


Individual differences in decoding skill, print exposure, and cortical structure in young adults

1854

Clinton L. Johns ^a, Andrew A. Jahn^a, Hannah R. Jones^b, Dave Kush^{a,c}, Peter J. Molfese^{a,d}, Julie A. Van Dyke^{a,e}, James S. Magnuson^{a,e,f,g}, Whitney Tabor^{a,e,g}, W. Einar Mencl^a, Donald P. Shankweiler^{a,g} and David Braze^{a,e}

^aHaskins Laboratories, New Haven, USA; ^bDepartment of Clinical and Social Sciences in Psychology, University of Rochester, Rochester, USA; ^cDepartment of Language and Literature, Norwegian University of Science and Technology, Trondheim, Norway; ^dDepartment of Health and Human Services, Section on Functional Imaging Methods, Laboratory of Brain and Cognition, National Institutes of Mental Health, National Institutes of Health, Bethesda, USA; ^eConnecticut Institute for the Brain and Cognitive Sciences, University of Connecticut, Storrs, USA; ^fBrain Imaging Research Center, University of Connecticut, Storrs, USA; ^gDepartment of Psychological Sciences, University of Connecticut, Storrs, USA

ABSTRACT

This exploratory study investigated relations between individual differences in cortical grey matter structure and young adult readers' cognitive profiles. Whole-brain analyses revealed neuroanatomical correlations with word and nonword reading ability (decoding), and experience with printed matter. Decoding was positively correlated with grey matter volume (GMV) in left superior temporal sulcus, and thickness (GMT) in right superior temporal gyrus. Print exposure was negatively correlated with GMT in left inferior frontal gyrus (pars opercularis) and left fusiform gyrus (including the visual word form area). Both measures also correlated with supramarginal gyrus (SMG), but in spatially distinct subregions: decoding was positively associated with GMV in left anterior SMG, and print exposure was negatively associated with GMT in left posterior SMG. Our comprehensive approach to assessment both confirms and refines our understanding of the novel relation between the structure of pSMG and proficient reading, and unifies previous research relating cortical structure and reading skill.

ARTICLE HISTORY

Received 18 October 2017
Accepted 4 May 2018

KEYWORDS

Individual differences; reading and literacy; structural MRI; phonological decoding; print exposure

1. Introduction

Proficient reading depends upon the efficient coordination of both language-specific processes (e.g. phonological, lexical, syntactic, semantic) and domain general processes (e.g. working memory, reasoning). Becoming literate entails changes to the brain's cortical structure, both in grey and white matter (Carreiras et al., 2009; Castro-Caldas et al., 1999; Petersson, Silva, Castro-Caldas, Ingvar, & Reis, 2007). There is substantial evidence for individual behavioural differences in reading comprehension and its components (for reviews see Long, Johns, & Morris, 2006; van den Broek, Mouw, & Kraal, 2015; Wagner, Piasta, & Torgesen, 2006), and that these differences often correlate with differences in functional activity in task-relevant brain regions (e.g. Clements-Stephens et al., 2012; Meyler, Keller, Cherkassky, Gabrieli, & Just, 2008; Shankweiler et al., 2008; Welcome & Joanisse, 2012). However, although it might be assumed that cortical structure may be similarly related to such behavioural differences, there is relatively little evidence available to support this

hypothesis (for review see Richardson & Price, 2009). Structural imaging may provide critical complementary information about the neural substrates underlying reading behaviour. Consequently, the goal of the current study is to explore potential relations between cortical grey matter structure and performance measures related to reading comprehension.

There are few studies that have directly assessed whether literacy-related skills correlate with indices of grey matter structure. In a recent example (Jednoróg et al., 2015), two such skills were assessed: rapid automatized naming (RAN) and decoding. The goal was to assess potential differences in grey matter volume (GMV) between two large groups of children, one with developmental dyslexia ($n = 130$), and the other an age-matched control group of unimpaired readers ($n = 106$). Performance on the RAN task – a speeded response task in which readers name characters that appear in a visual array – is often associated with fluent and efficient processing speed (although this is the subject of ongoing debate; for review see Norton & Wolf, 2012). *Decoding*

CONTACT Clinton L. Johns  Clinton.Johns@HaskinsLabs.org

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

refers to the ability to accurately map orthographic information – i.e. printed words in alphabetic languages – to a corresponding phonological representation. Both RAN and decoding ability are strong predictors of reading comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990; Norton & Wolf, 2012; Shankweiler et al., 1999). Jednoróg and colleagues found no evidence for group differences in cortical grey matter. There was also no evidence that either decoding or RAN were related to grey matter differences in dyslexic readers. However, they did observe individual differences within the control group: in unimpaired readers, decoding and GMV were correlated such that higher word reading accuracy corresponded to greater GMV in left supramarginal gyrus (SMG). Given its power, this study provides robust evidence that a skill that is critical to reading comprehension can also be directly related to variation in the structure of cortical grey matter in non-dyslexic readers.

The approach adopted by Jednoróg et al. (2015) is in some respects emblematic of most investigations of grey matter morphology and reading ability. Such studies are often primarily concerned with describing some specific group of interest (e.g. dyslexic or bilingual readers). Consistent with this, most studies of linguistic relations to cortical grey matter emphasise group-level comparisons with “typical” (i.e. non-dyslexic, or monolingual) readers, rather than specific measures of the participants’ reading-related skills. Furthermore, even when such measures are included, their scope is usually quite limited: it is most common to obtain only a single skill measure (or perhaps two). One consequence of these methodological emphases is that our ability to assess potential relations between distinct components of reading skill and neuroanatomical structure is limited by the reliance on only narrow information about readers’ literacy skills. Another is that we know comparatively little about the possible neurostructural correlates of literacy in so-called “typical” readers. Thus, we conducted an exploratory study of cortical grey matter structure in monolingual young adults without diagnosed reading disabilities. Further, in order to more fully characterise both the processes supporting reading comprehension in our participants and their neuroanatomical correlates, we administered a battery of behavioural tests indexing a wide range of reading-related skills.

There is some evidence that the structural relation to decoding skill that was observed in children (Jednoróg et al., 2015) may also be present in young adults. In one study, brain morphology was related to group differences in college students (Welcome, Chiarello, Thompson, & Sowell, 2011). Group membership was determined by testing not only decoding efficiency,

but also participants’ reading comprehension ability (i.e. their global understanding of a text). There were three groups: proficient readers ($n = 22$), whose scores indicated that their performance on both tests was commensurate with age; poor readers ($n = 12$), whose performance was below age-based norms; and “resilient” readers ($n = 21$), who had impaired decoding, but nonetheless exhibited age-appropriate reading comprehension. Decoding ability was related to hemispheric asymmetry in temporo-parietal regions (including SMG): grey matter thickness (GMT) in left hemisphere regions was greater relative to their homologues in the right hemisphere in proficient readers. However, this asymmetry was reduced in both groups with poor decoding skills. Although reductions in the typical leftward asymmetry are known to be associated with specific reading disability (Eckert, 2004; Heim & Keil, 2004; see also Chiarello, Lombardino, Kacinik, Otto, & Leonard, 2006), this was the first report of such a relation in individuals without a history of reading disability. In contrast to decoding, reading comprehension ability was not related to the structure of the left temporo-parietal region. Instead, it correlated only with structural aspects of right hemisphere brain regions: specifically, radial expansion – a measure of “local brain shape” related to cortical surface area – was smaller in frontal and parietal regions for poor comprehenders relative to the other groups. Overall, these findings provide important neurostructural information about the components of literacy skill: they corroborate the importance of assessing decoding ability; indicate a discrete role for measures of reading comprehension ability; and suggest that the relation between efficient decoding and the structure of cortical regions including SMG may be consistent across developmentally distinct age groups.

Another study of college students without histories of reading difficulty ($N = 28$) also assessed both decoding ability and reading comprehension skill, and additionally included measures of experience with printed material (Goldman & Manis, 2013). Significant correlations among these three literacy-related skills are well established. Skilled readers typically exhibit greater decoding ability and more extensive print exposure; conversely, decoding difficulty is associated with less print exposure and poorer reading comprehension (Cunningham & Stanovich, 1991, 1997; Shankweiler et al., 1999; Share, 1995; Stanovich, 1986; Stanovich & Cunningham, 1992; for review see Mol & Bus, 2011). Goldman & Manis defined regions of interest in the left hemisphere reading network (e.g. Bolger, Perfetti, & Schneider, 2005; Pugh, 2006; Pugh et al., 2013). Decoding ability was not related to GMT in any of these regions in this sample. However, both print exposure and reading

comprehension were positively correlated with GMT. In separate analyses, print exposure was related to GMT in left SMG, left fusiform gyrus, both pars opercularis and pars triangularis in inferior frontal gyrus (IFG), and angular gyrus (AG); reading comprehension ability, by contrast, was related to GMT only in the latter two regions. A subsequent analysis explored whether the overlapping correlations in IFG and AG indicated that each literacy skill was a unique predictor of GMT, or whether the correlations in the initial analyses might instead be based upon variance shared between the two measures. Although the latter was the case in IFG, print exposure emerged as a unique predictor of GMT in AG. These results corroborate the importance of assessing reading comprehension ability; suggest experience with printed material as a potentially important structural correlate; and highlight the importance of accounting for shared variance among multiple measures of literacy skills in analyses of individual differences and brain morphology.

Two additional studies report significant correlations between cortical grey matter and literacy skills in non-clinical populations. The principal findings concern vocabulary knowledge, which is known to be strongly correlated with reading comprehension (Anderson, Wilson, & Fielding, 1988; Joshi, 2005; Perfetti, 2007; Stanovich, 1986). Measures of vocabulary knowledge often emerge as unique predictors that capture reading-related variance beyond that of other cognitive assessments (for reviews see Braze et al., 2016; Protopapas, Mouzaki, Sideridis, Kotsolakou, & Simos, 2013; Tunmer & Chapman, 2012). Lee et al. (2007) found that greater vocabulary knowledge was positively related to grey matter density in bilateral SMG in adolescents ($N = 34$). Richardson, Thomas, Filippi, Harth, and Price (2010) confirmed this finding in adolescents, but did not observe the same correlation in either young children or adults. Moreover, in the later study, vocabulary knowledge was also positively associated with grey matter density in two other brain regions: left superior temporal sulcus (STS) and, in adults and adolescents only, in left posterior temporo-parietal cortex. Finally, both studies included some additional indices of individual differences, none of which correlated with cortical structure: Lee et al. included measures of verbal fluency and both verbal and performance IQ, and Richardson et al. included a measure of matrix reasoning ability. These results demonstrate the importance of assessing readers' word knowledge. In addition, both studies used their measures of general reasoning as nuisance variables in their regression analyses, so that any observed effects could be attributed specifically to linguistic factors, rather than to general cognitive ability.

Taken together, these studies provide evidence that at least a small set of cognitive skills related to reading comprehension may be correlated with cortical grey matter structure. However, it is equally clear that direct comparison of their results is not straightforward. For example, the studies of collegiate young adults included no measure of vocabulary knowledge, obviating a possible extension of the vocabulary findings. Both studies of college students (Goldman & Manis, 2013; Welcome et al., 2011) assessed decoding ability, but comparing their results is complicated: although one tested whether decoding related to cortical structure, the other employed it to differentiate groups of participants (and analysed broadly defined cortical regions rather than specific areas of the brain). Both assessed reading comprehension, but Goldman and Manis did not examine the right hemisphere, precluding confirmation of differences in hemispheric asymmetry found by Welcome and colleagues. The populations tested differ in each study, with two using convenience samples of college students (age 18–24, Goldman & Manis, 2013; age 18–34, Welcome et al., 2011), one using pre-collegiate adolescents (age 12–16; Lee et al., 2007), one using children from three different nations (age 8–13; Jednoróg et al., 2015), and one using a wide spectrum of ages (age 7–11, $n = 9$; age 12–17, $n = 17$; age 21–72, $n = 22$; Richardson et al., 2010). These sampling differences are non-trivial for comparing the results of these studies, as there is ample evidence for neurodevelopmental structural changes across these age groups (Giedd et al., 1999; Gogtay et al., 2004; Lu et al., 2009; Paus, 2005; Salat et al., 2004). The consequence of these methodological differences – assessing different indices of grey matter structure, and different measures of reading-related abilities, in different populations – is that each study presents a relatively narrow account of potential neurostructural links to the components of literacy skill. That is, although the available evidence does not contain clear contradictions, it also does not admit clear conclusions, either about the literacy skills that might be related to cortical structure, or the cortical structures to which they might relate.

