Cerebral Cortex, 2020;00: 1–11

doi: 10.1093/cercor/bhaa126
Original Article

ORIGINAL ARTICLE

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

Meaghan V. Perdue[1,2], Joshua Mednick[1], Kenneth R. Pugh[1,2] and Nicole Landi[1,2]

[1]University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT, 06269, USA and [2]Haskins 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; Warmington 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; Hoeft 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; Hoeft 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 (Hoeft 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 comparison 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 patterns 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 cerebellum 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 temporal cortex, left supramarginal gyrus, and left inferior temporal gyrus, along with positive correlations between reading comprehension and cortical thickness in left parahippocampus and right calcarine fissure. Finally, a recent large ($N = 404$) 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. Furthering 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 limitations 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). Measures 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 literature. 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 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-conflating 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 phonological 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; Koolschijn 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 associations 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. 2018). Nonetheless, given previous findings of disorder specific effects in gray matter structure (Hoeft et al., 2007; Krafnick 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].

---

[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

Measure	*n*	Mean	SD	Range
Age at MRI (months)	76	82.55	10.85	60–112
Age at cognitive testing (months)	76	80.22	11.55	56–114
Time between behavioral assessment and MRI scan (months)	76	2.93	2.28	0–9
WPPSI IV FSIQ composite score	47	112.62	12.61	78–133
WASI full scale IQ	27	106.56	9.76	85–126
CTOPP2 rapid digits raw score	74	26.12	9.95	12–77
WJIII Word Attack raw score	76	12.42	7.12	2–31
WJIII Letter-Word ID raw score	76	34.04	11.88	12–64
CTOPP2 PA composite raw score	74	19.86	5.44	4.5–29.5
CTOPP2 Elision standard score	74	10.97	2.44	5–17
CTOPP2 Blending words standard score	73	10.89	2.85	5–17
CTOPP2 rapid digits standard score	74	10.53	1.94	5–14
WJIII Word Attack standard score	76	113.66	11.68	83–138
WJIII Letter-Word ID standard score	76	114.20	14.15	84–148

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. WPPSI IV FSIQ composite score is a full-scale IQ composite score, which includes information, similarities, block design, matrix reasoning, picture memory, and bug 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:

"Woodcock-Johnson III Tests of Achievement (WJIII): Letter-Word Identification" subtest (LW) to test untimed sight-word reading ability (Woodcock et al. 2001).

"Woodcock-Johnson III Tests of Achievement (WJIII): Word Attack" subtest (WA) to test untimed decoding of phototactically plausible non-words (Woodcock et al., 2001).

"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 used in 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[2]. 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 $P = 0.031$)[3]. 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 1*b* 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[4] and cluster-corrected using Monte Carlo simulation did not reveal significant associations with the any of the reading-related measures.

---

[2] Due to variability among participants in timing between behavioral testing and MRI scanning sessions, the correlation models were also run with the number of months from behavioral testing session to MRI scanning session included as a nuisance variable. The results from these models were consistent with the original findings, so we report only the results from the original models.

[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.

Table 2 Correlations among behavioral variables

	Rapid digit naming	Word attack	Letter-word ID	PA
Rapid digit naming				
r	1	-0.59	-0.66	-0.65
P		<0.001	<0.001	<0.001
n		74	74	74
Word attack				
r		1	0.93	0.76
P			<0.001	<0.001
n			76	74
Letter-word ID				
r			1	0.75
P				<0.001
n				74
PA				
r				1
P				
n				

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

Cognitive measure	Structural metric	Cluster (LH)	Cluster Size (mm^2)	Cluster *P*	Peak MNI coordinates			Peak vertex statistics		
					X	Y	Z	*F*	*P*	*r*
LW	Thickness	superior temporal	1134.58	0.0014	-57.4	-10.4	0.9	33.664	<0.001	0.57
WA	Thickness	superior temporal	724.28	0.031	-59.2	-9.7	0.6	21.674	<0.001	0.486

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, Woodock Johnson III Word Attack subtest; Raw scores from cognitive measures centered and scaled prior to analysis.

