></td>
<td>WJIII LW Raw Score</td>
<td>44.000</td>
<td>7.724</td>
</tr>
<tr>
<td></td>
<td>WJIII LW Standard Score</td>
<td>114.737</td>
<td>9.893</td>
</tr>
<tr>
<td></td>
<td>WJIII WA Raw Score</td>
<td>17.895</td>
<td>6.420</td>
</tr>
<tr>
<td></td>
<td>WJIII WA Standard Score</td>
<td>111.632</td>
<td>8.036</td>
</tr>
<tr>
<td></td>
<td>CTOPP Elision Raw Score</td>
<td>25.158</td>
<td>6.353</td>
</tr>
<tr>
<td></td>
<td>CTOPP Elision Scaled Score</td>
<td>11.895</td>
<td>2.747</td>
</tr>
<tr>
<td></td>
<td>CTOPP Blending Words Raw Score</td>
<td>23.684</td>
<td>3.001</td>
</tr>
<tr>
<td></td>
<td>CTOPP Blending Words Scaled Score</td>
<td>11.526</td>
<td>2.366</td>
</tr>
<tr>
<td></td>
<td>CTOPP PA Composite Raw Score</td>
<td>24.421</td>
<td>4.372</td>
</tr>
<tr>
<td></td>
<td>CTOPP PA Composite Scaled Score</td>
<td>11.711</td>
<td>2.317</td>
</tr>
<tr>
<td></td>
<td>CTOPP Rapid Naming Digits Raw Score</td>
<td>20.000</td>
<td>4.607</td>
</tr>
<tr>
<td></td>
<td>CTOPP Rapid Naming Digits Scaled Score</td>
<td>10.789</td>
<td>2.200</td>
</tr>
</tbody>
</table>
ROI results

Descriptive statistics of ROI structural metrics are reported in Table 3. ROI analysis did not reveal any significant brain/behavior correlations for either the Concurrent or Follow-up analysis.

Table 3: ROI structural characteristics by analysis group. Surface area metrics represent total surface area of the anatomically defined ROI; Cortical Thickness metrics represent the average thickness across the ROI.

<table>
<thead>
<tr>
<th>Analysis Group</th>
<th>ROI</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concurrent</td>
<td>Estimated Total Intracranial Volume (mm³)</td>
<td>1486662.634</td>
<td>199526.817</td>
</tr>
<tr>
<td>N=15</td>
<td>Left TP Surface Area (mm²)</td>
<td>2015.933</td>
<td>342.304</td>
</tr>
<tr>
<td></td>
<td>Left OT Surface Area (mm²)</td>
<td>1342.400</td>
<td>284.045</td>
</tr>
<tr>
<td></td>
<td>Left STS Surface Area (mm²)</td>
<td>1173.333</td>
<td>217.120</td>
</tr>
<tr>
<td></td>
<td>Right TP Surface Area (mm²)</td>
<td>1729.400</td>
<td>233.318</td>
</tr>
<tr>
<td></td>
<td>Left TP Cortical Thickness (mm)</td>
<td>3.436</td>
<td>0.338</td>
</tr>
<tr>
<td></td>
<td>Left OT Cortical Thickness (mm)</td>
<td>3.391</td>
<td>0.205</td>
</tr>
<tr>
<td></td>
<td>Left STS Cortical Thickness (mm)</td>
<td>3.187</td>
<td>0.376</td>
</tr>
<tr>
<td></td>
<td>Right TP Cortical Thickness (mm)</td>
<td>3.484</td>
<td>0.309</td>
</tr>
<tr>
<td>Follow-up</td>
<td>Estimated Total Intracranial Volume (mm³)</td>
<td>1510099.638</td>
<td>213542.470</td>
</tr>
<tr>
<td>N=19</td>
<td>Left TP Surface Area (mm²)</td>
<td>2132.105</td>
<td>356.033</td>
</tr>
<tr>
<td></td>
<td>Left OT Surface Area (mm²)</td>
<td>1268.158</td>
<td>245.567</td>
</tr>
<tr>
<td></td>
<td>Left STS Surface Area (mm²)</td>
<td>1211.684</td>
<td>219.688</td>
</tr>
<tr>
<td></td>
<td>Right TP Surface Area (mm²)</td>
<td>1868.526</td>
<td>260.720</td>
</tr>
<tr>
<td></td>
<td>Left TP Cortical Thickness (mm)</td>
<td>3.426</td>
<td>0.277</td>
</tr>
<tr>
<td></td>
<td>Left OT Cortical Thickness (mm)</td>
<td>3.370</td>
<td>0.197</td>
</tr>
<tr>
<td></td>
<td>Left STS Cortical Thickness (mm)</td>
<td>3.180</td>
<td>0.364</td>
</tr>
<tr>
<td></td>
<td>Right TP Cortical Thickness (mm)</td>
<td>3.513</td>
<td>0.210</td>
</tr>
</tbody>
</table>

Exploratory whole-brain results

Exploratory whole-brain analysis in the Concurrent Group did not yield any significant brain/behavior correlations at the cluster-corrected threshold.