Despite these inconsistencies, there is one region whose grey matter structure appears to be consistently linked to components of reading comprehension: supramarginal gyrus. Furthermore, SMG's structural relations to literacy skills may provide unique information about the neural substrates of reading comprehension. Specifically, both studies of vocabulary knowledge localised the correlation with grey matter density to posterior SMG (pSMG), rather than anterior SMG (aSMG). This distinction is important since, as noted by both Lee et al. (2007) and Richardson et al. (2010), pSMG has not been functionally

related to language-specific processes. Rather, functional relations to language are typically reported in neighbouring areas such as aSMG and AG. These regions are associated with functional activity during phonological (aSMG) and semantic (AG) processing (Booth et al., 2002; Démonet et al., 1992; Devlin, Matthews, & Rushworth, 2003; Gathercole, Hitch, Service, & Martin, 1999; Price, Moore, Humphreys, & Wise, 1997; Tan, Laird, Li, & Fox, 2005). Additional support for this structural dissociation comes from a complementary analysis of white matter tractography, which revealed that pSMG has direct connections with both aSMG and AG, but that aSMG and AG are not themselves directly connected (Lee et al., 2007). Although Goldman and Manis (2013) did not specify MNI coordinates for their findings (making it unclear whether print exposure was correlated with a specific subregion of SMG), the other studies reporting links to SMG are broadly consistent with this pattern. The area of SMG in which GMT related to decoding in children was anteriorly located (Jednoróg et al., 2015); and although Welcome et al. (2011) also did not discriminate subregions of SMG, their figures suggest that the effects driven by group differences in decoding were in anterior SMG (see Figures 2 and 3 of Welcome et al., 2011, pp. 1199 & 1201). Importantly, this fine-grained subdivision of the structure of SMG according to specific components of literacy is not possible on the basis of functional associations alone.

This study advances our knowledge of possible links between brain morphology and skills related to literacy achievement in young adult readers. It is notable for at least three reasons. First, we assessed literacy skills using a large battery of cognitive assessments. This battery included measures assessed in previous studies, such as decoding ability, reading comprehension, vocabulary knowledge, and print exposure, as well as additional cognitive abilities that were not. This study simultaneously assessed the unique contributions of a broad range of specific cognitive measures, and is a necessary step for reconciling the diverse findings from previous research. Second, we assessed two indices of cortical structure: grey matter thickness and grey matter volume. Most previous research examining relations between cortical structure and reading ability (especially in clinical populations – see the General Discussion) has focused on only one of these, typically the latter, which is intuitive: GMV is derived from GMT and cortical surface area, and it might therefore be expected that differences observed in GMV would be reflected in its component measures. Yet recent evidence shows that this is only true for cortical surface area. In contrast, differences that manifest in GMT may not be reflected in GMV, and vice versa (Frye et al., 2010; Greve et al., 2013),

making it important to assess both in order to clearly characterise grey matter variation (cf. Winkler et al., 2010). Finally, although most previous work has recruited students of one kind or another, we chose to recruit a community-based sample of young adults who were not university students. The neurobiological bases of literacy skill have not been as extensively studied in young adults as in early language learners (Curtis, 2002), and even less is known about young adult readers who are not enrolled in (and may not plan to obtain) post-secondary education. Based on our previous work with this population, we expected a broad range in literacy-related skills across participants (Braze, Tabor, Shankweiler, & Mencl, 2007; Braze, Mencl et al., 2011; Braze, Katz et al., 2016; Johns, Matsuki, & Van Dyke, 2015; Kukona et al., 2016; Kuperman & Van Dyke, 2011; Li et al., 2017; Magnuson et al., 2011; Shankweiler et al., 2008; Van Dyke, Johns, & Kukona, 2014), which confers an advantage in our power to detect individual differences (for discussion see Peterson, 2001).

Given the small number of studies that constitute the current state of the field, we consider this an exploratory investigation. The scarcity of previous research, as well as its methodological heterogeneity and diverse, non-overlapping patterns of results, make it difficult to propose specific hypotheses about potential links between cortical structure and behavioural measures of literacy-related skills. Thus, although we were particularly interested in SMG (see above), we did not define any hypothesis-driven regions of interest *a priori*. Rather, we conducted a naïve whole brain analysis without assumptions or restrictions based on the size, location, or direction of potential neurostructural correlations with the behavioural battery measures.

2. Method

2.1. Participants

We obtained informed consent from 39 young people recruited from the local community. We recruited participants in several ways, including presentations at adult education centres, advertisements in local newspapers, and flyers placed on adult school campuses, community centres, public transportation hubs, and local retail and laundry facilities. Of the 39 participants, four were left-handed, and their data were excluded from further analysis. The remaining 35 participants (ages 16–24 years, mean 20.44; 17 female) were right-handed native English speakers with normal or corrected-to-normal vision. Participants reported no history of psychiatric or neurological disorder, no active use of psychoactive

medications, and no diagnosed reading or learning disability. Based on the Fast Reading subtest of the Stanford Diagnostic Reading Test (Karlson & Gardner, 1995), all participants demonstrated the ability to read well enough to comprehend basic texts (minimum 70% accuracy on attempted items). Each participant underwent two experimental sessions, each on a separate day. Participants received \$80 for one scanning session, which lasted no longer than 60 min, together with one behavioural testing session lasting no longer than three hours. Behavioural testing was completed prior to the MRI scan. The Yale University Human Investigation Committee approved this protocol.

2.2. Literacy-related cognitive assessments

We administered a battery of behavioural tests of literacy-related skills and abilities. Standardised instruments were chosen to optimise construct validity and test-retest reliability. The standardised measures are widely used for clinical assessment and diagnosis, and were administered individually during individual test sessions. Two skills – working memory and print exposure – were not derived from standardised assessments; in these cases, we employed test instruments identical in format to those that are commonly used in experimental research. The skills we examined, and the tests associated with them, included:

- Print exposure: *Magazine Recognition Test* (Acheson, Wells, & MacDonald, 2008; Cunningham & Stanovich, 1990), in which participants identify real magazine titles from a list that includes real and foil titles. We retained the original format of this test, but updated test items by replacing out-of-print titles with the names of current publications.
- Vocabulary knowledge: *Peabody Picture Vocabulary Test-Revised* (PPVT-R; Dunn & Dunn, 1997), a test of receptive vocabulary knowledge, in which participants hear a target word and select a picture (from a group of four possibilities) that best depicts its definition.
- Working memory capacity (WMC): assessed with the *Sentence Span* task (Daneman & Carpenter, 1980). Following the original format of this complex span assessment, participants heard sets of 2–6 sentences (number per set increases linearly), judging each as true or false; after each series, all sentence-final words must be recalled (in any order). We used an auditory variant of the task, permitting us to measure verbal working memory independent of the need to decode printed stimuli; moreover, we modified the sentence materials in order to make them more amenable for use with our community-based sample (for details, see Clark, McRoberts, Van Dyke, Shankweiler, & Braze, 2012).
- Non-linguistic reasoning ability: assessed using the *Weschler Abbreviated Scale of Intelligence* (WASI; Psychological Corp., 1999). We used the Matrix Reasoning subtest, in which participants completed visual analogical reasoning tasks, as a measure of general cognitive ability.
- Phonological awareness: *Comprehensive Test of Phonological Processing* (CTOPP; Wagner, Torgesen, & Rashotte, 1999). We used the composite phonological awareness measure, derived from the Elision core subtest (forming words by eliding a phonological segment from spoken word prompts) and the Blending Words core subtest (spoken sounds are combined to form words).
- Rapid naming: CTOPP (Wagner et al., 1999), Rapid Letter Naming core subtest. Scores on this test reflect time to name letters presented in a grid-like array. Because RAN scores are naming times, lower scores indicate better performance.
- Reading comprehension: *Peabody Individual Achievement Test-Revised* (PIAT-R; Markwardt, 1998). Participants read a series of sentences of increasing difficulty, choosing a picture (from a group of four possibilities) corresponding to the meaning of each. We administered odd numbered items to measure reading comprehension, reserving the even numbered items for a measure of speech sentence comprehension (Leach, Scarborough, & Rescorla, 2003; Spring & French, 1990).
- Listening comprehension: we created a listening comprehension assessment by splitting the PIAT-R (Markwardt, 1998), such that even numbered items were recorded and presented aurally in order to assess listening comprehension (Leach et al., 2003; Spring & French, 1990). The characteristics of the sentences, and the behavioural response task (i.e. selecting a picture from an array) are therefore identical to our reading comprehension measure.
- Reading Fluency: indexed using the WJ-III silent reading fluency subtest, from the reading and oral comprehension area subtests (WJ-III; Woodcock, McGrew, & Mather, 2001). This test measures the speed of reading sentences silently and answering yes/no questions about each. We also measured oral reading fluency through a subset of the *Gray Oral Reading Test*, fourth edition (GORT, passages 5, 7, and 9; Wiederholt & Bryant, 2001). Reading time for each passage was converted to a rate using the published tables; these were summed to yield a single score.

- Decoding skill: assessed using both the *Woodcock-Johnson-III Tests of Achievement*, reading and oral comprehension area subtests (WJ-III; Woodcock et al., 2001), Word Attack (reading a list of pseudowords aloud) and Letter-Word Identification (naming words from a list); and the *Test of Word Reading Efficiency* (TOWRE; Torgesen, Wagner, & Rashotte, 1999), sight word efficiency subtest (indexes the number of words that can be named in 45 s) and phonemic decoding efficiency subtests (indexes the number of pronounceable nonwords that can be named in 45 s).

2.3. Data preparation

We inspected the distribution of the raw scores for each measure using density and quantile-quantile plots in order to assess univariate normality, and using the Shapiro-Wilk test, which assesses both skewness and kurtosis (Shapiro & Wilk, 1965; implemented in the *stats* R package, R Core Team, 2016). Some measures showed significantly skewed distributions, which can inflate the influence of non-normal data. In order to correct this, we applied the Box-Cox transformation where appropriate (Box & Cox, 1964). The transformation equation is $y(\lambda) = (y^\lambda - 1)/\lambda$, given $\lambda \neq 0$; if $\lambda = 0$, then $\log(y)$. Using the *caret* R package (Kuhn, 2016), we calculated optimal lambda values for each measure. Lambda values close to 1.00 indicate that no transformation is necessary, since the distribution of the transformed data will be identical to the original data.

After addressing the distributional characteristics of the measures, all predictor variables were standardised: each was first mean-centred (i.e. the mean of each measure is subtracted from its value, setting its mean to zero, but leaving the standard deviation unchanged) and then scaled (i.e. the resulting values were converted to z-scores). Standardisation has several well-established analytic benefits (as described in, e.g. McElreath, 2016), such as allowing straightforward comparison of the relative influence of predictor variables, (which might not be possible with unstandardised data due to, for example, differences in measurement and/or scale), and reducing potential problems related to multicollinearity among predictor variables (see below).

Some measures in our test battery target the same theoretical constructs. For these, we built composite variables to improve reliability and to more robustly represent the underlying constructs. We created two composite variables: *Reading Fluency* (comprised of the WJ-III Reading Fluency subtest and the GORT, $r = .372$, $p = .028$) and *Word Decoding* (comprised of the WJ-III

Word Attack and Letter-Word Identification subtests, and the TOWRE subtests, all $r_s > .45$, all $p_s < .01$). Composites were derived by averaging component measures after they were first standardised, and then rescaling the resulting composite values. This approach is common in clinical and psycholinguistic studies of reading and reading-related skills (e.g. Braze et al., 2007; Guo, Roehrig, & Williams, 2011; Hua & Keenan, 2014; Kukona et al., 2016; Pugh et al., 2008; Sabatini, Sawaki, Shore, & Scarborough, 2010; Shankweiler et al., 2008; Van Dyke et al., 2014).

