Gray Matter Structure Is Associated with Reading Skill in Typically Developing Young Readers

Meaghan V. Perdue1,2, Joshua Mednick1, Kenneth R. Pugh1,2 and Nicole Landi1,2

1University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT, 06269, USA and 2Haskins Laboratories, 300 George St #900, New Haven, CT 06511, USA

Address correspondence to Meaghan V. Perdue, University of Connecticut, Department of Psychological Sciences, 406 Babbidge Road, Unit 1020, Storrs, CT, 06269, USA. Email: meaghan.perdue@uconn.edu.

Abstract

Research using functional and structural magnetic resonance imaging has identified areas of reduced brain activation and gray matter volume in children and adults with reading disability, but associations between cortical structure and individual differences in reading in typically developing children remain underexplored. Furthermore, the majority of research linking gray matter structure to reading ability quantifies gray matter in terms of volume, and cannot specify unique contributions of cortical surface area and thickness to these relationships. Here, we applied a continuous analytic approach to investigate associations between distinct surface-based properties of cortical structure and individual differences in reading-related skills in a sample of typically developing young children. Correlations between cortical structure and reading-related skills were conducted using a surface-based vertex-wise approach. Cortical thickness in the left superior temporal cortex was positively correlated with word and pseudoword reading performance. The observed positive correlation between cortical thickness in the left superior temporal cortex and reading may have implications for the patterns of brain activation that support reading.

Key words: children, individual differences, neuroimaging, cortical surface area, cortical thickness, reading

Introduction

Reading is a complex skill that depends on fundamental processes in the auditory, visual, and oral language domains (Ehri et al., 2001; Hulme et al. 2012; Preston et al., 2016; Schatschneider et al.2004; Wagner and Torgesen, 1987; Warington and Hulme, 2012). Consistent with this, reading and skill in reading have been associated with functional neural activation across a broad network of regions, including occipito-temporal, superior temporal, posterior parietal, inferior frontal, and supplementary motor regions, and the cerebellum (Christodoulou et al., 2014; D’Mello and Gabrieli, 2018; Hoeff et al., 2006; Maisog et al.2008; Martin et al.2015; Pugh et al., 2001; Richlan et al.2009; Rueckl et al., 2015; Taylor et al. 2013; Turkeltaub et al. 2003). With respect to cortical anatomy (gray matter volume, surface area, and thickness), a small but growing literature has identified reading skill-associated regions that appear consistent with functional studies. Indeed, studies of individuals with specific reading disability (SRD) (or developmental dyslexia) have revealed atypical neural structure in frontal, perisylvian, and occipito-temporal cortices as well as subcortical regions including thalamus and cerebellum (Eckert et al.2016; Hoeff et al., 2007; Krafnick et al. 2014; Linkersdörfer et al. 2012; Richlan et al. 2013; Xia et al. 2016). However, despite perceived consistency among reports on relations between SRD and cortical anatomy, a closer examination of the literature reveals variability among reports of neuroanatomical differences (Ramus et al., 2018). Indeed, three meta-analyses that included overlapping sets of studies identified three different sets of SRD-associated regional reductions in gray matter (Eckert et al., 2016; Linkersdörfer et al., 2014; Richlan et al., 2013). Inconsistency in this body of research may...
be attributed to variability in samples and analysis methods and
calls for replication of previous findings as well as application
of improved methods to characterize gray matter structure in
greater detail than standard gray matter volume measurements
allow (Ramus et al., 2018).
To date, the vast majority of studies relating neural anatomy
to reading have contrasted children with SRD and typically
developing children, rather than examining the relationship
between neural anatomy and reading skills directly. Studies that
compare groups with and without SRD provide an important
foundation for characterizing the neural basis of reading
difficulties, but group contrasts can be problematic because
there is no consensus for identification of SRD and different
methods are utilized to classify children across studies (Fletcher,
2009; Francis et al., 2005). Thus, it remains unknown whether the
same neuroanatomical features associated with SRD are related
to individual differences in reading and reading-related skills
in children with reading ability spanning the low-average to
above-average range. Studies that compare groups with SRD to
both age-matched and reading performance-matched control
groups have resulted in mixed reports of disorder-specific and
performance-related neuroanatomical characteristics (Hoef
et al., 2007; Krafnick et al., 2014; Xia et al., 2016). Using an
individual differences approach can build upon this research
to further test for disorder-independent associations between
reading performance and brain structure in typically developing
samples.
Anatomical studies that have applied both group compari-
sion and individual differences approaches have shown unique
brain-behavior associations in the context of continuous and
group-level analyses and highlight the value of examining pat-
terns of covariance among the variables using a continuous
analytic approach (Jednoróg et al., 2015; Pernet et al., 2009).
For example, research focusing on individual differences in reading
and cortical anatomy in adolescent and adult typical readers
points to primarily positive associations between reading and
related skills and gray matter structure in regions previously
associated with reading and/or SRD (Goldman and Manis, 2013;
He et al., 2013; Johns et al., 2018; Pernet et al.2009; Torre and Eden,
2019; Zhang et al., 2013). With regard to children, the research
targeting individual differences is quite sparse, and there is
considerable variation among studies in terms of age and the
language/writing system of the participants. For example, a study
of Polish, German, and French children ages 8–13 (Jednoróg et al.,
2015) found a positive correlation between reading accuracy and
gray matter volume in the left supramarginal gyrus and negative
correlation with between reading and volume in the left cerebel-
um for typically developing (TD) children, but not SRD children.
A study of Chinese children (ages 10–12) with typical reading
abilities (Xia et al., 2018) found positive correlations between
word reading and cortical thickness in bilateral superior tem-
poral cortex, left supramarginal gyrus, and left inferior tem-
poral gyrus, along with positive correlations between reading
comprehension and cortical thickness in left parahippocampus
and right calcineus fissure. Recently, a large sample (N = 404) study
of English speaking children and adolescents ranging in age
from 6 to 22 years failed to identify any significant associations
between gray matter volume and reading ability in the younger
participants, and effects in the older participants (ages 15–22)
were sex-specific, pointing to complex interactions among brain
structure, age, and sex with regard to reading ability (Torre and
Eden, 2019). These studies provide initial evidence that gray
matter structure in the reading network and its development
are related to individual differences in reading skills from the
first years of reading instruction, but additional work on young
typically developing readers is needed to chart the relationship
between cortical anatomy and reading across development. Fur-
thering this line of research is necessary because continuous
analytic approaches may capture unique associations between
brain and behavior that do not emerge from group designs.