Significant correlations between cortical structure and sight-word reading and phonological awareness were identified in the Follow-up Group (Table 4; Figure 1). Specifically, positive correlations were found between surface area and LW score (centered and scaled) in the
left prefrontal cortex (PFC; peak vertex \( r = .692, p = .004 \)), the left medial occipito-temporal cortex (M-OT; peak vertex \( r = .682, p = .005 \)), and the right superior temporal gyrus/superior temporal sulcus (STG/STS; peak vertex \( r = .731, p = .002 \)). A positive correlation between cortical thickness and PA score (centered, scaled, and transformed) was identified in the right postcentral gyrus/sulcus (PC; \( r = .773, p < .001 \)). No significant correlations among cortical structure and WA or RAN were observed.

<table>
<thead>
<tr>
<th>Behavioral Measure</th>
<th>Cortical Measure</th>
<th>Hemisphere</th>
<th>Cluster</th>
<th>Area (mm²)</th>
<th>( r )</th>
<th>( p )</th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>LW</td>
<td>SA</td>
<td>Left</td>
<td>Prefrontal Cortex (PFC)</td>
<td>1533.7</td>
<td>0.6921</td>
<td>0.0042</td>
<td>-35.6</td>
<td>53.1</td>
<td>-7.3</td>
</tr>
<tr>
<td>LW</td>
<td>SA</td>
<td>Left</td>
<td>Medial occipito-temporal cortex (M-OT)</td>
<td>1696.81</td>
<td>0.6824</td>
<td>0.0051</td>
<td>-28.6</td>
<td>-42.7</td>
<td>-8</td>
</tr>
<tr>
<td>LW</td>
<td>SA</td>
<td>Right</td>
<td>Superior temporal gyrus/sulcus (STG/STS)</td>
<td>1444.78</td>
<td>0.7314</td>
<td>0.0019</td>
<td>46.5</td>
<td>11</td>
<td>-22.6</td>
</tr>
<tr>
<td>PA</td>
<td>CT</td>
<td>Right</td>
<td>Postcentral Gyrus/Sulcus (PCG/PCS)</td>
<td>1237.5</td>
<td>0.7727</td>
<td>0.0007</td>
<td>49.8</td>
<td>-16.1</td>
<td>52.6</td>
</tr>
</tbody>
</table>

**Table 4:** Significant brain/behavior correlations and statistics at peak vertices for Follow-Up Testing group
SA = Surface Area, CT = Cortical Thickness

**Discussion**

An exploratory analysis of the whole cortex revealed associations between characteristics of cortical structure and reading-related skills assessed 8-12 months after MRI acquisition. No significant associations between cortical structure and reading-related skills assessed at a concurrent time point (within 0-3 months of MRI) were observed. Analysis of a priori ROIs did not reveal any significant associations among structure in the selected brain regions and reading-related skills at either time point.
Significant positive correlations with cortical structure were observed for outcomes in sight-word reading (LW) and phonological awareness (PA) skills measured 8-12 months following MRI acquisition. A pattern emerged such that LW scores were related to regional cortical surface area in both hemispheres while PA scores were related to regional cortical thickness in the right hemisphere. Specifically, higher surface area in the left M-OT, left PFC, and right STG/STS was associated with higher performance on the LW test. The present findings are consistent with existing evidence of structural and functional brain associations with reading and SRD in these regions.