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 prereaders 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 (Linkersdörfer 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

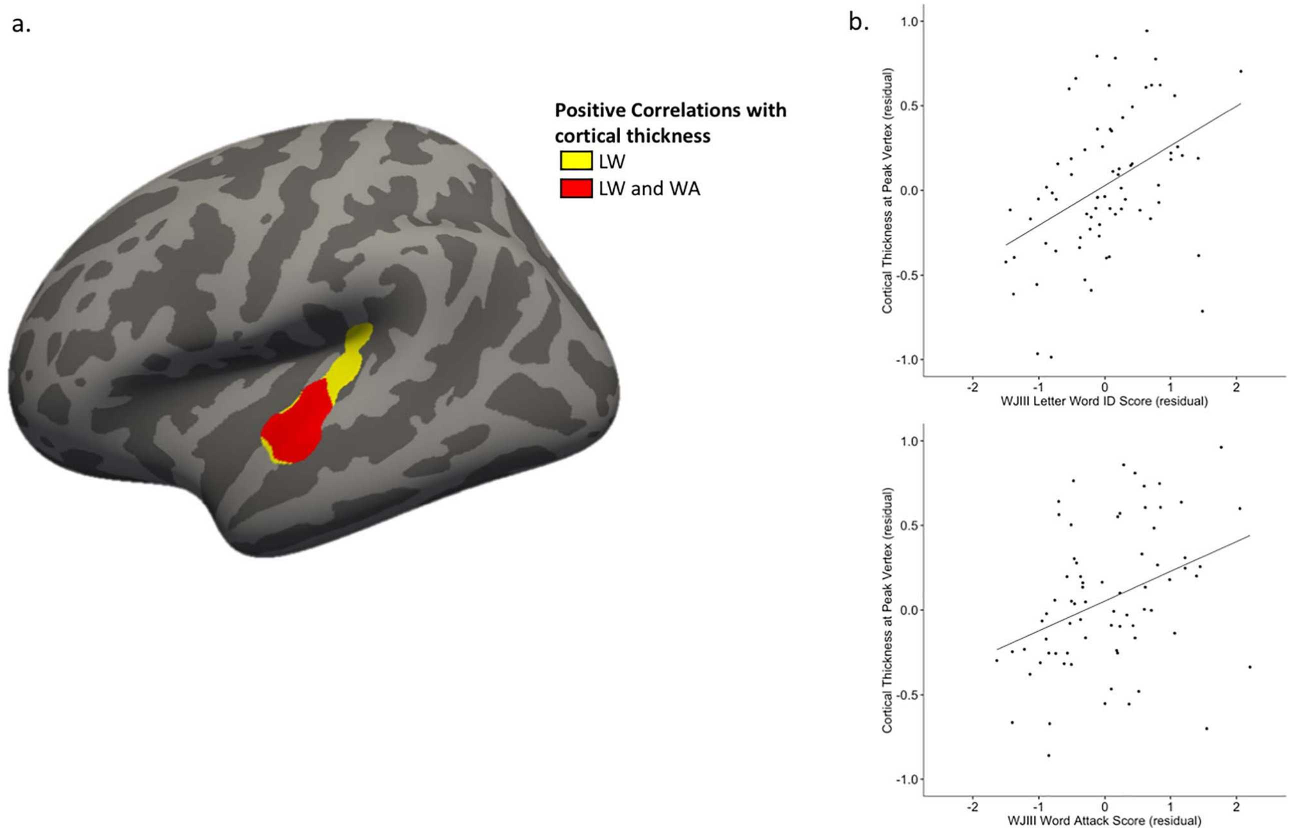

Figure 1. (*a*) Results from whole-brain vertex-wise analysis showing positive correlations between cortical thickness in the superior temporal cortex and reading performance, controlling for age, and sex projected onto the “fsaverage” inflated left hemisphere surface from Freesurfer. Results reported 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., 2019). 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 (Vandermosten et al. 2016). Interestingly, 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., 2010; Winkler et al., 2010; Yang et al., 2016). One possible explanation is that atypical neuroanatomy of 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 P01 HD001994-46 and NIH P50HD052120-11) and the

National Science Foundation (grant numbers DGE-1144399 and DGE-1747453).

References

Ahissar M, Protopapas A, Reid M, Merzenich MM. 2000. Auditory processing parallels reading abilities in adults. *Proc Natl Acad Sci USA*. 97(12):6832–6837.