Motivation for the Use of Surface-Based
Models
One way to improve characterization of cortical gray matter
structure and reading ability is to apply a surface-based model
of cortical structure. The surface-based model addresses limita-
tions of voxel-based morphometry (VBM), which quantifies gray
matter volume based on the number of voxels containing gray
matter in 3D space. VBM is appropriate for characterizing the
size and shape of subcortical structures, but it is less appropriate
for characterizing the structure of cortical gray matter, which
is better represented as a sheet with separate dimensions of
thickness and surface area. Accordingly, surface-based modeling
flattens the folds and curves of the cortex into a sheet to
more accurately represent the nature of cortical topography.
Boundaries between the white matter and gray matter, and
between the gray matter and outer dura/cerebrospinal fluid,
called the white matter surface and pial surface, respectively,
are delineated in the MR images, and each surface is modeled
as a triangle mesh (Dale et al. 1999; Fischl et al. 1999). Mea-
sures of several anatomical properties can be derived from this
model. Here, we focus on cortical thickness, a measure of the
distance between corresponding vertices on each surface, and
surface area, a sum of the area of the triangles.Independently
characterizing associations with cortical thickness and surface
area is important because these two properties of brain anatomy
are influenced by distinct sets of genes and neurodevelopmental
processes (Amlien et al., 2016; Lyall et al., 2015; Panizzon
et al., 2009; Rakic, 1995; Wierenga et al. 2014; Winkler et al.,
2010). Decomposing cortical volume into its constituent features
offers the advantage of characterizing the cortical anatomy
with greater specificity and identifying unique contributions
of cortical thickness and surface area to gray matter morphology
associated with a given trait or disorder. This approach has
been effectively utilized to identify separate effects of cortical
thickness and surface area in and samples with schizophrenia
and bipolar disorder and adolescents with a history of low
birth weight (Rimol et al., 2012; Winkler et al., 2018). Several
reports show that gray matter volume and surface area are more
closely related to each other than to cortical thickness (Frye
et al., 2010; Winkler et al., 2010; Yang et al. 2016), indicating that
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 litera-
ture. Several studies that have applied surface-based analysis
to study associations between cortical structure and reading
in adults have shown distinct effects in gray matter volume,
cortical thickness and surface area (Frye et al., 2010; Johns et al.,
2018). These findings point toward some degree of specificity
among features of cortical structure and domains of reading
ability that warrants inclusion of surface area and thickness
separately.
Given the independence of cortical surface area and
thickness characteristics, surface-based analysis techniques

Downloaded from https://academic.oup.com/cercor/advance-article-abstract/doi/10.1093/cercor/bhaa126/5850537 by University of Connecticut user on 11 June 2020
may provide insight to brain-reading relationships beyond what has been revealed through VBM and may help to disambiguate extant findings to 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.