**Positive correlation between surface area and follow-up LW score in the left M-OT**
The left occipito-temporal cortex (OT) is often referred to as the “visual word form area” due to observed specificity of functional activation of this region in response to printed words and letters in adults and children (Centanni et al., 2018; McCandliss, Cohen, & Dehaene, 2003). Reduced gray matter volume in the left OT has been reported in adolescents with SRD (Kronbichler et al., 2008) and pre-readers at familial risk of SRD (Raschle et al., 2011). Individual differences in sentence reading speed positively correlated with gray matter volume in the left OT in German-speaking adolescents (Kronbichler et al., 2008); on the other hand, a negative correlation between gray matter volume in the left OT and reading speed index was reported in French-speaking children (Simon et al., 2013), which may reflect a nuanced relationship among reading and cortical structure in this region based on language or developmental differences. The present effect of surface area in the left M-OT may indicate that these previously identified gray matter volume effects are driven by surface area, given the heightened association between gray matter volume and surface area, relative to cortical thickness (Frye et al., 2010; Winkler et al., 2010; Yang et al., 2016). The observed association with sight-word reading skills underpins the functional relevance of this region for visual word recognition, and reduced surface area in this region may be functionally linked to less efficient processing of text. Further investigation of structure/function relationships in the left OT is needed to address directional causality among structure and function of the left OT and its relation to reading ability.

*Positive correlation between surface area and follow-up LW score in the left PFC*

Gray matter structure in the left prefrontal cortex has also been previously linked to SRD and individual differences in reading skills. Notably, one meta-analysis of voxel-based studies in SRD revealed reduced gray matter volume in left PFC/inferior frontal gyrus in SRD, along with
positive correlations of gray matter volume in this region with sight-word reading, pseudoword reading, and passage comprehension (Eckert et al., 2016). In addition, reduced cortical thickness in the left PFC has been reported in children with SRD relative to typical readers (Clark et al., 2014). Lower gray matter volume in bilateral PFC has been associated with increased degree of maternal history of reading difficulties, pointing toward a potential hereditary effect in this region (Black et al., 2012). A recent study of individual differences in reading skills and surface-based cortical measures in young adults showed a positive correlation between a decoding composite score (consisting of sight-word and pseudo-word reading measures) and gray matter volume in the left PFC (Johns et al., 2018). The correspondence between this finding and the present finding in school-age children indicates that the association between left PFC cortical structure and word reading skills may remain stable across development. Accordingly, the left PFC may be a reasonable target for investigation of genetic contributions to reading ability due to the link to familial risk of SRD and the apparent stability of this brain-behavior association across age groups. Although the PFC is not typically thought of as a part of the reading network, it may play a role in reading through domain-general processes that are involved in learning to read and reading efficiently (e.g. cognitive control of semantic retrieval; A. D. Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). Further research is warranted to examine structure-function relationships in the left PFC and the role of this region in the acquisition of reading skills.

*Positive correlation between surface area and follow-up LW score in the right STG/STS*

In addition to the left hemisphere effects discussed above, surface area in the right STG/STS positively correlated with LW performance in the follow-up testing group. Although cortical structure of the left STG/STS is more commonly associated with reading (Clark et al.,
2014; Eckert et al., 2016; Linkersdörfer et al., 2014; Richlan et al., 2013), several studies have reported relationships among reading and structure in the right hemisphere homolog (Johns et al., 2018; Ma et al., 2015; Richlan et al., 2013; Williams, Juranek, Cirino, & Fletcher, 2018). Williams et al. (2018) and Johns et al. (2018) reported positive correlations between cortical thickness in the right STG/STS and word reading scores in children and young adults, respectively. These reports are contrary to an earlier finding of increased cortical thickness and increased rightward lateralization of cortical thickness in the right STG of children and adolescents with a history of SRD regardless of remediation status (Ma et al., 2015). The present results do not clarify this ambiguity because the association with word reading in this region was found only with surface area, and not thickness. However, the present findings are consistent with a previous meta-analysis showing reduced gray matter volume in the right STG in groups with SRD (Richlan et al., 2013), as well as an additional finding from Johns et al. (2018) that showed a positive correlation between decoding performance and gray matter volume in the right MTG/STG.

The link between right STG/STS and reading may arise from a role of this region bilaterally in phonological aspects of language processing. Hickok and Poeppel’s (2007) dual-stream model of the functional anatomy of language posits that early spectrotemporal analysis of speech input as well as phonological processing and representation occur in bilateral superior temporal regions. Accordingly, cortical structure of the STG/STS bilaterally may be related to mapping between phonological representations and orthographic representations as required for decoding text. The reason for an association with word reading in the right, but not the left, STG/STS in the present study is unclear, though it is possible that a similar effect was present in
the left hemisphere at a sub-threshold level that could not be detected due to a lack of statistical power.

