

Inter-subject correlation during long narratives reveals widespread neural correlates of reading ability

David C. Jangraw ^{a,b,c,*}, Emily S. Finn ^{a,d}, Peter A. Bandettini ^{a,e}, Nicole Landi ^f, Haorui Sun ^c, Fumiko Hoeft ^{f,g}, Gang Chen ^h, Kenneth R. Pugh ^{f,i}, Peter J. Molfese ^{e,f}

^a Section on Functional Imaging Methods, NIMH, Bethesda, MD, United States

^b Emotion and Development Branch, NIMH, Bethesda, MD, United States

^c Department of Electrical and Biomedical Engineering, University of Vermont, Burlington, VT, United States

^d Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, United States

^e Center for Multimodal Neuroimaging, NIMH, Bethesda, MD, United States

^f Haskins Laboratories, New Haven, CT, United States

^g Department of Psychological Sciences, University of Connecticut, Hartford, CT, United States

^h Statistical Computing Core, NIMH, Bethesda, MD, United States

ⁱ Department of Pediatrics, Yale University School of Medicine, New Haven, CT, United States

ARTICLE INFO

Keywords:

fMRI

Naturalistic

Reading

Narrative

Inter-subject correlation

ABSTRACT

Recent work using fMRI inter-subject correlation analysis has provided new information about the brain's response to video and audio narratives, particularly in frontal regions not typically activated by single words. This approach is very well suited to the study of reading, where narrative is central to natural experience. But since past reading paradigms have primarily presented single words or phrases, the influence of narrative on semantic processing in the brain – and how that influence might change with reading ability – remains largely unexplored. In this study, we presented coherent stories to adolescents and young adults with a wide range of reading abilities. The stories were presented in alternating visual and auditory blocks. We used a dimensional inter-subject correlation analysis to identify regions in which better and worse readers had varying levels of consistency with other readers. This analysis identified a widespread set of brain regions in which activity timecourses were more similar among better readers than among worse readers. These differences were not detected with standard block activation analyses. Worse readers had higher correlation with better readers than with other worse readers, suggesting that the worse readers had "idiosyncratic" responses rather than using a single compensatory mechanism. Close inspection confirmed that these differences were not explained by differences in IQ or motion. These results suggest an expansion of the current view of where and how reading ability is reflected in the brain, and in doing so, they establish inter-subject correlation as a sensitive tool for future studies of reading disorders.

Abbreviations: AG, angular gyrus; ACC, anterior cingulate cortex; AnnaK, Anna Karenina; BOLD, blood oxygen level dependent; CalG, calcarine gyrus; CerVer, cerebellar vermis; Cun, cuneus; DLPFC, dorsolateral prefrontal cortex; EPI, echo-planar image; FDR, false discovery rate; FEF, frontal eye fields; fMRI, functional magnetic resonance imaging; Fus, fusiform; FWHM, full-width half maximum; GLM, general linear model; IFG, inferior frontal gyrus; IFG-pOper-pTri, inferior frontal gyrus pars opercularis and pars triangularis IFG-pOrb, inferior frontal gyrus pars orbitalis; IPL, intraparietal lobule; IOG, inferior occipital gyrus; IQ, intelligence quotient; ISC, inter-subject correlation; IS-RSA, inter-subject representational similarity analysis; LingG, lingual gyrus; LME, linear mixed-effects; MidOG, middle occipital gyrus; MidTG, middle temporal gyrus; MidCC, middle cingulate cortex; MidFG, middle frontal gyrus; NN, nearest neighbor; Oc-T, occipito-temporal; PCA, principal components analysis; pCun, posterior cuneus; PreCG, precentral gyrus; Precun, precuneus; ROI, region of interest; SMA, supplementary motor area; SMG, supramarginal gyrus; SPL, superior parietal lobule; STG, superior temporal gyrus; SOG, superior occipital gyrus; TMP, temporal-parietal; TOWRE, Test of Word Reading Efficiency; TP, temporal pole; TR, repetition time; VWFA, visual word form area; WASI, Wechsler Abbreviated Scale of Intelligence; WJ3, Woodcock-Johnson 3.