Finally, it is well known that performance on individual difference measures tends to be correlated, sometimes highly, making it difficult to uniquely relate specific constructs to dependent variables (for discussion see Freed, Hamilton, & Long, 2017). However, neither the number nor the magnitude of bivariate correlations is an unambiguous indicator of troublesome multicollinearity: strongly correlated measures may not induce problematic multicollinearity, while high multicollinearity can occur even when all bivariate correlations in a set of variables are quite low (Belsley, 1991a; Flom, 1999). Thus, although standardising the predictor variables is known to reduce such dependencies (for discussion see McElreath, 2016), we nonetheless assessed the potential for problematic multicollinearity among our individual differences assessments. Using the *perturb* R package (Hendrickx, 2012), we calculated each predictor's condition number (κ), which provides information about how much the variance associated with an estimated regression coefficient is increased because of overlap with other predictors (Belsley, Kuh, & Welsch, 1980). Condition numbers are considered to be both more informative and precise than other estimates of multicollinearity, such as the more commonly used variance inflation factor (Belsley, 1991a). This is because condition numbers, unlike variance inflation factors, provide not only estimates of shared variance but information about the *ensembles* of variables which may be sharing variance. By contrast, variance inflation factors are relatively uninformative, because they cannot account for connections among variables (Harrell, 2001). For condition numbers that are "absolutely small, for example, 5 or 10 ... collinearity is not really a major problem" (Belsley, 1991b, p. 42), whereas $\kappa \geq 30$ suggests problematic multicollinearity (Belsley et al., 1980; Belsley, 1991a, 1991b; Faraway, 2014). Similarly, Baayen notes that for κ values "between 0 and 6, there is no collinearity to speak of. Medium collinearity is indicated by condition numbers around 15, and conditions numbers of 30 or more indicate potentially harmful collinearity" (Baayen, 2008, p. 182).

2.4. Structural imaging

2.4.1. Image acquisition and processing

We collected volumetric data from high-resolution 3D MPRAGE anatomical images, acquired on a Siemens 1.5 T Sonata MR system (192 sagittal slices; TE = 4.66 ms, TR = 2530 ms; FOV = 256 × 256 voxel matrix; resolution = 1.33 × 1.33 × 1.30 mm³). One whole-head, high resolution T1-weighted MPRAGE anatomical volume was acquired per participant. We used the FreeSurfer image analysis suite to perform cortical reconstruction and volumetric segmentation (Dale, Fischl, & Sereno, 1999; Dale & Sereno, 1993; Fischl & Dale, 2000; Fischl et al., 2001; <http://surfer.nmr.mgh.harvard.edu>, version 5.3). Prior to segmentation and classification, all images were visually inspected to identify potentially problematic motion artifacts. Subsequently, each segmentation was visually inspected to ensure (1) accurate skull stripping, (2) correct classification of grey/white matter boundaries, (3) appropriate separation of brain/non-brain matter. Minor adjustments to the automated segmentation and parcellation routines were made when necessary (e.g. adding control points to facilitate grey/white matter classification), but no major alterations were necessary. Non-brain tissue was removed using a hybrid watershed/surface deformation procedure (Ségonne et al., 2004). The resulting skull-stripped brain was processed using an automated Talairach transformation, segmentation of the subcortical white matter and deep grey matter volumetric structures (Fischl et al., 2002, 2004), intensity normalisation (Sled, Zijdenbos, & Evans, 1998), tessellation of the grey/white matter boundary, automated topology correction (Fischl et al., 2001; Ségonne, Pacheco, & Fischl, 2007), and surface deformation following intensity gradients to optimally place the grey/cerebrospinal and grey/white fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale et al., 1999; Dale & Sereno, 1993; Fischl & Dale, 2000). Both intensity and continuity information from the entire three-dimensional MR volume are used to produce representations of cortical thickness, calculated as the closest distance from the grey/white boundary to the grey/CSF boundary at each vertex on the tessellated surface (Fischl & Dale, 2000). The maps are not restricted to the voxel resolution of the original data, and are created using spatial intensity gradients across tissue classes (i.e. they are not simply reliant on absolute signal intensity). These procedures for the measurement of cortical thickness have been validated against histological analysis (Rosas et al., 2002) and manual measurements (Kuperberg et al., 2003; Salat et al., 2004).

The left and right hemispheres of all 35 participants were registered to the *fsaverage* atlas (common surface space) templates included in FreeSurfer, and smoothed with a Gaussian kernel of FWHM 10 mm. Each hemisphere was modelled separately. In contrast to non-surface based volumetric smoothing, surface-based smoothing only averages data from nearby vertices on the cortical mantle. This prevents the mixing of signal from nearby ridges and different tissue types and increases the spatial specificity of the averaged signal.

2.4.2. Spatial analysis

Differences in the measurements of grey matter volume and thickness were examined for both the left and right cerebral hemispheres with a vertex by vertex general linear model (GLM). Statistical analysis was performed at each vertex to test the significance of the correlation between the individual difference measurements and these structural measurements. The model included 11 regressors: 10 corresponding to the behavioural individual difference assessments, and an additional covariate of no interest – participant age (mean centred) – which was modelled to reduce error variance. Total intracranial volume (TIV) is often included as a covariate for between-group designs and regions-of-interest analyses, with the goal of normalising the data so that group-level effects are not confounded with individual differences in GMV. However, since we are specifically concerned with modelling correlations with individual differences in GMV, including TIV would explicitly remove variance relevant to this research question, making it an inappropriate covariate for our within-subjects design. In addition, there is broad agreement that GMT should not be normalised in any case (e.g. Fjell et al., 2009; Westman, Aguilar, Muehlboeck, & Simmons, 2013), and previous individual difference studies of GMT have thus not done so (e.g. Goldman & Manis, 2013; He et al., 2013).

Separate GLMs were constructed for GMV and GMT to estimate parameters for the covariates. These parameter estimates were then submitted to a group-level analysis and converted to *p*-value maps. Given the exploratory nature of this study, all vertex-wise results were thresholded at an individual vertex level of *p* < 0.05, and cluster extent thresholds corrected for multiple comparisons (*p* < 0.05) were calculated through Monte Carlo simulations of white noise on the cortical surface (Hagler, Saygin, & Sereno, 2006). These analyses permit the evaluation of the unique contributions of the regressors, in that they assess the proportion of variance in the dependent

variables (GMT and GMV) that is associated with one predictor but not any of the other predictors (i.e. their squared semi-partial correlations; see, e.g. Keith, 2014; Tabachnick & Fidell, 2012). Finally, it is worth noting that the appropriate parameters for the analysis of neuroimaging data is a topic of active, ongoing discussion. For example, although Eklund, Nichols, and Knutsson (2016) advocate more stringent cluster-forming thresholds for functional data, their results were not subsequently replicated (Cox, Chen, Glen, Reynolds, & Taylor, 2017a, 2017b), and may not account for elevated Type II error (Lohmann et al., 2017). At the time of this writing, the implications of this debate for the analysis of structural MRI data have not been tested. Therefore, our cluster-forming threshold was selected to be consistent with both the exploratory nature of our study and the current “state-of-the-science” threshold used in other recent studies examining structural imaging data (e.g. Bizzo et al., 2017; Gardumi, Ivanov, Havlicek, Formisano, & Uludağ, 2017; Jasińska et al., 2017).

Table 1. Descriptive statistics for all battery measures.

Measure	Range	M	SD	Max. possible
1. Magazine recognition test	0–23	10.31	5.66	40
2. Vocabulary knowledge	132–192	169.80	16.84	204
<i>Age equivalent</i>	10.4–23.1	18.93	4.51	>23.1
3. Working memory capacity	28–60	46.37	7.26	60
4. Matrix Reasoning	9–31	23.89	4.61	48
<i>Test Age equivalent</i>	6.5–25–29	13.93	4.67	25–29
5. Phonological awareness	61–118	93.23	15.66	150
<i>Percentile equivalent</i>	1–89	38.86	29.42	99
6. Rapid letter naming (<i>in seconds</i>)	16–34	23.91	4.18	<i>n/a</i>
<i>Percentile equivalent</i>	5–99	52.34	28.45	99
7. Listening comprehension	14–40	33.97	4.79	41
<i>Grade equivalent</i>	3–13	9.80	2.31	13
8. Reading comprehension	20–40	32.74	5.60	41
<i>Grade equivalent</i>	3.8–13	9.26	2.9	13
9. Oral reading fluency	12–30	23.46	4.69	30
10. Silent reading fluency	51–98	74.69	15.04	98
<i>Grade equivalent</i>	5.8–19	12.65	6.60	>18
<i>Rate (items/min., max. 180 s)</i>	17–43	25.68	6.68	<i>n/a</i>
11. Word identification	59–76	68.49	4.37	76
<i>Grade equivalent</i>	6.7–19	13.65	4.16	>18
12. Letter-word identification	20–31	27.23	2.91	32
<i>Grade equivalent</i>	4.3–19	10.67	4.26	>18
13. Sight word efficiency	75–104	91.71	8.92	104
<i>Grade equivalent</i>	5.8–12.6	10.68	2.12	>12.6
<i>Rate (items/min., max. 45 s)</i>	100–156	123.90	14.66	<i>n/a</i>
14. Phonemic decoding efficiency	22–62	52.23	8.73	63
<i>Grade equivalent</i>	2.8–12.6	10.63	2.76	>12.6
<i>Rate (items/min., max. 45 s)</i>	29–116	72.34	15.50	<i>n/a</i>

Note: 1: Print Exposure (adapted from Cunningham & Stanovich, 1990); 2: Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1997); 3: Listening span (Daneman & Carpenter, 1980); 4: Weschler Abbreviated Scales of Intelligence (Psychological Corp., 1999); 5–6: Comprehensive Test of Phonological Awareness (Wagner et al., 1999); 7–8: Peabody Individual Achievement Test-Revised (Markwardt, 1998); 9: Gray Oral Reading Test (Wiederholt & Bryant, 2001); 10: Woodcock-Johnson-III Tests of Achievement (Woodcock et al., 2001) silent reading fluency; 11–12: Woodcock-Johnson-III Tests of Achievement (Woodcock et al., 2001); 13–14: Test of Word Reading Efficiency (Torgesen et al., 1999).

3. Results

3.1. Literacy-related cognitive assessments

All battery measures were analysed with the R statistical software, version 3.3.1 (R Core Team, 2016). Descriptive information, including range, mean, standard deviation, are shown in Table 1. To aid interpretability and ease comparison with other studies, we include grade or age equivalents where possible. CTOPP scores do not have age or grade equivalents, but do include age-leveled percentile ranks. All analyses are based on raw scores except for the two TOWRE subtests and the WJ-III reading fluency subtest: because some participants completed these timed tests in less than the maximum time allotted, we converted scores on these measures into rates that index items-per-minute.

We next examined the distribution of the raw scores for each measure for normality and potential outliers. Density and quantile-quantile plots suggested that some battery data were non-normal. The results of the Shapiro-Wilks normality test indicate that the data from seven battery measures deviated from normality. After applying the Box-Cox transformation to these measures (as described above), all distributions but one no longer deviated from normality. The remaining measure, WJ-III (silent) reading fluency, had a lambda of 1.04, indicating no advantage to transformation. The details of the tests of normality and of the Box-Cox data transformations appear in Table 2.

Correlations among the 10 battery regressors are shown in Table 3. We observed a range of correlation strength, which we characterise according to the finer gradation proposed in Evans (1996). The correlations

Table 2. Results of the Shapiro-Wilk (*W*) test of normality before and after data transformation (and associated lambda (λ) values, where appropriate) for all battery and composite measures.