**Current Study**

The limited number of studies on reading-associated individual differences in neural anatomy in young typically developing children and the lack of convergence across extant studies on reading-brain structure relations motivate the present study. Two features make the current study a valuable contribution to the field: 1) the use of continuous sampling to determine the relationship between individual differences in reading and gray matter structure in children with low-average to above average reading ability, rather than identification of regions that differ between children with SRD and TD children, and 2) application of a surface-based model of cortical structure which improves upon VBM approaches by de-conflicting cortical thickness and surface area, which may independently vary as a function of reading skill.

Here, we examined brain structure associated with individual differences in reading skill in typically developing young children using an exploratory vertex-wise surface-based approach in FreeSurfer (Dale et al., 1999; Fischl et al., 1999). This approach affords characterization of distinct properties of cortical surface structure, indexed by cortical thickness and surface area, which are thought to reflect distinct features of the underlying neural architecture and may be independently associated with reading and reading-related skills (Frye et al., 2010; Ma et al., 2015; Wierenga et al., 2014). We investigated associations of cortical structure with word-level reading abilities as well as phenological awareness (PA) and rapid naming, two key predictors of reading ability that may show distinct relationships with brain structure and could contribute to more thorough characterization of the neuroanatomical underpinnings of reading ability (He et al., 2013; Johns et al., 2018). Importantly, we targeted an age range (5–9 years) in which links between individual differences reading ability to brain structure have not been extensively studied. This is significant because findings from older children, adolescents, and adults may not be generalizable to younger readers, as properties of cortical structure show a dynamic developmental trajectory through adolescence (Amlien et al., 2016; Koelschijn and Crone, 2013; Raznahan et al., 2011; Wierenga et al., 2014). Age-specific relationships may be especially relevant in brain regions that show experience-dependent functional specialization for reading that may have consequences for the development of the underlying brain structure, such as the left ventral occipito-temporal cortex (putative visual word form area) (Cohen and Dehaene, 2004; Saygin et al., 2016).

We predicted that cortical surface area would be positively correlated with measures of reading ability in the reading network, particularly in the left temporo-parietal cortex where an association between reading ability and gray matter volume was driven by typically developing children (Jednoróg et al., 2015). We expected consistency between findings in gray matter volume and surface area based on the previously noted relationship between surface area and volume measurements (Frye et al., 2010). Our analysis of cortical thickness was more exploratory, given that relatively few studies have investigated associations between cortical thickness and individual differences in reading ability. We expected to find positive correlations between cortical thickness and reading ability in regions typically associated with reading (e.g., temporo-parietal and occipito-temporal cortex), with additional correlations possible in other regions identified in studies of cortical thickness in SRD, such as frontal regions and regions around the central sulcus (Clark et al., 2014; Williams et al., 2019). Nonetheless, given previous findings of disorder-specific effects in gray matter structure (Hoeft et al., 2007; Krafick et al., 2014; Ma et al., 2015; Xia et al., 2016), we expected that effects in some regions showing gray matter differences in SRD may be absent in our continuous analysis of typically developing readers, though reports of disorder-specific findings are limited and do not provide sufficient evidence upon which to base specific predictions regarding these effects. Finally, we expected that cortical surface area and cortical thickness would show distinct patterns of association with reading-related skills due to their unique genetic and neurodevelopmental underpinnings (Amlien et al., 2016; Lyall et al., 2015; Panizzon et al., 2009; Rakic, 1995; Wierenga et al., 2014; Winkler et al., 2010).

**Methods**

**Participants**

The present study included 76 children (42 females, 34 males; 4.67–9.50 years old at behavioral testing; 69 right-handed; participant characteristics listed in Table 1) drawn from a larger longitudinal study (N = 186) of the behavioral and neural characteristics of reading acquisition from kindergarten through grade 3. The sample included in the present analyses was drawn from a sub-sample of participants who had completed both behavioral testing and magnetic resonance imaging (MRI) scanning at a minimum of one time point (N = 88). Twelve of these subjects were excluded from analysis during MRI quality inspection (detailed below) due to excessive motion artifacts. 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 75 on an age-appropriate test of full-scale IQ: either the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999) or the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (Wechsler, 2002). The racial and ethnic breakdown of the participants was as follows: 55 white, 8 African American, 3 more than 1 race, 5 unknown, 5 not reported; 11 Hispanic or Latino, 61 not Hispanic or Latino, 4 not reported. The study protocol was approved by the (Yale University Human Investigation Committee). Participants who completed the behavioral assessment battery and a structural MRI scan (N = 88) were considered for the present study. Twelve participants were excluded due to excessive motion artifacts in their T1-weighted MR images. Note that a few participants completed more than one scan; for these participants the choice of which scan to include was made based on two factors: 1) the quality of the MRI scan, and 2) scan proximity to behavioral assessment.