* Corresponding author at: Department of Electrical and Biomedical Engineering, University of Vermont, 33 Colchester Ave., Burlington, VT 05405, United States.

E-mail address: djangraw@uvm.edu (D.C. Jangraw).

<https://doi.org/10.1016/j.neuroimage.2023.120390>

Received 9 March 2023; Received in revised form 19 September 2023; Accepted 23 September 2023

Available online 24 September 2023

1053-8119/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Measuring narrative-related brain activity

In recent years, neuroimaging researchers have begun to gravitate toward “naturalistic” paradigms that present complex, dynamic stimuli such as films and audio narratives (Sonkusare et al., 2019; Finn et al., 2020). After decades of neuroimaging experiments using brief, simple stimuli to understand the brain’s processing of presumed basic elements and cognitive constructs, this trend toward naturalism represents a transformative new phase. This phase offers superior ecological validity and a chance to reexamine past assumptions about the most impactful elements of cognition. Thanks to these new studies, emerging evidence suggests that the brain may be more strongly tuned to naturalistic experiences (Schultz and Pilz, 2009; Gollo et al., 2015).

Inter-subject correlation (ISC) analysis has emerged as a powerful approach to analyzing naturalistic data, as it provides several advantages over traditional model-based linear regression techniques. ISC analysis measures task-evoked brain activity over the course of a narrative by taking one subject’s signal as a model for a second subject’s signal at the same spatial location (Hasson et al., 2004). Specifically, for a given voxel, ISC is defined as the temporal correlation between two subjects’ activation timecourses in that voxel (or, in some studies, the correlation between one subject’s timecourse and the average of the remaining n-1 subjects). Given an identical, time-locked stimulus across subjects, correlated activity in a given voxel implies that region is somehow involved in processing the stimulus. Greater ISC implies that responses are more “synchronized” or “consistent” across individuals. The advantages of ISC are that it is data-driven and relatively “model-free” in that it does not require specifying a model of either the task structure or the hemodynamic response, making it potentially more sensitive to nuanced task-evoked activation patterns (Finn et al., 2020).

These naturalistic methods have highlighted the substantial influence of narrative on the brain. Studies using ISC have shown that audio or visual narratives induce consistent responses in brain areas rarely studied in traditional tasks, especially frontal regions with longer “temporal receptive windows” (Lerner et al., 2011). ISC has been successfully leveraged with narratives to study group and individual differences, including autism (Hasson et al., 2009), depression (Guo et al., 2015) and paranoia (Finn et al., 2018). Recently, comparisons of between-subject patterns of ISC and patterns of behavioral similarity (such as mental disorder severity), an approach called inter-subject representational similarity analysis (IS-RSA), has extended ISC to the study of dimensional individual differences (Finn et al., 2020; Chen et al., 2017). These approaches share the advantages of single-group ISC and can be sensitive to nuanced differences in how groups or individuals process narrative stimuli over time-scales of several minutes.

1.2. The “Reading circuit” without narratives

Given these advantages, it is surprising that the study of reading has not yet benefited from such a naturalistic approach. Past studies have often used single words or sentences that are topically unrelated to each other, treating each trial as a brief, independent reading event. Researchers then use combinations of block, evoked, and mixed presentation patterns combined with traditional general linear model (GLM) analyses to identify brain regions modulated by reading, or by a person’s reading ability. This stimulus stands in contrast to how most people read in real-world settings, and this analysis misses the influence of any longer narrative arcs. The studies listed above suggest that this influence may be substantial and widespread in the brain.