Positive correlations between cortical thickness and follow-up PA score in right PCG/PCS

Cortical thickness in the right postcentral gyrus/sulcus (PCG/PCS) was positively correlated with phonological awareness (PA) at follow-up testing. One study showing functional activation in bilateral pre/postcentral cortex during a rhyming task in children with and without dyslexia points to a potential role of this region in phonological processing (Temple et al., 2001). Additional functional and structural imaging research has shown an association of the right precentral/postcentral cortex with word/pseudoword reading (Evans, Flowers, Napoliello, & Eden, 2014; Johns et al., 2018; Maisog et al., 2008). A meta-analysis of fMRI studies examining activation in response to visually presented words, pseudowords and/or letters showed reduced activation in the right PCG in adults and adolescents with SRD relative to typical readers (Maisog et al., 2008). With regard to cortical structure, Johns and colleagues (2018) reported a positive association between cortical thickness in the right pre/postcentral cortex and word/pseudoword reading ability in young adults, and Evans and colleagues (2014) observed reduced gray matter volume in this region in groups with SRD. These findings may be linked to the role of PA in the development of decoding skills that may have a cascading effect on reading performance. Additional research is needed to clarify the role of the right PCG/PCS in reading ability.

Distinct associations between structural properties and behavioral measures

The pattern of findings in the present study indicates relationships of surface area with sight-word reading ability and of cortical thickness with PA skills, along with distinct regions of effects for these relationships. The observed correlations among surface area and LW
performance are consistent with regions of prior reports of gray matter volume reduction in groups with/at risk for SRD, supporting the hypothesis that surface area largely accounts for the relationships between gray matter volume and SRD (Frye et al., 2010). The association between cortical thickness and PA performance was observed in a region that has not been consistently linked to SRD in the voxel-based structural neuroimaging literature, and may reveal a structural relationship specific to cortical thickness that is not well-captured by volumetric measures.

The regions of effects observed in this study have been previously linked to functional activation during reading tasks, but the specific association between brain structure and function, and possible differential relationships of surface area and cortical thickness to function remain poorly understood. A recent multi-modal imaging analysis showed that different properties of cortical structure (i.e., surface area, thickness, volume, gyrification) covaried with specific graph theory metrics of resting state functional connectivity, but this study could not identify the biological mechanisms underlying these relationships (Yang et al., 2016). Multimodal human neuroimaging studies employing measures of structure, function, neurochemistry, and neuroimaging genetic approaches, as well as animal studies linking histological, functional, and genotypic traits will be required to characterize the nature of structure-function relationships in the brain. In the meantime, the evidence of differential relationships of surface area and cortical thickness with behavior and function supports the independent examination of these structural properties. The findings reported here should be further explored in larger samples in order to confirm whether these patterns are characteristic of brain-reading relationships.

The differential findings for LW and PA performance may also be explained by methods of grouping participants for case-control studies of SRD. Case-control studies typically use a
measure or set of measures of word/pseudoword reading skills to group participants, and seldom include measures of reading sub-skills such as PA to characterize SRD. As a result, associations between cortical structure and individual differences in word reading ability are expected to align more closely with findings reported in the case-control literature than would associations with PA performance. This highlights an advantage of the continuous analysis approach to reveal associations with reading-related skills that may not emerge when examining unidimensional profiles of reading ability in group designs. Further examination of individual differences may thus help to build a more comprehensive characterization of brain-reading relationships.

The present study did not reveal significant relationships of cortical structure with performance on the RAN or WA measures. Johns et al. (2018) also failed to detect significant structural associations with RAN using a similar approach in young adults; On the other hand, Raschle et al., (2011) reported significant positive correlations between gray matter volume in left TP and OT ROIs and RAN performance in pre-readers. These contrasting findings could be attributed to age differences between the two studies or to the use of different RAN sub-tests across studies (Johns et al. (2018) included rapid letter naming, Raschle et al. (2011) included a composite of rapid object naming and rapid color naming, and the present study included rapid digit naming). Further research is required to determine the nature of these effects. The null findings for WA performance in the present study are difficult to explain, as pseudoword reading has been associated with functional activation in response to print (Pugh et al., 2013; Simos et al., 2001). This functional link motivates the expectation that properties of brain structure may also correlate with pseudoword reading. One previous report linked gray matter volume of the left STG to timed pseudoword reading in a small group of typical adult readers, but not those
with SRD (Steinbrink et al., 2008). Larger sample sizes may be required to determine the existence of structural relationships with pseudoword reading or confirm null effects.