Measure	<i>W</i>	λ	Box-Cox <i>W</i>
1. Magazine recognition test	0.972		
2. Vocabulary knowledge	0.930*	4.36	0.956
3. Working memory capacity	0.978		
4. Matrix Reasoning	0.869***	2.93	0.974
5. Phonological awareness	0.965		
6. Rapid letter naming	0.978		
7. Listening comprehension	0.767****	4.81	0.970
8. Reading comprehension	0.890**	3.49	0.946
9. Oral reading fluency	0.902**	2.94	0.957
10. Silent reading fluency	0.900**	1.04 [†]	
11. Word identification	0.976		
12. Letter-word identification	0.916*	4.80	0.958
13. Sight word efficiency	0.965		
14. Phonemic decoding efficiency	0.967		
<i>Composite measures</i>			
9–10. Reading Fluency Composite	0.980		
11–14. Decoding Composite	0.975		

* $p < .05$; ** $p < .01$; *** $p < .001$; **** $p < .0001$; [†]Transform unnecessary.

Table 3. Condition numbers (κ) for and correlations (r) among the individual differences regressors.

Measure	κ	1	2	3	4	5	6	7	8	9	10
1. Print Exposure	2.29	–	.001	.314	.728	.047	.574	.066	.099	.009	.015
2. Vocabulary	4.64	.552	–	.002	.057	.000	.884	.000	.000	.000	.000
3. Working Memory	3.98	.175	.506	–	.089	.006	.933	.002	.022	.046	.006
4. Reasoning	3.16	.061	.324	.291	–	.002	.049	.004	.000	.257	.117
5. Phonological Awareness	2.85	.339	.640	.456	.508	–	.178	.000	.000	.012	.000
6. Rapid Naming	3.68	-.098	-.026	-.015	.335	.233	–	.481	.123	.173	.028
7. Reading Comprehension	6.87	.315	.813	.506	.470	.630	.123	–	.000	.006	.003
8. Listening Comprehension	5.01	.283	.747	.385	.569	.605	.266	.778	–	.000	.000
9. Reading Fluency Composite	8.26	.435	.581	.339	.197	.420	-.236	.560	.454	–	.000
10. Decoding Composite	10.18	.408	.684	.452	.270	.624	-.372	.669	.483	.654	–

$|r| \geq .335, p < .05$; $|r| \geq .435, p < .01$; $|r| \geq .552, p < .001$; $|r| \geq .624, p < .0001$.

Note: Correlations appear below the diagonal; their associated p -values appear above the diagonal.

ranged from very weak ($|r| \leq .19$), weak ($.20 \geq |r| \leq .39$), and moderate ($.40 \geq |r| \leq .59$) to strong ($.60 \geq |r| \leq .79$) and very strong ($|r| \geq .80$). We observed moderate to strong correlations between those measures previously included in studies of cortical grey matter in non-dyslexic populations. For example, vocabulary knowledge was strongly correlated to decoding ($r = .684, p < .0001$), and moderately to print exposure ($r = .552, p < .001$); the latter measures were themselves also moderately correlated ($r = .408, p = .015$). Overall, the observed correlations are consistent with many other studies that have measured a broad range of literacy skills (e.g. Braze et al., 2007, 2016; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Freed et al., 2017; Kukona et al., 2016; Li et al., 2017; Long, Prat, Johns, Morris, & Jonathan, 2008; Macaruso & Shankweiler, 2010; Van Dyke et al., 2014).

Finally, we assessed multicollinearity among the predictors by calculating their condition numbers (κ). Our analysis showed that multicollinearity among our predictors is not problematic for our subsequent analyses: all κ s were substantially below 30, with all but three measures below 6. The condition number for each predictor appears in Table 3.¹

3.2. Structural imaging

Our whole brain analysis revealed numerous correlations between literacy skills and cortical structure in both hemispheres.² Significant correlations between our behavioural measures and grey matter structure appear in Table 4 (GMT; see Figure 1) and Table 5 (GMV; see Figure 2). Most these correlations were negative, indicating that more effective performance on a given measure was correlated with thinner, rather than thicker, GMT and/or reduced, rather than increased, GMV. The exception to this was the decoding composite, for which increased decoding efficiency corresponded to grey matter increases. The only predictors that were uncorrelated with any aspect of cortical structure in our analysis were participant age, vocabulary knowledge and rapid naming.

4. Discussion

Our discussion focuses on the subset of the overall findings that converge with previous studies. This encompasses SMG, left IFG, left STS, and areas in right frontal and parietal areas. Given SMG's prominence in previous studies, we highlight significant correlations between

Table 4. Grey matter thickness: individual difference measures, peak t -values, and centroid coordinates.

Measure	Cluster	Area (mm ²)	t	p	MNI Coordinates		
					x	y	z
Print Exposure	LH ITG	1617	-4.01	.0001	-40.8	-59.4	-6.2
	LH IFG (pars opercularis)	1202	-5.75	.0037	-54.3	19.4	16.4
	LH SMG	1150	-2.53	.0051	-38.8	-44.8	35.5
	LH caudal middle frontal	901	-2.20	.0251	-38.3	0.2	46.5
	RH rostral middle frontal	3280	-3.83	.0001	42.3	28.4	21.2
	RH MTG	1229	-3.66	.0040	47.8	-61.1	3.9
Working Memory	RH precentral	1581	-3.15	.0004	34.8	-10.0	56.0
Phonological Awareness	RH MTG	1129	-3.07	.0069	60.8	-40.5	-5.3
Listening Comprehension	RH lingual	1420	3.58	.0008	32.8	-50.5	-6.1
Reading Comprehension	LH rostral middle frontal	2325	-2.92	.0001	-19.3	40.8	33.7
	LH ITG	1401	-3.39	.0007	-40.8	-59.4	-6.2
	LH superior frontal	860	-3.27	.0333	-9.3	57.5	10.3
	RH transverse temporal	3093	-5.97	.0001	39.2	-29.9	11.2
	RH rostral middle frontal	1380	-4.52	.0010	41.9	26.1	21.6
	RH inferior parietal	849	-3.54	.0470	52.0	-48.1	24.8
Decoding Composite	RH caudal middle frontal	841	-2.79	.0492	39.3	9.7	47.0
	RH STG	1720	4.23	.0001	43.8	-32.3	9.5
	RH precentral	1357	3.76	.0012	23.1	-28.4	51.2
	RH lateral occipital	1191	4.60	.0047	29.4	-88.0	11.7

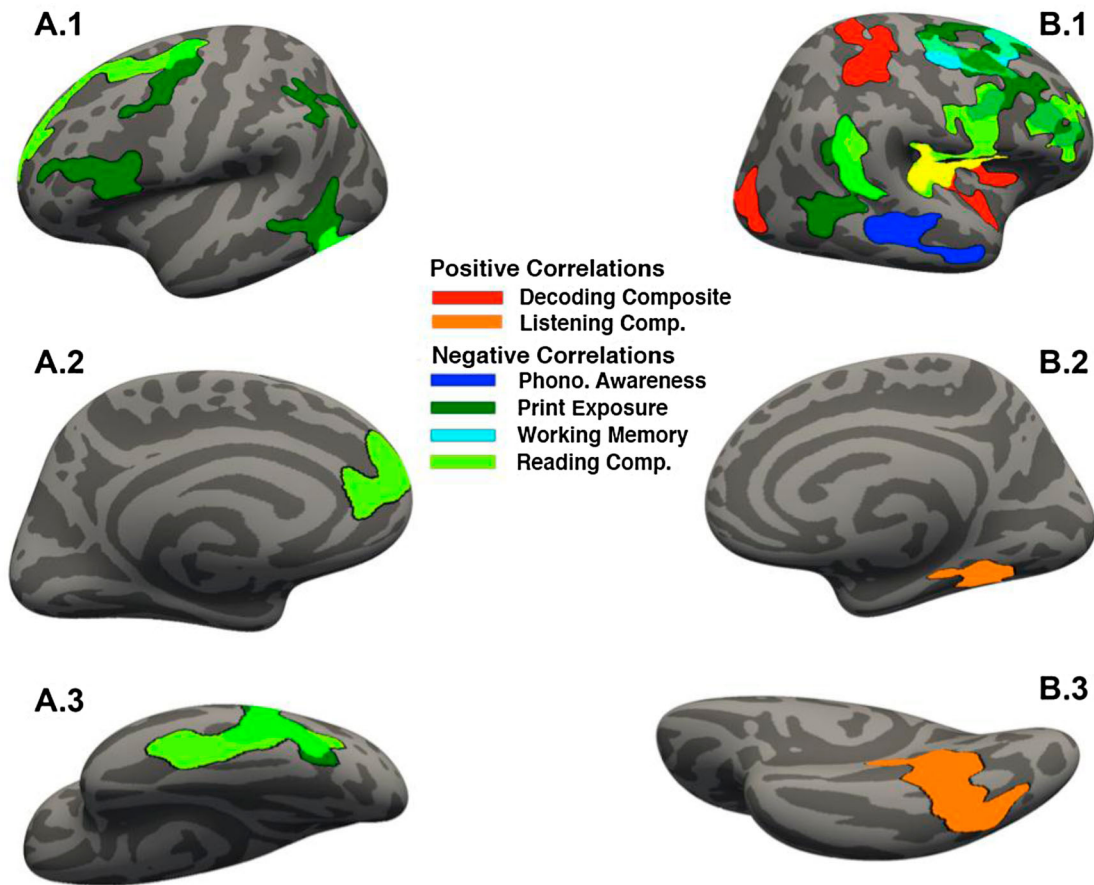


Figure 1. Cluster-corrected results for individual difference measures correlated with GMT projected onto the *fsaverage* template in FreeSurfer. All cluster-corrected results depicted at $p < 0.05$. (A) Results projected onto the left hemisphere; (B) results projected onto the right hemisphere. From top to bottom: (1) lateral view, (2) medial view, (3) ventral view.

Note: yellow indicates overlap between reading comprehension skill and the decoding composite.

the grey matter structure of this region and literacy skills in Figure 3.

4.1. Supramarginal gyrus

Our analysis confirms and extends several previous findings in this region, as well as producing significant novel

results. In our study, decoding ability was positively correlated with GMV in SMG. This confirms a previous finding in children (Jednoróg et al., 2015), and is also in line with our interpretation of the group differences observed between good and both poor and resilient readers (Welcome et al., 2011). In addition, we observed a significant negative correlation between print exposure

Table 5. Grey matter volume: individual difference measures, peak t -values, and centroid coordinates.