---

1 Date of MRI scan acquisition ranged from 5 months before to 9 months after behavioral assessment (M = 2.93, SD = 2.28).
Table 1  Participant demographics and behavioral characteristics

<table>
<thead>
<tr>
<th>Measure</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at MRI (months)</td>
<td>76</td>
<td>82.55</td>
<td>10.85</td>
<td>60–112</td>
</tr>
<tr>
<td>Age at cognitive testing (months)</td>
<td>76</td>
<td>80.22</td>
<td>11.55</td>
<td>56–114</td>
</tr>
<tr>
<td>Time between behavioral assessment and MRI scan (months)</td>
<td>76</td>
<td>2.93</td>
<td>2.28</td>
<td>0–9</td>
</tr>
<tr>
<td>WASI full scale IQ</td>
<td>27</td>
<td>106.56</td>
<td>9.76</td>
<td>85–126</td>
</tr>
<tr>
<td>WJIII Word Attack raw score</td>
<td>74</td>
<td>26.12</td>
<td>9.95</td>
<td>12–77</td>
</tr>
<tr>
<td>WJIII Letter-Word ID raw score</td>
<td>76</td>
<td>12.42</td>
<td>7.12</td>
<td>2–31</td>
</tr>
<tr>
<td>CTOPP2 rapid digits raw score</td>
<td>74</td>
<td>19.86</td>
<td>5.44</td>
<td>4.5–29.5</td>
</tr>
<tr>
<td>CTOPP2 Elision standard score</td>
<td>74</td>
<td>10.97</td>
<td>2.44</td>
<td>5–17</td>
</tr>
<tr>
<td>CTOPP2 Blending words standard score</td>
<td>73</td>
<td>10.89</td>
<td>2.85</td>
<td>5–17</td>
</tr>
<tr>
<td>CTOPP2 rapid digits standard score</td>
<td>74</td>
<td>10.53</td>
<td>1.94</td>
<td>5–14</td>
</tr>
<tr>
<td>WJIII Word Attack standard score</td>
<td>76</td>
<td>113.66</td>
<td>11.68</td>
<td>83–138</td>
</tr>
<tr>
<td>WJIII Letter-Word ID standard score</td>
<td>76</td>
<td>114.20</td>
<td>14.15</td>
<td>84–148</td>
</tr>
</tbody>
</table>

Notes: Brain-behavior analyses were conducted using raw scores, centered and scaled. CTOPP2 PA composite raw score is an average of raw scores on the Elision and Blending subtests. WJIII IV FSIQ composite score is a full-scale IQ composite score, which includes information, similarities, block design, matrix reasoning, picture memory, and digit search subtests. Standard scores are provided for comparison with previous literature.

Cognitive Assessment
Participants completed a battery of standardized assessments of reading and language. Assessments that reflect early individual differences in word reading skill and that have been shown to be strong predictors of later word reading and have been widely studied in previous neuroimaging literature of reading and SRD were selected for the present analysis including:


“Comprehensive Test of Phonological Processing 2 (CTOPP2): Elision and Blending Words” subtests to test PA (Wagner et al. 1999). The Elision subtest 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 PA score for each participant. The Blending words subtest score was unavailable for one participant, so the Elision score was used as the PA score for that child.

“Comprehensive Test of Phonological Processing (CTOPP2): Rapid Digit Naming” subtest (RD) to test rapid automatized naming of digits (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.

Two participants did not complete the CTOPP2 and were excluded from the brain-behavior analysis of PA and RD.

Brain Image Acquisition and Preprocessing
High-resolution T1-weighted 3D MPRAge anatomical images were acquired using a Siemens 3 T Trio MR system (TE = 2.77 ms, TR = 2530 ms; FOV = 256 × 256 voxel matrix; voxel size = 1.0 × 1.0 × 1.0 mm). Preprocessing and analysis of anatomical images was conducted using FreeSurfer v. 5.3 software (Dale et al., 1999) via parallel processing in GNU parallel (Tange, 2018). The automated pipeline for surface-based cortical reconstruction and volumetric segmentation including skull stripping, volumetric labeling, intensity normalization, white matter segmentation, surface atlas registration, surface extraction, and gyral labeling was applied to each participant individually. Individual subjects’ data were resampled onto the Freesurfer average subject (fsaverage) and smoothed at default full width half maximum (FWHM) values of 0, 5, 10, 15, 20, and 25 mm. Intersubject registration to an adult-based average template (i.e., fsaverage) using surface-based registration has been validated in children ages 4–11 and performs well without introducing age-related biases (Ghosh et al., 2010). Data smoothed using the FWHM of 10 mm were used in subsequent analyses. 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. Quality control of preprocessed MR images is particularly important in samples of young children who tend to move during scanning, especially because measures of interest (such as cortical thickness) are known to be underestimated when motion artifacts are present (Reuter et al., 2015), and motion may account for apparently thinner cortex in younger children (Ducharme et al., 2016). We addressed this concern by carefully examining raw and preprocessed MR images and excluding subjects with excessive motion artifacts (N = 12) and manually editing errors in the preprocessed images. Specifically, 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).