Amlien IK, Fjell AM, Tamnes CK, Grydeland H, Krogsrud SK, Chaplin TA, MGP R, Walhovd KB. 2016. Organizing principles of human cortical development—thickness and area from 4 to 30 years: insights from comparative primate neuroanatomy. *Cereb Cortex*. 26:257–267.

Brem S, Bach S, Kucian K, Guttorm TK, Martin E, Lyytinen H, Brandeis D, Richardson U. 2010. Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proc Natl Acad Sci*. 107(17):7939–7944.

Centanni TM, Norton ES, Park A, Beach SD, Halverson K, Ozernov-Palchik O, Gaab N, Gabrieli J. 2018. Early development of letter specialization in left fusiform is associated with better word reading and smaller fusiform face area. *Dev Sci*. 21(5):e12658.

Christodoulou JA, Del Tufo SN, Lymberis J, Saxler PK, Ghosh SS, Triantafyllou C, Whitfield-Gabrieli S, Gabrieli JD. 2014. Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLoS One*. 9(7). doi: 10.1371/journal.pone.0100552.

Church JA, Coalson RS, Lugar HM, Petersen SE, Schlaggar BL. 2008. A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. *Cereb Cortex*. 18(9):2054–2065.

Chyl K, Kossowski B, Dębska A, Łuniewska M, Banaszkiewicz A, Żelechowska A, Frost SJ, Mencl WE, Wypych M, Marchewka A, et al. 2018. Prereader to beginning reader: changes induced by reading acquisition in print and speech brain networks. *J Child Psychol Psychiatry Allied Discip*. 59(1):76–87.

Chyl K, Kossowski B, Dębska A, Łuniewska M, Marchewka A, Pugh KR, Jednoróg K. 2019. Reading acquisition in children: developmental processes and dyslexia-specific effects. *J Am Acad Child Adolesc Psychiatry*. 58(10):948–960.

Clark KA, Helland T, Specht K, Narr KL, Manis FR, Toga AW, Hugdahl K. 2014. Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain*. 137:3136–3141.

Cohen L, Dehaene S. 2004. Specialization within the ventral stream: the case for the visual word form area. *Neuroimage*. doi: 10.1016/j.neuroimage.2003.12.049.

D'Mello AM, Gabrieli JDE. 2018. Cognitive neuroscience of dyslexia. *Lang Speech Hear Serv Sch*. 49(4):798.

Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis I. Segmentation and surface reconstruction. *Neuroimage*. 9:179–194.

Ducharme S, Albaugh MD, Nguyen T-V, Hudziak JJ, Mateos-Pérez JM, Labbe A, Evans AC, Karama S, Brain Development Cooperative Group. 2016. Trajectories of cortical thickness maturation in normal brain development—the importance of quality control procedures. *Neuroimage*. 125:267–279.

Eckert MA, Berninger VW, Vaden KIJ, Gebregziabher M, Tsu L. 2016. Gray matter features of reading disability: a combined meta-analytic and direct analysis approach. *ENeuro*. 3(1): 1–15.

Ehri L, Nunes S, Willows DM, Valeska Schuster B, Yaghoub-Zadeh Z, Shanahan T. 2001. Phonemic awareness instruction helps students learn to read: evidence from the National Reading Panel's meta-analysis. *Read Res Q*. 36(3):250–287.

Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*. 9:195–207.

Fletcher JM. 2009. Dyslexia: the evolution of a scientific concept. *J Int Neuropsychol Soc*. 15(4):501–508.

Francis DJ, Fletcher JM, Stuebing KK, Lyon GR, Shaywitz BA, Shaywitz SE. 2005. Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *J Learn Disabil*. 38(2):98–108.

Frye RE, Liederman J, Malmberg B, McLean J, Strickland D, Beauchamp MS. 2010. Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. *Cereb Cortex*. 20(11):2625–2635.

Ghosh SS, Kakunoori S, Augustinack J, Nieto-Castanon A, Kovelman I, Gaab N, Christodoulou JA, Triantafyllou C, Gabrieli JD, Fischl B. 2010. Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11years of age. *Neuroimage*. 53(1):85–93.

Gialluisi A, Newbury DF, Wilcutt EG, Olson RK, DeFries JC, Brandler WM, Pennington BF, Smith SD, Scerri TS, Simpson NH, et al. 2014. Genome-wide screening for DNA variants associated with reading and language traits. *Genes Brain Behav*. 13(7):686–701.