Measure	Cluster	Area (mm ²)	t	p	MNI Coordinates		
					x	y	z
Print Exposure	RH superior frontal	1133	-4.84	.0130	22.0	1.3	56.1
Working Memory	LH superior frontal	905	-3.03	.0432	-17.2	16.1	57.7
Reasoning	LH postcentral	895	-2.05	.0470	-53.3	-20.7	33.1
Reading Fluency Composite	LH lateral occipital	1261	-2.29	.0048	-22.8	-94.0	7.0
Decoding Composite	LH lateral occipital	1308	2.54	.0036	-21.5	-94.7	5.3
	LH SMG	1257	3.03	.0048	-52.3	-26.4	20.2
	LH lateral orbitofrontal	1244	2.89	.0058	-27.2	25.6	0.2
	LH medial orbitofrontal	1223	3.27	.0061	-7.7	41.7	-16.6
	LH superior frontal	1103	2.75	.0118	-8.7	11.5	54.4
	LH STS	920	3.23	.0394	-50.2	-45.4	8.2
	RH postcentral	3448	3.90	.0001	37.7	-7.1	17.7
	RH superior frontal	2889	5.05	.0001	7.9	50.5	17.6
	RH MTG	1148	3.01	.0121	57.6	-53.4	7.1
	RH lingual	1030	3.46	.0246	22.4	-53.6	6.6
	RH superior parietal	1002	2.81	.0291	22.0	-61.3	37.8

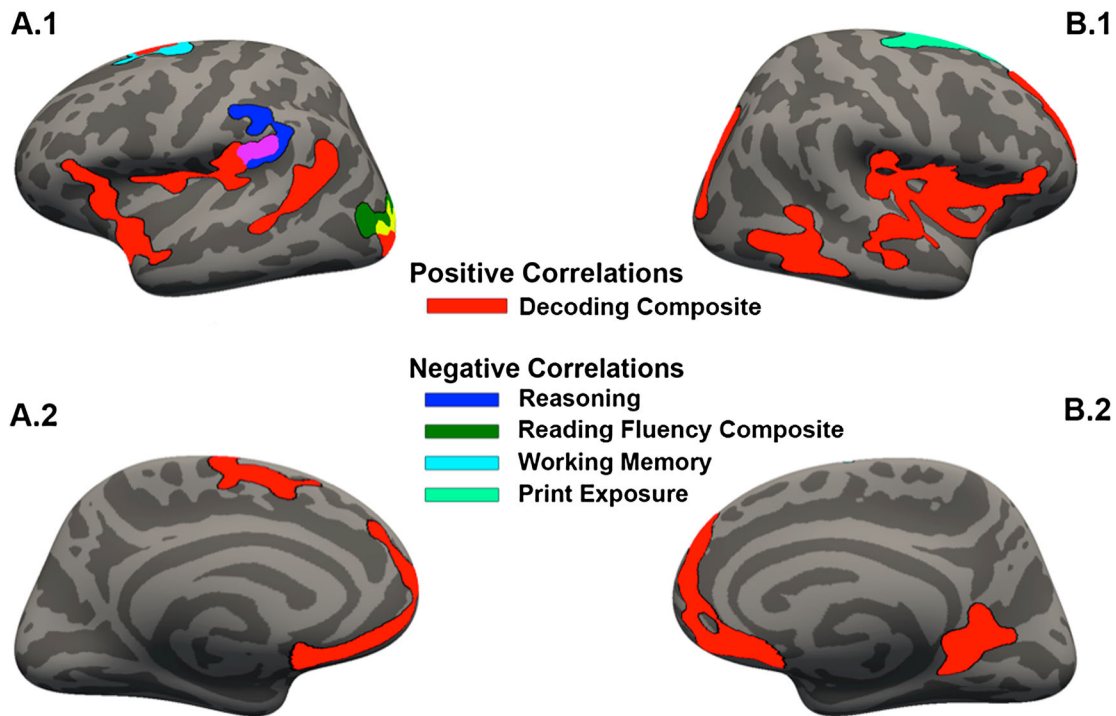


Figure 2. Cluster-corrected results for individual difference measures correlated with GMV projected onto the *fsaverage* template in FreeSurfer. All cluster-corrected results depicted at $p < 0.05$. (A) Results projected onto the left hemisphere; (B) results projected onto the right hemisphere. From top to bottom: (1) lateral view, (2) medial view.

Note: purple indicates overlap between matrix reasoning ability and the decoding composite; yellow indicates overlap between the reading fluency and decoding composite measures.

and GMT in left SMG, extending posteriorly to AG, indicating that greater experience with printed material was associated with *thinner* grey matter in our sample. This confirms the previous report of a relation between print exposure and GMT in SMG, although the correlation in that study was positive (Goldman & Manis, 2013). This discrepancy may stem from methodological differences, e.g. our use of a broad, community-based sample of participants, rather than a relatively skilled subgroup (i.e.

college students) of the population. Furthermore, our finding is also consistent with a hypothesis initially proposed by Goldman and Manis: that a negative correlation between cortical thickness and print exposure is compatible with longitudinal evidence that cortical thinning is a byproduct of maturation, possibly related to skill consolidation (as proposed in, e.g. Lu et al., 2007; see also Sowell, Thompson, Tessner, & Toga, 2001; Sowell et al., 2004). Finally, we found no relation between vocabulary

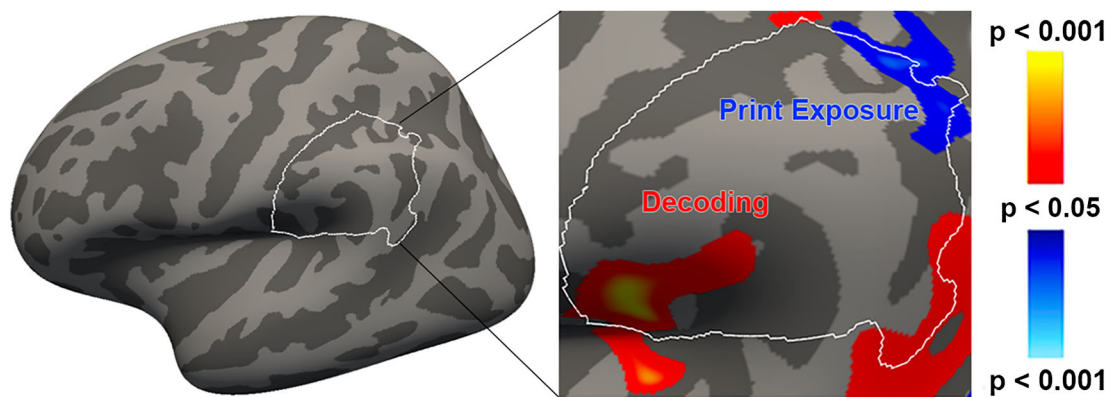


Figure 3. Left supramarginal gyrus: the bounded area is the Desikan-Killiany parcellation of SMG in MNI space. Positive correlations with GMV are depicted on a red-yellow scale, and negative correlations with GMV are depicted on a blue-white scale. Results depicted at an uncorrected vertex-wise threshold of $p < 0.05$.

knowledge and the structure of SMG in our young adult readers. This is consistent with previous findings indicated that such a relation was only present in adolescents between 12 and 17 years of age (Lee et al., 2007; Richardson et al., 2010), a range with which our sample barely overlaps.

Our results also dovetail with previous studies indicating that SMG can be bisected into distinct anterior and posterior regions. Here, the correlation with decoding efficiency was centred in anterior SMG, and the centroid of the correlation with print exposure was in the posterior region of SMG. Although ours is the first study to directly relate decoding to the structure of aSMG in monolingual young adult readers, this result is analogous to previous findings in children (Jednoróg et al., 2015) and bilingual young adults (L2 English; He et al., 2013). In the latter case, decoding was positively correlated with GMV in aSMG in a large sample ($N=253$) of native Chinese speakers. Our study is also the first to directly relate readers' experience with printed material to structural aspects of pSMG. Furthermore, print exposure is known to be correlated with vocabulary knowledge (Acheson et al., 2008; Long et al., 2008; Stanovich, 1993; Stanovich & West, 1989; Stanovich, West, & Harrison, 1995), as indeed it is in our data ($r = .552, p < .001$). Reading facilitates acquisition of novel word forms and the development of skilled spelling (Mol & Bus, 2011). Indeed, readers with greater print experience are more likely to encounter rare words in print than in spoken language (e.g. Carroll, Davies, & Richman, 1971), and potentially see ten times as many words as readers with more limited exposure (Nagy & Anderson, 1984). It is therefore no surprise that experience with printed matter is a significant vehicle by which new words are acquired (Cunningham, Stanovich, & West, 1994; Stanovich, 1986). Thus, the relation between grey matter in pSMG and print exposure that we report here is clearly not inconsistent with previously reported structural correlations between pSMG and vocabulary knowledge (Lee et al., 2007; Richardson et al., 2010).

4.2. Left inferior frontal gyrus

Our whole-brain analysis also found a relation between literacy skills and pars opercularis in left IFG. Pars opercularis – as part of Broca's Area – has long been associated with many language-related processes (for review, see Friederici, 2011). As in SMG, print exposure was negatively correlated with GMT in this area, indicating that greater experience with printed matter was associated with thinner cortex (see Figure 1). This confirms the previously reported link between print exposure and left IFG but, again, the direction of the correlation in our study

differs from the first report (Goldman & Manis, 2013). As previously noted, we suspect that the results diverge in this way due to methodological and analytic differences between the two studies.

Our findings align with those of Goldman and Manis (2013) in one regard: neither study obtained evidence for a correlation between decoding ability and GMT in IFG. Such a relation might have been expected because of two prior findings. First, longitudinal research suggests that the cortex thickens in IFG (and bilateral STG) over time in young children (Sowell et al., 2004). Second, thickening in pars opercularis during normal development in childhood has been related to increasingly proficient phonological awareness (Lu et al., 2007). Phonemic awareness is a necessary precursor to skilled decoding ability, and corresponds to the capacity to exploit knowledge about individual phonemes in a language (Scarborough & Brady, 2002). However, as with decoding ability, we found no evidence that phonological awareness correlated with either GMT or GMV in IFG, suggesting that the correlation observed in children reflects active neural development that may not be present in our more developmentally mature sample. Instead, in our participants, phonological awareness was negatively correlated with GMT in right middle temporal gyrus (see Figure 1), a region that is functionally important for discriminating sublexical speech sounds (Boets et al., 2013).

The absence of a relation between decoding skill and cortical structure in left IFG in our participants may also be related to the fact that decoding skill – which is strongly correlated with phonological awareness (in this study, $r = .624, p < .0001$) – decreases in importance as children grow into competent comprehenders, explaining progressively less variance related to comprehension (Goldman & Manis, 2013; Wagner et al., 1997). In contrast, the importance of experience with printed matter increases throughout maturation, from explaining 12% of the variance in oral language comprehension in kindergarten-age children to 34% in college-aged young adults (Mol & Bus, 2011). The pattern of relations that we observed for both decoding and print exposure in this region is consistent with these developmental shifts.

4.3. Left superior temporal sulcus

We observed a positive correlation between decoding ability and GMV in left STS (Figure 2). There is no analog for this finding in the previous literature investigating cortical structure in non-dyslexic readers (but see our discussion of clinical similarities below). Rather, this region was previously related to vocabulary knowledge in this population (Richardson et al., 2010), such

that greater vocabulary knowledge was associated with greater grey matter density. In addition, because the correlation was consistent across age groups, Richardson and colleagues suggested that this region might be related to processes of vocabulary acquisition that are independent of formal instruction (as opposed to their proposed relation between pSMG and vocabulary, to which we return below). We obtained no evidence that vocabulary was related to grey matter structure in this region. However, vocabulary knowledge in our sample was strongly correlated with decoding ($r = .684$, $p < .0001$), and this may be a case in which variance that might otherwise have been assigned to vocabulary knowledge may instead have been attributed to other measures that were not modelled in Richardson et al.

4.4. Right frontal/parietal regions

Several correlations emerged between cortical structure in right frontal and parietal areas and measures in our individual differences battery. As shown in Table 4, these include relations between GMT and print exposure, working memory capacity, phonological awareness, comprehension, and decoding ability; and per Table 5, between GMV and both print exposure and decoding ability. We know of no previous reports of structural relations between these brain regions and behavioural assessments of these literacy-related skills. Nevertheless, Welcome et al. (2011) observed group differences in radial expansion, a measure related to cortical surface area, in broadly-defined right frontal/parietal regions. Specifically, poor comprehenders showed smaller radial expansion in these areas than both proficient and resilient readers. To the extent that either GMV or GMT are related to radial expansion, some of our results loosely correspond with those of Welcome and colleagues. However, whereas Welcome and colleagues reported that radial expansion was relatively small in poor readers, some of our findings are in the opposite direction, with thinner GMT and/or smaller GMV associated with better performance on individual difference measures. In light of the unclear relation between radial expansion and more traditional indices of cortical grey matter structure (discussed in greater detail by Welcome et al., 2011, p. 1203), and the associated difficulty in interpreting the group differences in Welcome and colleagues' study relative to our continuous approach, we refrain from offering any interpretation of this difference.