Structural MRI Analysis
Whole-brain vertex-wise analyses of relationships among reading-related test scores and cortical thickness (CT) and surface area (SA) were conducted in the Freesurfer neuroimaging analysis suite (Dale et al., 1999; Fischl et al., 1999). CT is...
quantified as the distance between the inner (white matter/gray matter) and outer (gray matter/exterior cerebrospinal fluid) surfaces of the cortex, and is measured at each corresponding vertex across the triangle mesh models of the surfaces. SA is quantified locally by summing the areas of adjacent triangle faces within the triangle mesh. This measurement is conducted in each subject’s native space, allowing for individual variation in the size of each triangle in the mesh. SA is represented at each vertex as one-third of the sum of the faces that share that vertex, which facilitates vertex-wise analysis of SA (Winkler et al., 2018). Our analyses were limited to cortical gray matter and did not include subcortical structure or cerebellum because we were most interested in improving the characterization of cortical gray matter structure using surface-based models that are not appropriate for measuring subcortical structures. Raw scores of the behavioral assessments were entered into the brain-behavior analyses in order to evaluate the association between cortical structure and reading ability, rather than the association between cortical structure and reading ability relative to peers (standard scores). Raw scores were centered by subtracting the group mean from each value and scaled by dividing each value by the group standard deviation using the scale function in R (version 3.5.1; R Core Team, 2016) prior to entry into whole brain analyses to address differences in measurement and scale of the behavioral variables and the brain structure metrics while maintaining the distribution of each variable. This is a necessary step for handling covariates of different measurement scales in Freesurfer. Correlation models were built using Freesurfer’s mri_glmfit function with sex and age at MRI (centered and scaled) as variables of no interest and behavioral test scores (centered and scaled raw scores) as independent variables. Independent models were built for each reading-related measure (LW, WA, PA and RAN) for each structural metric (CT and SA), and were run separately for right and left hemispheres. The statistical models are estimated at each vertex across the triangle mesh. A vertex-wise significance threshold of \( P < 0.01 \) was applied to resulting statistical maps and cluster-wise correction for multiple comparisons was conducted using Freesurfer’s mri_glmfit-sim function to perform Monte Carlo simulations of white noise on the cortical surface. Monte Carlo simulations (10,000 iterations) were conducted with a 10 mm FWHM smoothing kernel and a vertex-wise cluster forming threshold of \( P < 0.01 \). The cluster-wise \( P \)-value of \( P < 0.05 \) was corrected for multiple comparisons induced by running the analyses separately in each hemisphere (2 tests) using Bonferroni correction. Results were evaluated at a corrected cluster-wise threshold of \( P < 0.05 \).

In studies targeting group-level differences in gray matter volume, it is common practice to include total brain volume (TBV) or intracranial volume (ICV) as a covariate to ensure that local group differences are not simply driven by group differences in TBV, but we deemed this inappropriate for our analyses for several reasons. First, we were interested in examining individual differences in neuroanatomy, and including a correction for TBV/ICV would remove variance of interest to our research aims. Second, prior research has shown that normalized CT measures (using global mean thickness or intracranial volume) gave significantly lower prediction accuracies in a classification study of dementia (Westman et al., 2013). Additionally, it has been suggested that correcting a 1D measure such as CT with a 3D measure such as intracranial volume could result in over-correction (Wierenga et al., 2014). Nonetheless, we repeated the analyses with estimated total intracranial volume as an additional covariate for comparison with published literature.

**Results**

Behavioral characteristics of the sample are reported in Table 1. Raw scores (centered and scaled) were entered into the brain-behavior analyses. Standard scores are reported for purposes of showing normed sample characteristics and comparison with published literature.

Correlations among the behavioral variables were tested using the centered and scaled raw scores. Significant correlations among all behavioral variables were found and are reported in Table 2.

**Cortical Thickness Analyses**

Whole-brain vertex-wise analyses controlling for age and sex revealed significant positive correlations with cortical thickness in the left superior temporal cortex, including Heschl's Gyrus, for word reading (peak vertex: \( r = 0.570, \ P < 0.001 \); cluster-corrected \( P = 0.0014 \)) and pseudoword reading (peak vertex: \( r = 0.486, \ P < 0.001 \); cluster-corrected \( P = 0.031 \)). Significant clusters are reported in Table 3 and depicted in Figure 1. Scatter plots showing the correlation between reading scores and the cortical thickness of the peak vertex of each cluster are shown in Figure 1b and c. Scatter plots showing the correlation between reading scores and mean thickness of each cluster are included in supplementary materials (Supplementary Figure S1). Note that the effect of WA may be considered marginally significant based on recent recommendations for cluster thresholding data with vertex-wise \( P < 0.01 \) at a cluster-wise threshold of \( P < 0.02 \) (Greve and Fischl, 2018). No significant associations between cortical thickness and PA or rapid naming scores were observed.