Gialluisi A, Guadalupe T, Francks C, Fisher SE. 2017. Neuroimaging genetic analyses of novel candidate genes associated with reading and language. *Brain Lang*. 172:9–15.

Goldman JG, Manis FR. 2013. Relationships among cortical thickness, reading skill, and print exposure in adults. *Sci Stud Read*. 17(3):163–176.

Greve DN, Fischl B. 2018. False positive rates in surface-based anatomical analysis. *Neuroimage*. 171:6–14.

He Q, Xue G, Chen C, Chen C, Lu Z-L, Dong Q. 2013. Decoding the neuroanatomical basis of reading ability: a multivoxel morphometric study. *J Neurosci*. 33(31):12835–12843.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci*. 8:393–402.

Hoeft F, Hernandez A, McMillon G, Taylor-Hill H, Martindale JL, Meyler A, Keller TA, Siok WT, Deutsch GK, Just MA, et al. 2006. Neural basis of dyslexia: a comparison between dyslexic and nondyslexic children equated for reading ability. *J Neurosci*. 26(42):10700–10708.

Hoeft F, Meyler A, Hernandez A, Juel C, Taylor-Hill H, Martindale JL, Merzenich MM. 2007. Functional and morphometric brain dissociation between dyslexia and reading ability. *Proc Natl Acad Sci*. 104(10):4234–4239.

Hulme C, Bowyer-Crane C, Carroll JM, Duff FJ, Snowling MJ. 2012. The causal role of phoneme awareness and letter-sound knowledge in learning to read: combining intervention studies with mediation analyses. *Psychol Sci*. 23(6):572–577.

James KH. 2010. Sensori-motor experience leads to changes in visual processing in the developing brain. *Dev Sci*. 13(2):279–288.

Jednoróg K, Marchewka A, Altarelli I, Monzalvo Lopez AK, van Ermingen-Marbach M, Grande M, Grabowska A, Heim S, Ramus F. 2015. How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? Insights from a large-scale voxel-based morphometry study. *Hum Brain Mapp*. 36(5). doi: 10.1002/hbm.22734.

Johns CL, Jahn AA, Jones HR, Kush D, Molfese PJ, Van Dyke JA, Magnuson JS, Tabor W, Mencl WE, Shankweiler DP, et al.

2018. Individual differences in decoding skill, print exposure, and cortical structure in young adults. *Language, Cognition and Neuroscience*. 1–21. doi: 10.1080/23273798.2018.1476727.

Koolschijn PCMP, Crone EA. 2013. Sex differences and structural brain maturation from childhood to early adulthood. *Dev Cogn Neurosci*. 5:106–118.

Krafnick AJ, Flowers DL, Luetje MM, Napoliello EM, Eden GF. 2014. An investigation into the origin of anatomical differences in dyslexia. *J Neurosci*. 34(3):901–908.

Linkersdörfer J, Jurcoane A, Lindberg S, Kaiser J, Hasselhorn M, Fiebach CJ, Lonnemann J. 2014. The association between gray matter volume and reading proficiency: a longitudinal study of beginning readers. *J Cogn Neurosci*. 27(2):308–318.

Linkersdörfer J, Lonnemann J, Lindberg S, Hasselhorn M, Fiebach CJ. 2012. Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: an ALE meta-analysis. *PLoS One*. 7(8):1–10.

Lyall AE, Shi F, Geng X, Woolson S, Li G, Wang L, Hamer RM, Shen D, Gilmore JH. 2015. Dynamic development of regional cortical thickness and surface area in early childhood. *Cereb Cortex*. 25:2204–2212.

Ma Y, Koyama MS, Milham MP, Castellanos FX, Quinn BT, Pardoe H, Wang X, Kuzniecky R, Devinsky O, Thesen T, et al. 2015. Cortical thickness abnormalities associated with dyslexia, independent of remediation status. *NeuroImage Clin*. 7. doi: 10.1016/j.nicl.2014.11.005.

Maisog JM, Einbinder ER, Flowers DL, Turkeltaub PE, Eden GF. 2008. A meta-analysis of functional neuroimaging studies of dyslexia. *Ann N Y Acad Sci*. 1145(1):237–259.