Taken together, the present findings show that cortical structure in beginning readers is associated with sight-word reading and PA skills measured 8-12 months later. The presence of the effects over a lagged period indicate that aspects of cortical structure in the left M-OT, left PFC, right STG/STS, and right PCG/PCS may affect activity in these regions that supports the development of reading-related skills. Longitudinal research examining the structure/function relationships of these regions across development is needed, as bidirectional effects among structure and function, as well as aspects of connectivity among the regions, likely contribute to a complex set of influences on reading ability.

Limitations

Several limitations must be addressed with regard to the present study. First, the small samples included in the analysis (Concurrent Group, n = 15; Follow-up Group, n = 19) yielded low statistical power to detect medium-small effects. For target power of .80, these sample sizes are sufficient to detect effects with a correlation coefficient ≥ .60 (Algina & Olejnik, 2003). This lack of statistical power may explain the null findings for ROI analyses in both groups and for the exploratory analysis in the Concurrent Group. Accordingly, this study does not provide conclusive evidence of a lack of associations between brain structure and reading-related skills assessed at a concurrent time point. Follow-up analysis with larger sample sizes will be required to determine whether the null effects are maintained or whether the failure to identify effects in the present study is a result of type II error.
The non-normal distribution of some of the behavioral scores was a second limitation that is likely related to the small sample sizes. The distribution of PA scores in the follow-up analysis group remained significantly skewed following transformation ($W = 0.899, p = 0.048$; skewness = -0.611). The reported CT/PA relationship should be replicated with a larger sample and normal distribution of scores to confirm that the effect was not driven by the skew of the data.

Finally, statistical thresholding and methods of correction for multiple comparisons in neuroimaging analysis present a challenging limitation as these methods are currently under scrutiny in the literature and consensus on a solution has not yet been reached. Specifically, concerns have recently been raised regarding inflated false-positive rates in studies using voxel- or vertex-wise analysis approaches in functional and structural MRI data (Cox, Chen, Glen, Reynolds, & Taylor, 2017; Eklund, Nichols, & Knutsson, 2016; Greve & Fischl, 2018). Greve and Fischl (2018) have suggested applying more conservative cluster-forming $p$-values ($p < .01, p < .005$, or $p < .001$) or using non-parametric permutation testing for cluster-wise correction for multiple comparisons, however these suggestions were made based on evaluation of group-wise comparison analysis, and optimal methods for continuous analysis have not yet been evaluated. Nonetheless, caution should be taken in interpreting the exploratory results reported herein, as they were reported at a less-stringent cluster-forming threshold of $p < .05$ with Monte Carlo correction for multiple comparisons. Replication of findings must be relied upon to confirm the observed brain-behavior associations.

ROI analysis limitations

The ROI analysis conducted as a first step in this study did not reveal any significant effects. Because significant brain-behavior associations were observed in the exploratory
analysis, it is worth considering limitations specific to the ROI analysis. First, atlas-based anatomical ROIs may not align closely with regions of the cortex that are functionally relevant to the behavior of interest. Second, atlas-based ROIs are fairly large and the statistics derived from the ROIs reflect averages (for cortical thickness) and sums (for surface area) across the entire region, which may wash out effects that are limited to only a portion of the ROI (e.g., only the posterior portion of the OT region). Third, the ROIs were selected based on meta-analyses of case-control studies that included predominantly voxel-based characterization of brain measures, while a continuous approach including surface-based measures of cortical structure was applied in the present study. These methodological differences may reduce precision in replicating regional effects. Notably, two of the clusters of significant association observed in the exploratory analysis were spatially close to or overlapping with selected ROIs: the left M-OT cluster that was associated with LW performance lies adjacent to the selected left OT ROI, and the right PCG/PCS cluster that was associated with PA performance extends into the right IPL ROI. This indicates that slight differences in ROI localization may influence the detection of significant relationships with behavior.

**Conclusions**

The present study revealed positive correlations between properties of cortical structure and individual differences in reading-related skills measured 8-12 months after brain imaging. Accordingly, regional measures of cortical surface structure show promise as salient predictors of reading outcomes over time. Regionally and behaviorally distinct associations with surface area and cortical thickness were observed, such that surface area in right and left hemisphere regions was associated with LW performance, and cortical thickness in the right postcentral
cortex was associated with PA performance. The differential effects show the value of investigating properties of gray matter structure independently, and further research should be conducted to elucidate patterns of behavioral associations that may be specific to surface area and/or cortical thickness. Regions showing significant effects in the present study are consistent with previous reports of gray matter reductions in SRD samples. These findings support behavioral evidence that SRD represents the lower end of a continuum of reading ability rather than a qualitatively independent phenotype (Fletcher, 2009; Francis et al., 2005). The continuous analytic approach applied here may, in fact, more accurately reflect the nature of brain-behavior associations in reading, and further application of this approach is merited.