**Surface Area Analyses**

Whole-brain vertex-wise analysis of surface area controlling for age and sex and cluster-corrected using Monte Carlo simulation did not reveal significant associations with the any of the reading-related measures.

---

3 These models were repeated in the sub-sample of right-handed children (\( N = 69 \)), and the correlation between cortical thickness and LW remained significant at the cluster-corrected threshold, but the cluster showing a correlation between cortical thickness and WA was not significant with cluster-correction, but remained present at an uncorrected vertex-wise threshold of \( P < 0.01 \). Furthermore, including estimated total intracranial volume as an additional covariate in the analysis did not affect the results.

4 A separate model was run in addition including age, sex and total surface area as covariates to control for potential influences of global surface area, and this model also resulted in null effects for surface area. Likewise, a model including age, sex and estimated total intracranial volume resulted in null effects for surface area.
Discussion

In the present study, we investigated associations between brain anatomy and individual differences in reading-related skills in typically developing young children using a surface-based model to independently analyze distinct features of cortical structure: cortical thickness and surface area. Vertex-wise analyses of cortical surface structure showed that greater cortical thickness in the left superior temporal cortex was associated with better performance on word and pseudoword reading in children during the years of initial reading instruction. Prior work linking reading ability to brain anatomy is largely based on studies contrasting groups with and without SRD, and these findings expand upon that literature to show some consistency with individual differences in typical readers.

The present findings are consistent with several previous reports of atypical neuroanatomy in the left superior temporal cortex in SRD relative to typical readers. Clark et al. (2014) found reduced cortical thickness in the left Heschl’s gyrus in pre-readers who subsequently went on to develop RD, and the left anterior superior temporal cortex/middle temporal cortex in 11–12-year-old children with RD. Recently, Williams et al. (2018) identified reduced cortical thickness in the left superior temporal sulcus of children and adolescents with RD. Furthermore, reduced gray matter volume in the bilateral superior temporal cortex has been associated with RD using VBM (see Richlan et al. 2013 for a meta-analysis). Importantly, we observed an association between left superior temporal cortical thickness and individual differences in reading ability in a group of children with a range of reading skill, most of whom scored in the low average to above average range on tests of word and pseudoword reading. This indicates that cortical structure in the left superior temporal cortex is not specifically disrupted in SRD, but rather is associated with reading skill in a continuous fashion, such that better readers have thicker cortex in this region. This finding is consistent with a positive correlation between word reading and cortical thickness in the bilateral superior temporal gyrus, among several other regions, reported in 10–12-year-old typically developing Chinese children (Xia et al., 2018) and a positive correlation between gray matter volume in the left superior temporal gyrus and subsequent gains in reading proficiency over 1 year in typically developing German children (Linkensdörfer et al., 2014). The convergence of these findings in children with and without SRD and across different languages supports a universal role of the superior temporal cortex in reading (Rueckl et al., 2015).

With respect to mechanism, cortical thickness in the left superior temporal cortex is likely related to reading ability through its functional role in auditory processing of speech, which is important for the development of phonological analysis (Hickok and Poeppel, 2007). Indeed, extant work finds that basic auditory processing and phonological processing contribute to skilled reading (Ahissar et al. 2000; Vellutino and Scanlon, 1987). Functional neuroimaging research shows that children who read at typical and advanced levels engage the left superior temporal cortex during reading and reading-related tasks (Church et al. 2008; Chyl et al. 2019; Hoeft et al., 2006; Shaywitz et al., 2002; Yamada et al., 2011), and children with SRD tend to show reduced activation in this

---

Table 2 Correlations among behavioral variables

<table>
<thead>
<tr>
<th>Rapid digit naming</th>
<th>Word attack</th>
<th>Letter-word ID</th>
<th>PA</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>-0.59</td>
<td>-0.66</td>
<td>-0.65</td>
</tr>
<tr>
<td>P</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>n</td>
<td>74</td>
<td>74</td>
<td>74</td>
</tr>
</tbody>
</table>

Notes: Pearson’s correlations were conducted using raw scores, centered, and scaled. Correlations were corrected for multiple comparisons using false discovery rate.