Naturalism’s chance to advance reading research is a potentially impactful one. Reading is a fundamental skill that can have a major impact on an individual’s success both during and after childhood (Pugh et al., 2001). Although some reading difficulties are related to general intelligence, reading-specific deficits decouple from IQ (Ferrer

et al., 1999). Specific reading disability, often referred to as developmental dyslexia, is a common and chronic condition that can cause stress in children and adults, leading to increased risk of depression and anxiety disorders (Goldston et al., 2007). Although progress has been made in the assessment and education of children with dyslexia, most treatments produce only modest effects (Toffalini et al., 2021). But there is nevertheless reason for hope: an advancing understanding of the brain’s “reading circuit” brings an objective reference point to discussions of how reading advances through development, and this may eventually inspire brain-based interventions for struggling readers (Kearns et al., 2018). Because this circuit was identified using non-narrative stimuli, naturalistic stimuli could reveal additional reading-related regions which, by extension, may have implications for intervention.

The canonical reading circuit consists of three major hubs: (1) the inferior frontal gyrus (IFG); (2) a temporal-parietal area (TMP) composed of the posterior superior temporal gyrus, middle temporal gyrus, angular gyrus, and supramarginal gyrus; and (3) the occipital-temporal area (Oc-T), including the inferior temporal gyrus and fusiform gyrus (which overlaps with the often-cited visual word form area (VWFA)) (Pugh et al., 2001; Dehaene et al., 2002; Price and Devlin, 2011; Dehaene and Cohen, 2011; Yeatman et al., 2021). These three areas are believed to be responsible for phonological recoding, orthographic decoding of words, and fast automated word recognition/pattern matching, respectively (Pugh et al., 2010).

When processing non-narrative stimuli, individual differences in reading ability and development have been reflected in this circuit. These (usually) left-hemisphere regions are thought to develop sequentially, with early readers relying heavily on the “dorsal” stream, of the IFG and TMP. As reading becomes more proficient, processing of reading shifts to rely more on the ventral stream of the Oc-T (Pugh et al., 2010; Schlaggar and McCandliss, 2007). Activation of these regions has been found to correlate with behavioral measures of reading ability (for reviews, see Price 2012, Pugh et al. 2013, Norton et al. 2015, Ozenov-Palchik et al. 2016, Vandermosten et al. 2016).

1.3. Adding narratives to the study of reading ability

Because most past experimental designs have lacked an extended narrative, naturalistic reading paradigms and ISC analyses offer a new opportunity to expand our understanding of the reading circuit. Their increased ecological validity offers a better view of real-life reading experience and brain activity. They also allow researchers to study how information representations at higher levels of the semantic hierarchy – including regions that respond to longer-term context – might be altered by reading ability.

In this study, we brought the strengths of naturalism and ISC to the study of reading ability. We presented long visual and auditory stories during fMRI scans to 75 adolescents and young adults with varying reading ability based on commonly used reading assessments. We assessed the consistency of various brain regions’ fMRI response timecourses across participants (Nastase et al., 2019; Finn et al., 2020), and we quantified how this inter-individual consistency depended on reading ability. We tested the hypothesis that better and worse readers would show different levels of consistency in regions across the brain, including some not typically implicated in reading research. We also used IS-RSA to test the hypothesis that worse readers would have decreased consistency in these areas not because they produced less brain activity during reading, but because this activity tended to be more idiosyncratic. We also investigated the timecourses of BOLD activity and sliding-window ISC in these clusters to explore the possible stimulus drivers of these effects.

2. Materials & methods

2.1. Subjects and inclusion criteria

Our analyses made secondary use of a dataset collected as part of a study of how spoken and written word learning are altered in reading disability. This larger data collection effort included multiple imaging tasks, behavioral tasks, and surveys collected over multiple days. Other outcomes from this dataset have been the subject of a previous publication (Ryherd et al., 2018). We will make use of data from the “passage task” described in Section 2.3, which we have repurposed for our analysis of individual differences in inter-subject correlation because it involved the presentation of extended, time-locked narratives to adolescents with well-characterized reading ability levels. (While the paradigm’s use of auditory blocks interspersed with visual (i.e., written) ones introduces unnatural modality changes, it did provide a benefit by enabling contrasts between auditory and visual blocks.)