Martin A, Schurz M, Kronbichler M, Richlan F. 2015. Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum Brain Mapp*. 36(5):1963–1981.

Mascheretti S, De Luca A, Trezzi V, Peruzzo D, Nordio A, Marino C, Arrigoni F. 2017. Neurogenetics of developmental dyslexia: from genes to behavior through brain neuroimaging and cognitive and sensorial mechanisms. *Transl Psychiatry*. 7:1–15.

Maurer U, Schulz E, Brem S, der Mark S, Bucher K, Martin E, Brandeis D. 2011. The development of print tuning in children with dyslexia: evidence from longitudinal ERP data supported by fMRI. *Neuroimage*. 57(3):714–722.

Noble S, Scheinost D, Constable RT. 2020. Cluster failure or power failure? Evaluating sensitivity in cluster-level inference. *NeuroImage*. 209. doi: 10.1016/j.neuroimage.2019.116468.

Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, Jacobson K, Lyons MJ, Grant MD, Franz CE, et al. 2009. Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex*. 19(11):2728–2735.

Pernet C, Andersson J, Paulesu E, Demonet JF. 2009. When all hypotheses are right: a multifocal account of dyslexia. *Hum Brain Mapp*. 30(7):2278–2292.

Pleisch G, Karipidis II, Brauchli C, Röthlisberger M, Hofstetter C, Stämpfli P, Walitza S, Brem S. 2019. Emerging neural specialization of the ventral occipitotemporal cortex to characters through phonological association learning in preschool children. *Neuroimage*. 189. doi: 10.1016/j.neuroimage.2019.01.046.

Preston JL, Molfese PJ, Frost SJ, Mencl WE, Fulbright RK, Hoeft F, Landi N, Shankweiler D, Pugh KR. 2016. Print-speech convergence predicts future reading outcomes in early readers. *Psychol Sci*. 27(1). doi: 10.1177/0956797615611921.

Pugh KR, Mencl WE, Jenner AR, Katz L, Frost SJ, Lee JR, Shaywitz S, Shaywitz BA. 2001. Neurobiological studies of reading and reading disability. *J Commun Disord*. 34(6):479–492.

Rakic P. 1995. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci*. 18(9):383–388.

Ramus F, Altarelli I, Jednoróg K, Zhao J, Scotto di Covella L. 2018. Neuroanatomy of developmental dyslexia: pitfalls and promise. *Neurosci Biobehav Rev*. 84:434–452.

Raschle NM, Chang M, Gaab N. 2011. Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage*. 57:742–749.

Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, Clasen L, Gogtay N, Giedd JN. 2011. How does your cortex grow? *J Neurosci*. 31(19):7174–7177.

Reuter M, Tisdall MD, Qureshi A, Buckner RL, van der Kouwe AJW, Fischl B. 2015. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *Neuroimage*. 107:107–115.

Richlan F, Kronbichler M, Wimmer H. 2009. Functional abnormalities in the dyslexic brain: a quantitative meta-analysis of neuroimaging studies. *Hum Brain Mapp*. 30(10):3299–3308.

Richlan F, Kronbichler M, Wimmer H. 2013. Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. *Hum Brain Mapp*. 34(11):3055–3065.

Rimol LM, Nesvåg R, Hagler DJ, Bergmann Ø, Fennema-Notestine C, Hartberg CB, Haukvik UK, Lange E, Pung CJ, Server A, et al. 2012. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. *Biol Psychiatry*. 71(6):552–560.

Rueckl JG, Paz-Alonso PM, Molfese PJ, Kuo W-J, Bick A, Frost SJ, Hancock R, Wu DH, Mencl WE, Duñabeitia JA, et al. 2015. Universal brain signature of proficient reading: evidence from four contrasting languages. *Proc Natl Acad Sci*. 112(50):15510–15515.

Saygin ZM, Osher DE, Norton ES, Youssoufian DA, Beach SD, Feather J, Gaab N, Gabrieli JD, Kanwisher N. 2016. Connectivity precedes function in the development of the visual word form area. *Nat Neurosci*. 19(9):1250–1255.

Schatschneider C, Fletcher JM, Francis DJ, Carlson CD, Foorman BR. 2004. Kindergarten prediction of reading skills: a longitudinal comparative analysis. *J Educ Psychol*. 96(2):265–282.