Table 3 Reading-related measures associated with regional cortical structure: cluster and peak vertex statistics

<table>
<thead>
<tr>
<th>Cognitive measure</th>
<th>Structural metric</th>
<th>Cluster (LH)</th>
<th>Cluster Size (mm²)</th>
<th>Cluster P</th>
<th>Peak MNI coordinates</th>
<th>Peak vertex statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X        Y        Z</td>
<td>F        P        r</td>
</tr>
<tr>
<td>LW</td>
<td>Thickness</td>
<td>superior temporal</td>
<td>1134.58</td>
<td>0.0014</td>
<td>-57.4  -10.4  0.9</td>
<td>33.664  &lt;0.001  0.57</td>
</tr>
<tr>
<td>WA</td>
<td>Thickness</td>
<td>superior temporal</td>
<td>724.28</td>
<td>0.031</td>
<td>-59.2  -9.7   0.6</td>
<td>21.674  &lt;0.001  0.486</td>
</tr>
</tbody>
</table>

Notes: Results reported for a vertex-wise threshold P < 0.01, with Monte Carlo simulation cluster-wise corrected P < 0.05. Coordinates reported for vertex of strongest effect in MNI 305 space. Clusters identified according to the Desikan-Killiany atlas (Desikan et al., 2006). LH, left hemisphere; LW, Woodcock Johnson III Letter-Word subtest; WA, Woodcock Johnson III Word Attack subtest; Raw scores from cognitive measures centered and scaled prior to analysis.
Figure 1. (a) Results from whole-brain vertex-wise analysis showing positive correlations between cortical thickness in the superior temporal cortex and reading performance, controlling for age, and sex projected onto the “fsaverage” inflated left hemisphere surface from Freesurfer. Results reported at a vertex-wise threshold of $P < 0.01$, corrected for multiple comparisons with a cluster-wise threshold of $P < 0.05$. (b) Scatterplot depicting positive association between WJIII Letter-Word ID score (raw score, centered and scaled) and cortical thickness at the peak vertex of the significant cluster (yellow, including overlap with red) identified in the vertex-wise analysis. Values shown are residuals controlling for age and sex. (c) Scatterplot depicting positive association between WJIII Word Attack Score (raw score, centered, and scaled) and cortical thickness at the peak vertex of the significant cluster (red) identified in the vertex-wise analysis. Values shown are residuals controlling for age and sex.

region in comparison (e.g., Shaywitz et al., 2002). Moreover, activation of the bilateral superior temporal gyrus during reading was positively associated with reading ability in 5–8-year-old children, supporting functional relevance of this region for reading in the age range that we studied (Chyl et al., 2018). Accordingly, reduced cortical thickness in the superior temporal cortex may reflect reduced neuroanatomical resources for reading and its constituent auditory processes (Clark et al., 2014).

In our study, cortical thickness in the left superior temporal cortex was associated with skill in reading both real words and pseudowords, indicating that this region is associated with phonological decoding. Notably, the correlation between raw scores on word reading (WJIII Letter-Word ID) and pseudoword reading (WJIII Word Attack) is quite high in this sample ($r = 0.93$), which may indicate that these children process both words and pseudowords in a similar manner, and the overlapping effects observed in the brain likely arise from similar underlying neural mechanisms. Having a thinner superior temporal cortex during the early years of reading instruction when developing phonological analysis abilities must be linked to printed text may be detrimental to developing reading skills. Indeed, Clark et al., (2014) finding of reduced cortical thickness of Heschl’s gyrus in pre-readers suggests that reduced cortical thickness in the superior temporal cortex precedes reading difficulties, though causal conclusions cannot be drawn from these studies. Furthermore, the association between cortical thickness and neural functioning as measured by MRI is not well understood, and further research is needed to investigate mechanisms underlying this relationship.

The mechanisms underlying the link between reading and gray matter structure in the superior temporal cortex are unknown. A recent neuroimaging genetics study revealed that the minor allele of one variant on RBFOX2, a gene involved in brain development that has been linked to reading and language traits, was associated with reduced cortical thickness in the bilateral superior temporal gyrus, along with several other regions in the reading network (Gialluisi et al., 2014, 2017). Associations with other SRD candidate genes and/or genes that have not yet been implicated in reading ability are possible, as the application of neuroimaging genetics methods to investigate reading and language traits remains in its infancy. Future work using this approach to is needed to establish convergent links between genes, brain structure, and reading ability.