5. General discussion

The goal of this study was to assess potential relations between a broad array of the components of literacy

and cortical grey matter structure in young adult readers. Although there are few previous research reports addressing this question, many aspects of our study make contact with earlier work, while also offering some methodological or analytic extensions. This study clearly demonstrates that mastery of fundamental skills related to proficient reading comprehension can be reflected in neurostructural characteristics of an array of language-relevant brain regions. The breadth of our cognitive assessments allowed us to both identify correlations between cortical structure and literacy skills that are consistent across previous studies in spite of methodological differences, and to assess whether individual literacy skills contributed uniquely to any structural variance observed in our imaging results. In addition, by measuring both GMV and GMT, we were able to make a detailed assessment of neuroanatomical relations to literacy skills – one which potentially accounts for differing patterns of results in these two measures. Finally, our use of a community-based sample confers several advantages to our study. This sample is more representative of the population at large than a convenience sample of undergraduates, and consequently the range of literacy and literacy-related abilities is greater than is typical in most neuroimaging studies of adult readers, which increases our ability to detect meaningful relationships between our assessments and grey matter structure.

On balance, our findings indicate that two measures – print exposure and decoding ability – are associated with grey matter structure in regions known to be related to language comprehension in general and, as we will describe below, to reading-specific behaviours in particular. This latter point is not trivial, since both print exposure and decoding skill reflect a reader's interaction with *orthographic* input. Both are particularly important to the development of orthographic processing skill (i.e. the ability to construct, encode, and retrieve orthographic representations), which is thought to explain unique variance in word recognition (Cunningham & Stanovich, 1990, 1997; Stanovich, 1986; Stanovich & West, 1989). Skilled orthographic processing strengthens readers' awareness of the phonological parameters of their language (Ehri, 1984; Nation & Hulme, 2011), enabling a tight link between words' orthographic and phonological representations (Barron, 1986; Ehri, 1980, 1987) and, ultimately, facilitating efficient lexical access by means of these representations.

It is therefore notable that many of the cortical regions that we have discussed thus far have been functionally linked to establishing relations between orthographic and phonological information. For example, aSMG (but not pSMG) is active during the conversion of

orthographic input to licit phonological information (e.g. Booth et al., 2002; Hartwigsen et al., 2010; Richlan, Kronbichler, & Wimmer, 2009; Tan et al., 2005; see also Heim et al., 2010, discussed below). In IFG, indices of print exposure have been functionally related to this area's role in print-speech convergence, such that greater print exposure was related to greater functional overlap during the processing of spoken and written sentences (Shankweiler et al., 2008). Left STS has been implicated in processes of orthography-to-phonology conversion (van Atteveldt, Formisano, Goebel, & Blomert, 2004), with dysfunction in this region indicating impaired translation of novel graphemes to licit phonemes. (Perhaps unsurprisingly, left STS is also proximate to heteromodal language processing regions associated with print-speech convergence; Braze et al., 2011; Frost et al., 2009; Shankweiler et al., 2008.) And the importance of left fusiform gyrus (and the VWFA) to word reading is of course well known (McCandliss, Cohen, & Dehaene, 2003).

Further, many of the right hemisphere regions in which we observed structural correlations with literacy skills have been functionally implicated in studies of semantic and syntactic relations, sentence comprehension, and discourse processing (e.g. Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006; Petersson, Folia, & Hagoort, 2012; Prat & Just, 2011; Prat, Mason, & Just, 2011; Robertson et al., 2000; St. George, Kutas, Martinez, & Sereno, 1999; Snijders et al., 2009). Right hemisphere homologs of "classical" left hemisphere regions associated with language processing are often concurrently activated during processing, and as such they might best be considered as part of an extended cortical network related to language comprehension in general (Hagoort, 2009).

In addition to the correspondence between the structural relations to literacy-related skills in this study and previous functional research, our results also support previous reports of a unique structural relation between pSMG and reading-related skills that has no functional analog. In previous research, grey matter density in this region was correlated with vocabulary, with greater density indicative of a more extensive vocabulary knowledge (Lee et al., 2007; Richardson et al., 2010; see also Mechelli et al., 2004). Given its direct connections to both aSMG and AG, these studies concluded that pSMG is likely a binding site for orthographic, phonological, and semantic information. In addition, they proposed that because this relation was obtained only in adolescents, it might be a cortical reflection of formal instruction in the service of vocabulary acquisition (see also Richardson & Price, 2009).

We concur with the broader conclusion of these studies regarding SMG's potential relevance to the assembly of high-dimensional lexical representations. However, our observation of correlations between decoding ability and aSMG, and print exposure and pSMG, suggest a revised account of the link between pSMG and vocabulary. We agree that aSMG may be chiefly concerned with phonological aspects of grapheme-to-phoneme translation. We suggest that pSMG, in contrast, may be primarily concerned with orthographic aspects of this process; that is, the structure of pSMG may not reflect the number of lexical forms added to a reader's vocabulary, or the way in which these forms were acquired (Lee et al., 2007; Richardson & Price, 2009; Richardson et al., 2010), but the *amount* of reading that a person actually does.³ This explanation is more parsimonious than the proposal that pSMG is related to vocabulary instruction – an explanation that was always somewhat problematic given that the relation with vocabulary knowledge was not present in children, in whom increasing vocabulary knowledge has instead been related to cortical *thinning* in left parietal regions (Sowell et al., 2004; see also Linkersdörfer et al., 2015). This explanation is also consistent with SMG's well-established functional role in linking graphemes to their phonemic equivalents during reading.

It is worth noting that our study reports different cortical grey matter measurements than those studies examining the pSMG and vocabulary knowledge. We report correlations with GMT and GMV, whereas Lee et al. (2007) and Richardson et al. (2010) assessed relations to grey matter density, an index of the ratio of grey matter voxels to those of other types of brain tissue in a particular cortical region (for a detailed discussion, see Mechelli, Price, Friston, & Ashburner, 2005). Because grey matter density, GMT, and GMV are partially independent measures, it is reasonable to expect them to show different patterns of results. Further, as we discussed in section 4.1 above, experience with printed material is an established precursor to vocabulary knowledge (see also Cunningham & Stanovich, 1991; Stanovich & Cunningham, 1992; Stanovich & West, 1989). By assessing print exposure and vocabulary together, our study examined whether these measures explained unique variance in our indices of cortical structure. Our findings in this region (and others, such as left STS) suggests that it is important to assess skills linked to multiple components of reading ability, rather than a single measure (or a small set of measures), in order to determine whether effects are related to common variance shared among the measures, or specific contributions of particular cognitive abilities.

Finally, there is also substantial agreement between our findings and the results of neuropsychological studies of group differences between dyslexic and typical readers. Although there is considerable heterogeneity among the results of such studies, two recent meta-analyses clearly identify several points of convergence. For example, Linkersdörfer, Lonnemann, Lindberg, Hasselhorn, and Fiebach (2012, $N=277$) identified four cortical areas in which dyslexic readers consistently exhibit smaller GMV: bilateral SMG, left fusiform gyrus (including VWFA) and the right superior temporal gyrus (STG). In contrast, Richlan, Kronbichler, and Wimmer (2013, $N=266$) found only two: right STG and left STS. In our study, we observed significant relations between literacy skills and *all* of these regions that are consistent with this literature. Decoding ability was positively correlated with GMV in left SMG and left STS, and with GMT in right STG in our community-based sample, such that poor decoders also had smaller GMV/thinner cortex than those with more efficient decoding ability. Given the obvious relation between dyslexia and decoding ability, the parallel between our results and the clinical pattern is clear, with some functional imaging research explicitly suggesting that these group differences are related to impaired grapheme-to-phoneme conversion (e.g. Heim et al., 2010). Furthermore, we also observed a relation between GMT and both print exposure and reading comprehension in an area of left inferior temporal gyrus (ITG) corresponding to the fusiform gyrus. Here the correlation was negative, with greater experience with printed matter and greater reading comprehension ability coinciding with thinner grey matter. However, we do not believe this difference suggests any meaningful conflict between ours and the neuropsychological studies. Rather, the relation of GMV to group differences related to dyslexia and, in our study, of GMT and the development of skilled reading comprehension are parallel, rather than contradictory, findings.

6. Conclusions and future directions

Skilled reading is the result of the fluent orchestration of numerous component processes, some of which are thought to be linguistically-based, whereas others are domain-general. The results of this exploratory study confirm, extend, and refine many of the diverse findings of previously published work. In addition, this study clearly demonstrates the importance of taking a multifaceted approach to investigating the neurostructural correlates of skilled reading comprehension. This is true not just for future investigations of cortical grey matter, but for investigations of other aspects of brain morphology

as well. For example, there are a small number of studies in which literacy skills have been correlated with white matter structure. We have mentioned one such study, in which vocabulary knowledge informed our understanding of the structural significance of pSMG (Lee et al., 2007). There are others exploring relations between the connectivity of regions in the left-hemisphere reading network and, for example, decoding ability (Welcome & Joanisse, 2014; Zhang et al., 2014; in dyslexic readers, see Pugh et al., 2000; Steinbrink et al., 2008). Taking white matter structure into account could be particularly important since increased white matter myelination decreases the space between the cortex and the skull, potentially contributing to grey matter thinning (Lu et al., 2007; Sowell et al., 2004). However, as in previous investigations of cortical grey matter, most such studies have been narrowly construed, suggesting that a broad-based approach to assessing participants' literacy skills could be similarly applied, with the same potential benefits to analysis and interpretation (e.g. Van Dyke et al., 2015).

Indeed, we view this approach to assessing individual cognitive differences as essential to future research into the neural architecture supporting reading comprehension. This is consistent with recent behavioural research on the relation between individual cognitive variation and skilled reading. Although our own battery of assessments is more extensive than previous work, it is hardly exhaustive; large-scale behavioural studies typically administer many more assessments than were included in our project. For example, studies such as Van Dyke et al. (2014) and Freed et al. (2017) both assessed performance on more than two dozen measures. The benefits offered by such an approach are twofold. First, expanding the battery of assessments allows the testing of theoretically important constructs for which previous research (including our own) does not account. For example, as noted by Freed and colleagues, individual differences in both perceptual speed and suppression/inhibition ability, neither of which have been assessed in studies such as ours, are both likely contributors to comprehension performance. Second, more extensive test batteries allow theoretical constructs to be measured with multiple assessments, which would permit the construction of more stable composite predictors (cf. Braze et al., 2016). This is particularly important for non-standardised instruments, such as the complex span tasks that are in common use to assess working memory capacity: the test-retest reliability of these measures is known to be relatively low, and a composite based on two or three assessments (e.g. sentence span, operation span, symmetry span, etc.) has much greater predictive stability (Waters & Caplan, 2003).

Finally, as noted in our introduction, most of this research has been conducted with adult readers. However, studies such as Lee et al. (2007) and Richardson et al. (2010) make it clear that assessing developmentally distinct groups of readers is important for understanding neurostructural correlates of literacy-related skills. For example, as discussed above, experience with printed material *increases* in importance as readers age, while decoding ability experiences a corresponding decrease in importance (Mol & Bus, 2011); this yields testable predictions about both when and where these abilities should have correlates in the brain. In light of evidence that both grey and white matter are sensitive to efforts at remediation in poor readers (e.g. Keller & Just, 2009; Krafnick, Flowers, Napoliello, & Eden, 2011), a comprehensive assessment of both cortical structure and the precise cognitive capacities that support the achievement of literacy is essential for our understanding of the neural architecture supporting skilled reading comprehension.

Notes

1. Although, as noted in the Method, we have reservations about using variance inflation factors (VIFs) to index multicollinearity, we nonetheless calculated them for our predictors. The VIFs converged with the results of the condition numbers: their range (1.43–6.08) was not only below the commonly recommended threshold of 10 (e.g. Cohen, Cohen, West, & Aiken, 2013; James, Witten, Hastie, & Tibshirani, 2013; Kutner, Nachtsheim, Neter, & Li, 2005), but also below a more conservative recent recommendation of 7 (Keith, 2014).
2. In line with recent recommendations (e.g. Eklund, Nichols, & Knutsson, 2017), all structural MRI data and analysis scripts are available at [OpenNeuro.org](https://openneuro.org), a free online repository for neuroimaging data: <https://openneuro.org/datasets/ds001365/versions/00001>.
3. Similarly, Goldman and Manis (2013) observed a relation between SMG and the amount of pleasure reading their participants engaged in outside of instructional settings. Although the absence of MNI coordinates in this study does not permit assessment of the locus of their SMG correlations, we mention it here as weak converging evidence for our account of pSMG's structural relation to reader characteristics.