This work constitutes an important incremental step in establishing approaches to more specifically characterize relationships among brain structure and reading skills beyond the foundation of the existing literature. These methods may be applied to future efforts to identify brain-based predictors of reading skills across a continuum of ability and to assess the stability of structural brain predictors over time. Furthermore, analysis of fine-grained properties of brain structure will be instrumental for linking genes to reading behavior via intermediate phenotypes at the neuroanatomical level, considering the putative distinct genetic underpinnings of cortical thickness and surface area (Panizzon et al., 2009; Winkler et al., 2010). Such application of these structural methods will inform theoretical models of causal pathways from genes to reading ability in individuals with and without SRD.
References


ASSOCIATIONS BETWEEN CORTICAL STRUCTURE AND READING

Metabolism, 110(3), 201–212.


Insights from DTI and VBM at 3.0 T. *Neuropsychologia, 46*, 3170–3178.


Appendix: Supplementary Information

Table S1: Shapiro-Wilk test statistics for measures exhibiting non-normal distributions

<table>
<thead>
<tr>
<th>Analysis Group</th>
<th>Measure (centered &amp; scaled)</th>
<th>W</th>
<th>p</th>
<th>skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concurrent</td>
<td>CTOPP RD (RAN)</td>
<td>0.887</td>
<td>0.06</td>
<td>0.598</td>
</tr>
<tr>
<td></td>
<td>Age at MRI</td>
<td>0.886</td>
<td>0.059</td>
<td>-0.646</td>
</tr>
<tr>
<td>Follow-up</td>
<td>CTOPP PA composite</td>
<td>0.872</td>
<td>0.015</td>
<td>-0.781</td>
</tr>
<tr>
<td></td>
<td>Age at MRI</td>
<td>0.91</td>
<td>0.075</td>
<td>-0.743</td>
</tr>
<tr>
<td></td>
<td>Left TP cortical thickness</td>
<td>0.898</td>
<td>0.045</td>
<td>-1.039</td>
</tr>
</tbody>
</table>

Table S2: Shapiro-Wilk test statistics following transforms for measures exhibiting non-normal distributions

<table>
<thead>
<tr>
<th>Analysis Group</th>
<th>Measure (transformed)</th>
<th>Transform</th>
<th>W</th>
<th>p</th>
<th>skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concurrent</td>
<td>CTOPP RD (RAN)</td>
<td>Box-Cox</td>
<td>0.919</td>
<td>0.189</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>Age at MRI</td>
<td>Box-Cox</td>
<td>0.894</td>
<td>0.077*</td>
<td>-0.559</td>
</tr>
<tr>
<td></td>
<td>Age at MRI</td>
<td>Tukey</td>
<td>0.955</td>
<td>0.608</td>
<td>0.35</td>
</tr>
<tr>
<td>Follow-up</td>
<td>CTOPP PA composite</td>
<td>Box-Cox</td>
<td>0.899</td>
<td>0.048*</td>
<td>-0.611</td>
</tr>
<tr>
<td></td>
<td>CTOPP PA composite</td>
<td>Tukey</td>
<td>0.893</td>
<td>0.0367*</td>
<td>-0.874</td>
</tr>
<tr>
<td></td>
<td>Age at MRI</td>
<td>Box-Cox</td>
<td>0.92</td>
<td>0.114</td>
<td>-0.659</td>
</tr>
<tr>
<td></td>
<td>Left TP cortical thickness</td>
<td>Box-Cox</td>
<td>0.927</td>
<td>0.153</td>
<td>-0.715</td>
</tr>
</tbody>
</table>

*Distribution remained skewed following transform; for CTOPP PA Composite, the Box-Cox transformed variable was used in subsequent analysis because it showed the lowest degree of skew under this transform relative to untransformed and Tukey transformed versions
Figure S1: Density plots of RAN scores in Concurrent group (a) before and (b) after transform

Figure S2: Density plots of MRI age in Concurrent group (a) before and (b) after transform
**Figure S3:** Density plots of PA Composite scores in Follow-up group (a) before and (b) after transform

**Figure S4:** Density plots of MRI Age in Follow-up group (a) before and (b) after transform
Figure S5: Density plots of left TP cortical thickness in Follow-up group (a) before and (b) after transform.