The observed relationship with reading ability in the left superior temporal cortex converges with several previous reports of gray matter reductions in SRD (Richlan et al., 2013); however, we did not find an association between reading
ability and cortical structure in other regions that have been implicated in SRD, such as the left occipito-temporal cortex or the left inferior frontal gyrus. Given findings in adults and older children that show left occipito-temporal CT correlations with word reading and print exposure (Goldman and Manis, 2013; Johns et al., 2018; Xia et al., 2018; Zhang et al., 2013), it is possible that associations between left occipito-temporal neuroanatomy and reading ability emerge with age- and reading experience-driven plasticity. This is consistent with a view of increasing functional specialization of this region for word reading with age and reading experience (Brem et al., 2010; Centanni et al., 2018; James, 2010; Maurer et al., 2011; Pleisch et al., 2015). Notably, previous studies in pre-readers at risk of SRD or who went on to develop SRD have shown reduced CT and gray matter volume in left occipito-temporal regions, so the absence of findings in this region in our sample may indicate that those reductions are specific to SRD and/or SRD-risk or only detectable when more severe reading difficulties are represented in the sample at this early age range (Clark et al., 2014; Raschle et al. 2011). Regarding the left inferior frontal gyrus, reading associations with gray matter structure in adults have been specific to measures of print exposure (Goldman and Manis, 2013; Johns et al., 2018), and functional differences in this region in SRD are often observed in older children and adults (Maisog et al. 2008; Richlan et al., 2009), but not young children with an elevated risk of SRD (Vandermoten et al. 2016).

Intriguingly, Clark et al. (2014) found CT reductions in children with SRD at age 11, but not in retrospective analysis of the same children scanned at the pre-reading stage. Thus, engagement of the left inferior frontal gyrus may also develop with reading experience. Nonetheless, our null findings do not allow us to conclude that relationships among cortical surface structure and reading-related abilities do not exist in young children, as small or medium effects could be present but not detected in our study due to the sample size and/or threshold applied for cluster-based correction (see Noble et al. 2020 for a discussion of sensitivity in cluster-level thresholding of neuroimaging data).

We also did not find any significant associations between reading-related skills and cortical surface area. This was surprising given the previously noted correspondence between surface area and volume (Frye et al., 2016; Winkler et al., 2010; Yang et al., 2016). One possible explanation is that atypical neuroanatomy in other regions in the reading network and/or individual differences in surface area may only be associated with reading difficulties, rather than individual differences among typically developing readers—although our sample included low-average performing children and a couple of “poor readers,” it did not include any very poor readers. Another possibility is that structural associations with other properties of reading network anatomy emerge with reading experience, and the children in our sample were too young to show these relationships. Torre and Eden (2019) failed to find continuous associations between gray matter volume and reading ability in typically developing children, but identified sex-driven associations between reading and gray matter volume in the left fusiform gyrus and right superior temporal gyrus in older adolescents and adults. In contrast, we observed brain-reading correlations in young children that were limited to measures of cortical thickness, which may not have been well represented in the volumetric approach applied by Torre and Eden (2019). Across studies, the pattern of associations between reading and cortical structure point toward a complex relationship among reading ability, cortical structure, and development in which specific features of cortical neuroanatomy (structural properties or brain regions) are sensitive to individual differences in reading ability, while others are sensitive to group differences, and distinct effects may emerge based on age/experience or sex. However, larger scale studies that include both SRD and typically developing children across a wide age range are needed to confirm these speculations.

Limitations
The surface-based models that we have used do not include subcortical brain structures, so our findings should be integrated with research focused on associations between reading and anatomy of subcortical structures and cerebellum, which can be characterized using voxel-based volumetric approaches, in order to contribute to a comprehensive understanding of the neuroanatomical foundations of reading. Further, as noted above, our sample was limited to children with low-average to above-average reading skills, so the results do not account for potential SRD-specific effects.

Conclusion
We identified positive correlations among cortical thickness in the left superior temporal cortex and individual differences in both word and pseudoword reading performance in young children. These findings support the application of an individual differences approach to studying associations between reading and cortical structure, and examining distinct properties of cortical structure independently. Given the complex genetic underpinnings of SRD (Mascheretti et al., 2017) and the distinct mechanisms and genetics of cortical thickness and surface area development (Panizzon et al., 2009; Rakic, 1995), it will be important to understand differential relationships of cortical thickness and surface area with reading ability to inform models of causal pathways from genes through neurobiology to reading phenotypes. In the context of the extant literature, these findings highlight the complexity of associations between brain structure and reading ability and call for careful investigation of these relationships that account for the heterogeneity inherent to complex cognitive traits (Ramus et al., 2018).

Supplementary Material
Supplementary material is available at Cerebral Cortex online.

Acknowledgements
The authors would like to acknowledge the research team at Haskins Laboratories who contributed to data collection and the families who dedicated their time to participate in this study. Conflict of interest: None declared.

Funding
This work was supported by National Institutes of Health (grant numbers 01 HD001994-46 and NIH 50HD052120-11). and the
National Science Foundation (grant numbers DGE-1144399 and DGE-1747453).

References


Vellutino FR, Scanlon DM. 1987. Phonological coding, phonological awareness, and reading ability: evidence from a