Acknowledgements

We are grateful to Jessica Grittner for assistance with data collection, to Morgan L. Bontrager, Alexis R. Johns, Leonard Katz, Victor Kuperman, and Ashley G. Lewis for insightful comments and advice on earlier drafts of this paper, and to P. R. Nelson for support and inspiration.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the following grants to Haskins Laboratories from the Eunice Kennedy Shriver National Institute of Child Health and Human Development: P01 HD-01994 (Jay G. Rueckl, PI), R01 HD-040353 (Donald P. Shankweiler, PI), R01 HD-071988 (David Braze, PI), R01 HD-073288 (Julie A. Van Dyke, PI), and NRSA HD-080331 (Dave Kush, PI).

ORCID

Clinton L. Johns  <http://orcid.org/0000-0003-0593-7508>

References

- Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278–289.
- Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, *23*(3), 285–303.
- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Barron, R. W. (1986). Word recognition in early reading: A review of the direct and indirect access hypotheses. *Cognition*, *24*(1–2), 93–119.
- Belsley, D. A. (1991a). *Conditioning diagnostics: Collinearity and weak data in regression*. New York, NY: John Wiley & Sons, Inc.
- Belsley, D. A. (1991b). A guide to using the collinearity diagnostics. *Computational Economics*, *4*(1), 33–50.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity* (Vol. 571). New York, NY: John Wiley & Sons.
- Bizzo, B. C., Sanchez, T. A., Tukamoto, G., Zimmermann, N., Netto, T. M., & Gasparetto, E. L. (2017). Cortical thickness and episodic memory impairment in systemic lupus erythematosus. *Journal of Neuroimaging*, *27*(1), 122–127.
- Boets, B., de Beeck, H. P. O., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., ... Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, *342*, 1251–1254.
- Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping*, *25*(1), 92–104.
- Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2002). Functional anatomy of intra- and cross-modal lexical tasks. *NeuroImage*, *16*(1), 7–22.
- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*, 211–252.
- Braze, D., Katz, L., Magnuson, J. S., Mencl, W. E., Tabor, W., Van Dyke, J. A., ... Shankweiler, D. P. (2016). Vocabulary does not complicate the simple view of reading. *Reading and Writing*, *29*(3), 435–451.
- Braze, D., Mencl, W. E., Tabor, W., Pugh, K. R., Constable, R. T., Fulbright, R. K., ... Shankweiler, D. P. (2011). Unification of sentence processing via ear and eye: An fMRI study. *Cortex*, *47*(4), 416–431.

- Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary reading skill differences in young adults. *Journal of Learning Disabilities, 40*(3), 226–243.
- Carreiras, M., Seghler, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature, 461*(7266), 983–986.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. Boston, MA: Houghton Mifflin.
- Castro-Caldas, A., Miranda, P. C., Carmo, I., Reis, A., Leote, F., Ribeiro, C., & Ducla-Soares, E. (1999). Influence of learning to read and write on the morphology of the corpus callosum. *European Journal of Neurology, 6*(1), 23–28.
- Chiarello, C., Lombardino, L. J., Kacinik, N. A., Otto, R., & Leonard, C. M. (2006). Neuroanatomical and behavioral asymmetry in an adult compensated dyslexic. *Brain and Language, 98*(2), 169–181.
- Clark, N. B., McRoberts, G. W., Van Dyke, J. A., Shankweiler, D. P., & Braze, D. (2012). Immediate memory for pseudowords and phonological awareness are associated in adults and pre-reading children. *Clinical Linguistics & Phonetics, 26*(7), 577–596.
- Clements-Stephens, A. M., Materek, A. D., Eason, S. H., Scarborough, H. S., Pugh, K. R., Rimrodt, S., ... Cutting, L. E. (2012). Neural circuitry associated with two different approaches to novel word learning. *Developmental Cognitive Neuroscience, 2*, S99–S113.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017a). fMRI clustering in AFNI: False-positive rates redux. *Brain Connectivity, 7*(3), 152–171.
- Cox, R. W., Chen, G., Glen, D. R., Reynolds, R. C., & Taylor, P. A. (2017b). fMRI clustering and false-positive rates. *Proceedings of the National Academy of Sciences, 114*(17), E3374–E3375.
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 102*(3), 687–700.
- Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733–740.
- Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*(2), 264–274.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934–945.
- Cunningham, A. E., Stanovich, K. E., & West, R. F. (1994). Literacy environment and the development of children's cognitive skills. In E. M. H. Assink (Ed.), *Literacy acquisition and social context: Approaches, emphases, and questions* (pp. 70–90). London: Harvester-Wheatsheaf.
- Curtis, M. E. (2002). *Adolescent reading: A synthesis of research*. Paper presented as U.S. Department of Education and National Institute of Child Health and Human Development Conference on Adolescent Literacy. Retrieved from <http://216.26.160.105/conf/nichd/synthesis.asp>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage, 9*(2), 179–194.
- Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience, 5*(2), 162–176.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466.
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115*(6), 1753–1768.
- Devlin, J. T., Matthews, P. M., & Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience, 15*(1), 71–84.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Services.
- Eckert, M. (2004). Neuroanatomical markers for dyslexia: A review of dyslexia structural imaging studies. *The Neuroscientist, 10*(4), 362–371.
- Ehri, L. C. (1980). The development of orthographic images. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 311–338). London: Academic Press.
- Ehri, L. C. (1984). How orthography alters spoken language competencies in children learning to read and spell. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 119–147). New York, NY: Springer.
- Ehri, L. C. (1987). Learning to read and spell words. *Journal of Literacy Research, 19*(1), 5–31.
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences, 113*(28), 7900–7905.
- Eklund, A., Nichols, T. E., & Knutsson, H. (2017). Reply to Brown and Behrmann, Cox, et al., and Kessler et al.: Data and code sharing is the way forward for fMRI. *Proceedings of the National Academy of Sciences, 114*(17), E3374–E3375.
- Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing Company.
- Faraway, J. J. (2014). *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models* (2nd ed.). New York, NY: CRC press.
- Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences, 97*(20), 11050–11055.
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron, 33*(3), 341–355.
- Fischl, B., Salat, D., Kennedy, D., Makris, N., Albert, M., Killiany, R., & Dale, A. (2001). Automatic segmentation of the structures in the human brain. *NeuroImage, 13*(6), 118.
- Fischl, B., Salat, D. H., van der Kouwé, A. J., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage, 23*, S69–S84.

- Fjell, A. M., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., ... Walhovd, K. B. (2009). High consistency of regional cortical thinning in aging across multiple samples. *Cerebral Cortex*, 19(9), 2001–2012.
- Flom, P. (1999). *Multicollinearity diagnostics for multiple regression: A Monte Carlo study*. Unpublished doctoral dissertation, New York University, NYC.
- Freed, E. M., Hamilton, S. T., & Long, D. L. (2017). Comprehension in proficient readers: The nature of individual variation. *Journal of Memory and Language*, 97, 135–153.
- Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91(4), 1357–1392.
- Frost, S. J., Landi, N., Mencl, W. E., Sandak, R., Fulbright, R. K., Tejada, E. T., ... Pugh, K. R. (2009). Phonological awareness predicts activation patterns for print and speech. *Annals of Dyslexia*, 59(1), 78–97.
- Frye, R. E., Liederman, J., Malmberg, B., McLean, J., Strickland, D., & Beauchamp, M. S. (2010). Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. *Cerebral Cortex*, 20(11), 2625–2635.
- Gardumi, A., Ivanov, D., Havlicek, M., Formisano, E., & Uludağ, K. (2017). Tonotopic maps in human auditory cortex using arterial spin labeling. *Human Brain Mapping*, 38(3), 1140–1154.
- Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, 13(1), 65–77.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2(10), 861–863.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., & Rapoport, J. L. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, 101(21), 8174–8179.
- Goldman, J. G., & Manis, F. R. (2013). Relationships among cortical thickness, reading skill, and print exposure in adults. *Scientific Studies of Reading*, 17(3), 163–176.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10.
- Greve, D. N., Van der Haegen, L., Cai, Q., Stufflebeam, S., Sabuncu, M. R., Fischl, B., & Brysbaert, M. (2013). A surface-based analysis of language lateralization and cortical asymmetry. *Journal of Cognitive Neuroscience*, 25(9), 1477–1492.
- Guo, Y., Roehrig, A. D., & Williams, R. S. (2011). The relation of morphological awareness and syntactic awareness to adults' reading comprehension: Is vocabulary knowledge a mediating variable? *Journal of Literacy Research*, 43(2), 159–183.
- Hagler, D. J., Saygin, A. P., & Sereno, M. I. (2006). Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage*, 33(4), 1093–1103.
- Hagoort, P. (2009). Reflections on the neurobiology of syntax. In D. Bickerton & E. Szathmáry (Eds.), *Biological foundations and origin of syntax* (pp. 279–296). Cambridge, MA: MIT Press.
- Harrell, F. R. (2001). *Regression modeling strategies*. Berlin: Springer.
- Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right Supramarginal Gyri. *Proceedings of the National Academy of Sciences*, 107(38), 16494–16499.
- 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. *The Journal of Neuroscience*, 33(31), 12835–12843.
- Heim, S., Grande, M., Meffert, E., Eickhoff, S. B., Schreiber, H., Kukolja, J., & Amunts, K. (2010). Cognitive levels of performance account for hemispheric lateralisation effects in dyslexic and normally reading children. *NeuroImage*, 53(4), 1346–1358.
- Heim, S., & Keil, A. (2004). Large-scale neural correlates of developmental dyslexia. *European Child & Adolescent Psychiatry*, 13(3), 125–140.
- Hendrickx, J. (2012). *perturb: Tools for evaluating collinearity* [Software]. R package version 2.05. Retrieved from <https://CRAN.R-project.org/package=perturb>
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127–160.
- Hua, A. N., & Keenan, J. M. (2014). The role of text memory in inferring and in comprehension deficits. *Scientific Studies of Reading*, 18(6), 415–431.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York, NY: Springer.
- Jasińska, K. K., Molfese, P. J., Kornilov, S. A., Mencl, W. E., Frost, S. J., Lee, M., & Landi, N. (2017). The BDNF Val 66 Met polymorphism is associated with structural neuroanatomical differences in young children. *Behavioural Brain Research*, 328, 48–56.
- Jednoróg, K., Marchewka, A., Altarelli, I., Monzalvo Lopez, A. K., van Ermingen-Marbach, M., Grande, M., & 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. *Human Brain Mapping*, 36(5), 1741–1754.
- Johns, C. L., Matsuki, K., & Van Dyke, J. A. (2015). Poor readers' retrieval mechanism: Efficient access is not dependent on reading skill. *Frontiers in Psychology*, 6, 233.
- Joshi, R. M. (2005). Vocabulary: A critical component of comprehension. *Reading & Writing Quarterly*, 21(3), 209–219.
- Karlsen, B., & Gardner, E. (1995). *Stanford diagnostic reading test* (4th ed.). San Antonio, TX: Psychological Corp.
- Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling* (2nd ed.). New York, NY: Routledge.
- Keller, T. A., & Just, M. A. (2009). Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. *Neuron*, 64(5), 624–631.
- Krafnick, A. J., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *NeuroImage*, 57(3), 733–741.
- Kuhn, M. (2016). *caret: Classification and Regression Training* [Software]. R package version 6.0-71. Retrieved from <https://CRAN.R-project.org/package=caret>
- Kukona, A., Braze, D., Johns, C. L., Mencl, W. E., Van Dyke, J. A., Magnuson, J. S., ... Tabor, W. (2016). The real-time prediction and inhibition of linguistic outcomes: Effects of language and literacy skill. *Acta Psychologica*, 171, 72–84.
- Kuperberg, G. R., Broome, M. R., McGuire, P. K., David, A. S., Eddy, M., Ozawa, F., ... Fischl, B. (2003). Regionally localized thinning of the cerebral cortex in schizophrenia. *Archives of General Psychiatry*, 60(9), 878–888.

- Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *NeuroImage*, 33(1), 343–361.
- Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, 65(1), 42–73.
- Kutner, M. H., Nachtsheim, C., Neter, J., & Li, W. (2005). *Applied linear regression models* (5th ed.). New York, NY: McGraw-Hill/Irwin.
- Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, 95(2), 211–224.
- Lee, H., Devlin, J. T., Shakeshaft, C., Stewart, L. H., Brennan, A., Glensman, J., & Price, C. J. (2007). Anatomical traces of vocabulary acquisition in the adolescent brain. *Journal of Neuroscience*, 27(5), 1184–1189.
- Li, M., Braze, D., Kukona, A., Johns, C. L., Tabor, W., Van Dyke, J. A., Mencl, W. E., ... Magnuson, J. S. (2017). Individual differences in subphonemic sensitivity and reading-related abilities. Manuscript under review, Haskins Laboratories. <http://doi.org/10.17605/OSF.IO/3FNCY>
- Linkersdörfer, J., Jurcoane, A., Lindberg, S., Kaiser, J., Hasselhorn, M., Fiebach, C. J., & Lonnemann, J. (2015). The association between gray matter volume and reading proficiency: A longitudinal study of beginning readers. *Journal of Cognitive Neuroscience*, 27(2), 308–318.
- Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., & Fiebach, C. J. (2012). Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: An ALE meta-analysis. *PLoS one*, 7(8), e43122.
- Lohmann, G., Stelzer, J., Mueller, K., Lacosse, E., Buschmann, T., Kumar, V. J., & Scheffler, K. (2017). Inflated false negative rates undermine reproducibility in task-based fMRI. *bioRxiv*. doi:10.1101/122788
- Long, D. L., Johns, C. L., & Morris, P. E. (2006). Comprehension ability in mature readers. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 801–833). Burlington, MA: Academic Press.
- Long, D. L., Prat, C., Johns, C., Morris, P., & Jonathan, E. (2008). The importance of knowledge in vivid text memory: An individual-differences investigation of recollection and familiarity. *Psychonomic Bulletin & Review*, 15(3), 604–609.
- Lu, L. H., Dapretto, M., O'Hare, E. D., Kan, E., McCourt, S. T., Thompson, P. M., ... Sowell, E. R. (2009). Relationships between brain activation and brain structure in normally developing children. *Cerebral Cortex*, 19(11), 2595–2604.
- Lu, L. H., Leonard, C. M., Thompson, P. M., Kan, E., Jolley, J., Welcome, S. E., ... Sowell, E. R. (2007). Normal developmental changes in inferior frontal gray matter are associated with improvement in phonological processing: A longitudinal MRI analysis. *Cerebral Cortex*, 17(5), 1092–1099.
- Macaruso, P., & Shankweiler, D. (2010). Expanding the simple view of reading in accounting for reading skills in community college students. *Reading Psychology*, 31(5), 454–471.
- Magnuson, J., Kukona, A., Braze, D., Johns, C. L., Van Dyke, J. A., Tabor, W., ... Shankweiler, D. (2011). Phonological instability in young adult poor readers: Time course measures and computational modeling. In P. McCardle, B. Miller, J. R. Lee, & O. J. L. Tzeng (Eds.), *Dyslexia across languages* (pp. 184–201). Baltimore, MD: Paul H. Brookes Publishing Co.
- Markwardt, F. C., Jr. (1998). *Peabody individual achievement test-revised*. Circle Pines, MN: American Guidance Service.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293–299.
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2nd ed.). New York, NY: CRC Press.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature*, 431, 757.
- Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-based morphometry of the human brain: Methods and applications. *Current Medical Imaging Reviews*, 1(2), 105–113.
- Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D., & Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia*, 46(10), 2580–2592.
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267–296.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19(3), 304–330.
- Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science*, 14(4), 649–659.
- Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, 63, 427–452.
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, 9(2), 60–68.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383.
- Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, 28, 450–461.
- Petersson, K. M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, 120(2), 83–95.
- Petersson, K. M., Silva, C., Castro-Caldas, A., Ingvar, M., & Reis, A. (2007). Literacy: A cultural influence on functional left-right differences in the inferior parietal cortex. *European Journal of Neuroscience*, 26(3), 791–799.
- Prat, C. S., & Just, M. A. (2011). Exploring the neural dynamics underpinning individual differences in sentence comprehension. *Cerebral Cortex*, 21(8), 1747–1760.
- Prat, C. S., Mason, R. A., & Just, M. A. (2011). Individual differences in the neural basis of causal inferencing. *Brain and Language*, 116(1), 1–13.
- Price, C. J., Moore, C. J., Humphreys, G. W., & Wise, R. J. (1997). Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience*, 9(6), 727–733.
- Protopapas, A., Mouzaki, A., Sideridis, G. D., Kotsolakou, A., & Simos, P. G. (2013). The role of vocabulary in the context of the simple view of reading. *Reading & Writing Quarterly*, 29(2), 168–202.

- Psychological Corp. (1999). *WASI: Wechsler abbreviated scale of intelligence*. San Antonio, TX: Author.
- Pugh, K. (2006). A neurocognitive overview of reading acquisition and dyslexia across languages. *Developmental Science*, 9(5), 448–450.
- Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Rueckl, J. G., Constable, R. T., ... Mencl, W. E. (2008). Effects of stimulus difficulty and repetition on printed word identification: An fMRI comparison of nonimpaired and reading-disabled adolescent cohorts. *Journal of Cognitive Neuroscience*, 20(7), 1146–1160.
- Pugh, K. R., Landi, N., Preston, J. L., Mencl, W. E., Austin, A. C., Sibley, D., ... Frost, S. J. (2013). The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain and Language*, 125(2), 173–183.
- Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., ... Liberman, A. M. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *Psychological Science*, 11(1), 51–56.
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Richardson, F. M., & Price, C. J. (2009). Structural MRI studies of language function in the undamaged brain. *Brain Structure and Function*, 213(6), 511–523.
- Richardson, F. M., Thomas, M. S., Filippi, R., Harth, H., & Price, C. J. (2010). Contrasting effects of vocabulary knowledge on temporal and parietal brain structure across lifespan. *Journal of Cognitive Neuroscience*, 22(5), 943–954.
- Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping*, 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. *Human Brain Mapping*, 34(11), 3055–3065.
- Robertson, D. A., Gernsbacher, M. A., Guidotti, S. J., Robertson, R. R., Irwin, W., Mock, B. J., & Campana, M. E. (2000). Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. *Psychological Science*, 11(3), 255–260.
- Rosas, J., Ramos-Casals, M., Ena, J., García-Carrasco, M., Verdu, J., Cervera, R., & Pascual, E. (2002). Usefulness of basal and pilocarpine-stimulated salivary flow in primary Sjogren's syndrome. Correlation with clinical, immunological and histological features. *Rheumatology*, 41(6), 670–675.
- Sabatini, J. P., Sawaki, Y., Shore, J. R., & Scarborough, H. S. (2010). Relationships among reading skills of adults with low literacy. *Journal of Learning Disabilities*, 43(2), 122–138.
- Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., & Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral Cortex*, 14(7), 721–730.
- Scarborough, H. S., & Brady, S. A. (2002). Toward a common terminology for talking about speech and reading: A glossary of the “phon” words and some related terms. *Journal of Literacy Research*, 34(3), 299–336.
- Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, 22(3), 1060–1075.
- Ségonne, F., Pacheco, J., & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Transactions on Medical Imaging*, 26(4), 518–529.
- Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., ... Shaywitz, B. A. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading*, 3(1), 69–94.
- Shankweiler, D., Mencl, W. E., Braze, D., Tabor, W., Pugh, K. R., & Fulbright, R. K. (2008). Reading differences and brain: Cortical integration of speech and print in sentence processing varies with reader skill. *Developmental Neuropsychology*, 33(6), 745–775.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.
- Share, D. L. (1995). Phonological recoding and self-teaching: *Sine qua non* of reading acquisition. *Cognition*, 55(2), 151–218.
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, 17(1), 87–97.
- Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, 19(7), 1493–1503.
- Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience*, 24(38), 8223–8231.
- Sowell, E. R., Thompson, P. M., Tessner, K. D., & Toga, A. W. (2001). Mapping continued brain growth and gray matter density reduction in dorsal frontal cortex: Inverse relationships during postadolescent brain maturation. *Journal of Neuroscience*, 21(22), 8819–8829.
- Spring, C., & French, L. (1990). Identifying children with specific reading disabilities from listening and reading discrepancy scores. *Journal of Learning Disabilities*, 23(1), 53–58.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.
- Stanovich, K. E. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. *Advances in Child Development and Behavior*, 24, 133–180.
- Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, 20(1), 51–68.
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24(4), 402–433.
- Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology*, 31(5), 811–826.
- Steinbrink, C., Vogt, K., Kastrup, A., Müller, H. P., Juengling, F. D., Kassubek, J., & Riecker, A. (2008). The contribution of white and gray matter differences to developmental dyslexia: Insights from DTI and VBM at 3.0T. *Neuropsychologia*, 46(13), 3170–3178.
- St. George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: Engagement of the right

- hemisphere during discourse processing. *Brain*, 122(7), 1317–1325.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Tan, L. H., Laird, A. R., Li, K., & Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Human Brain Mapping*, 25(1), 83–91.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency*. Austin, TX: Pro-Ed.
- Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, 45(5), 453–466.
- van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, 43(2), 271–282.
- van den Broek, P., Mouw, J. M., & Kraal, A. (2015). Individual differences in reading comprehension. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 138–150). New York, NY: Routledge.
- Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, 131(3), 373–403.
- Van Dyke, J. A., Matsuki, K., Jones, H. R., Molfese, P. J., Jahn, A. A., Johns, C. L., ... Bontrager, M. L. (2015). *A random forests analysis of adding-related skills and white matter tractography*. Poster presented at the 7th annual Society for the Neurobiology of Language Conference, 15 October.
- Wagner, R. K., Piasta, S. B., & Torgesen, J. K. (2006). Learning to read. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 1111–1142). Burlington, MA: Academic Press.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-Ed.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., ... Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, 33(3), 468–479.
- Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, 35(4), 550–564.
- Welcome, S. E., Chiarello, C., Thompson, P. M., & Sowell, E. R. (2011). Reading skill is related to individual differences in brain structure in college students. *Human Brain Mapping*, 32(8), 1194–1205.
- Welcome, S. E., & Joanisse, M. F. (2012). Individual differences in skilled adult readers reveal dissociable patterns of neural activity associated with component processes of reading. *Brain and Language*, 120(3), 360–371.
- Welcome, S. E., & Joanisse, M. F. (2014). Individual differences in white matter anatomy predict dissociable components of reading skill in adults. *NeuroImage*, 96, 261–275.
- Westman, E., Aguilar, C., Muehlboeck, J. S., & Simmons, A. (2013). Regional magnetic resonance imaging measures for multivariate analysis in Alzheimer's disease and mild cognitive impairment. *Brain Topography*, 26(1), 9–23.
- Wiederholt, J. L., & Bryant, B. R. (2001). *Gray oral reading tests (GORT-4)*. Austin, TX: Pro-Ed.
- Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., ... Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *NeuroImage*, 53(3), 1135–1146.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.
- Zhang, M., Chen, C., Xue, G., Lu, Z. L., Mei, L., Xue, H., ... Dong, Q. (2014). Language-general and -specific white matter microstructural bases for reading. *NeuroImage*, 98, 435–441.