Shaywitz BA, Shaywitz SE, Pugh KR, Mencl WEE, Fulbright RK, Skudlarski P, Constable RT, Marchione KE, Fletcher JM, Lyon GR, et al. 2002. Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol Psychiatry*. 52(2):101–110.

Tange O. 2018. *GNU Parallel*. doi: 10.5281/zenodo.1146014.

Taylor JSH, Rastle K, Davis MH. 2013. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychol Bull*. 139(4):766–791.

R Core Team. 2016. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Torre GA, Eden GF. 2019. Relationships between gray matter volume and reading ability in typically developing children, adolescents, and young adults. *Dev Cogn Neurosci*. 36:100636.

Turkeltaub PE, Gareau L, Flowers DL, Zeffiro TA, Eden GF. 2003. Development of neural mechanisms for reading. *Nat Neurosci*. 6(6):767–773.

Vandermosten M, Hoeft F, Norton ES. 2016. Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: a review and quantitative meta-analysis. *Curr Opin Behav Sci*. 10:155–161.

Vellutino FR, Scanlon DM. 1987. Phonological coding, phonological awareness, and reading ability: evidence from a longitudinal and experimental study. *Merrill-Palmer Quart*. 33(3):321–363.

Wagner RK, Torgesen JK, Rashotte CA. 1999. *Comprehensive Test of Phonological Processing*. Austin, TX: PRO-ED.

Wagner RK, Torgesen JK. 1987. The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol Bull*. 101(2):192–212.

Warmington M, Hulme C. 2012. Phoneme awareness, visual-verbal paired-associate learning, and rapid automatized naming as predictors of individual differences in reading ability. *Sci Stud Read*. 16(1):45–62.

Wechsler D. 1999. *Wechsler Abbreviated Scale of Intelligence*. New York: The Psychological Corporation and Harcourt Brace & Company.

Wechsler D. 2002. Wechsler Preschool and Primary Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

Westman E, Aguilar C, Muehlboeck JS, Simmons A. 2013. Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topogr*. 26(1):9–23.

Wierenga LM, Langen M, Oranje B, Durston S. 2014. Unique developmental trajectories of cortical thickness and surface area. *Neuroimage*. 87:120–126.

Williams VJ, Juranek J, Cirino P, Fletcher JM. 2018. Cortical thickness and local gyrification in children with developmental dyslexia. *Cereb Cortex*. 28:963–973.

Winkler AM, Greve DN, Bjuland KJ, Nichols TE, Sabuncu MR, Håberg AK, Skranes J, Rimol LM. 2018. Joint analysis of cortical area and thickness as a replacement for the analysis of the volume of the cerebral cortex. *Cereb Cortex*. 28(2):738–749.

Winkler AM, Kochunov P, Blangero J, Almasy L, Zilles K, Fox PT, Duggirala R, Glahn DC. 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage*. 53(3): 1135–1146.

Woodcock RW, McGrew KS, Mather N. 2001. *Woodcock Johnson III*. Itasca, IL: Riverside Publishing.

Xia Z, Hoeft F, Zhang L, Shu H. 2016. Neuroanatomical anomalies of dyslexia: disambiguating the effects of disorder, performance, and maturation. *Neuropsychologia*. 81:68–78.

Xia Z, Zhang L, Hoeft F, Gu B, Gong G, Shu H. 2018. Neural correlates of oral word reading, silent reading comprehension, and cognitive subcomponents. *Int J Behav Dev*. 42(3): 342–356.

Yamada Y, Stevens C, Dow M, Harn BA, Chard DJ, Neville HJ. 2011. Emergence of the neural network for reading in five-year-old beginning readers of different levels of pre-literacy abilities: an fMRI study ☆. *Neuroimage*. 57:704–713.

Yang Z, Qiu J, Wang P, Liu R, Zuo XN. 2016. Brain structure–function associations identified in large-scale neuroimaging data. *Brain Struct Funct*. 221(9):4459–4474.

Zhang M, Li J, Chen CC, Mei L, Xue G, Lu Z, Chen C, He Q, Wei M, Dong Q. 2013. The contribution of the left mid-fusiform cortical thickness to Chinese and English reading in a large Chinese sample. *Neuroimage*. 65(2): 250–256.