Only data from subjects who completed the passage task and full complement of reading assessments described below were included in the analysis. Using these criteria, data from 75 monolingual adolescent and young adult English speakers were included in the present study. In keeping with the policies of the Yale University Human Investigation Committee, all participants over 18 gave informed consent. For participants under 18, a parent or guardian provided written consent, while children were asked to give verbal assent.

2.2. Behavioral measures

Participants completed a battery of tests to evaluate their reading ability. These tests included the Woodcock-Johnson 3 (WJ3; Woodcock et al., 2001) Letter-Word ID and Word Attack subtests; as well as the Test of Word Reading Efficiency (TOWRE; Torgesen et al., 1999) Sight Word Efficiency and Phonetic Decoding Efficiency subtests. These tests were later combined to create a composite “reading score” that combines all measures of their word and sub-word level reading ability, as we will describe.

Raw scores were normed for age (i.e., transformed to standard scores), then normalized to mean 0 and standard deviation 1. Similar to Pugh et al. (2013), we computed a principal components analysis (PCA) on these reading measures, where the first principal component across scores was used as the composite “reading score” for each participant. This component showed relatively uniform positive weightings of the four sub-test scores and was conceived as a measure of general reading

ability that we believe is more robust to measurement error than any one subtest score (Fig. 1).

The Wechsler Abbreviated Scale of Intelligence II (WASI; Wechsler, 2011) was used to measure performance IQ. Because handedness is related to language lateralization (Szaflarski et al., 2002), subjects also completed an Edinburgh Handedness Inventory to evaluate the extent to which their left or right hand was dominant (Oldfield et al., 1971). However, we opted not to exclude any participants based on handedness to increase the generalization and representativeness of our results (Bailey et al., 2020).

2.3. MRI data collection

fMRI data were collected using a Siemens TIM-Trio 3 Tesla MRI system with a 12-channel head coil. T1-weighted MPRAGE structural images were acquired with the following parameters: matrix size = 256 × 256; voxel size = 1 × 1 × 1 mm; FoV = 256 mm; TR = 2530 ms; TE = 3.66 ms; flip angle = 7°. These were followed by gradient echo-planar image (EPI) functional scans with the following parameters: 32 slices; 4 mm slice thickness; no gap, matrix size = 64 × 64; voxel size = 3.4375 × 3.4375 × 4 mm; FoV = 220 mm; TR = 2000 ms; TE = 30 ms; flip angle = 80°.

During functional runs, participants completed the “passage task” described in (Ryherd et al., 2018), as seen in Fig. 2. In this task, they passively read and listened to two stories by Hans Christian Andersen across two scan runs. Relatively little-known stories (“Jack the Dullard” and “Brave Tin Soldier”) were selected, so participants were unlikely to have prior knowledge of them. Each run was organized in visual and auditory blocks (with the narrative continuing across blocks). In the visual blocks, the story text was presented phrase by phrase at the center of the screen. Each phrase was presented for 2 s, with an average of 6.46 words per phrase. In the auditory blocks, participants listened to a narrator read the story through headphones. The speech in the auditory blocks was a continuous stream whose speed roughly matched the visual block’s presentation rate (194 words per minute). Thus, both visual and auditory presentation rates were controlled by the experiment and similar to each other. Each run consisted of six blocks, as seen in Fig. 2. The block order in the first run was: baseline, visual, auditory, visual, auditory, baseline. The block order in the second run was: baseline, auditory, visual, auditory, visual, baseline. Stimuli and blocks were not randomized: they were presented the same way for every subject to make ISC analysis possible. Blocks had a mean duration of 59 s (range 48–64), such that each story lasted about 4 min. (We still consider this

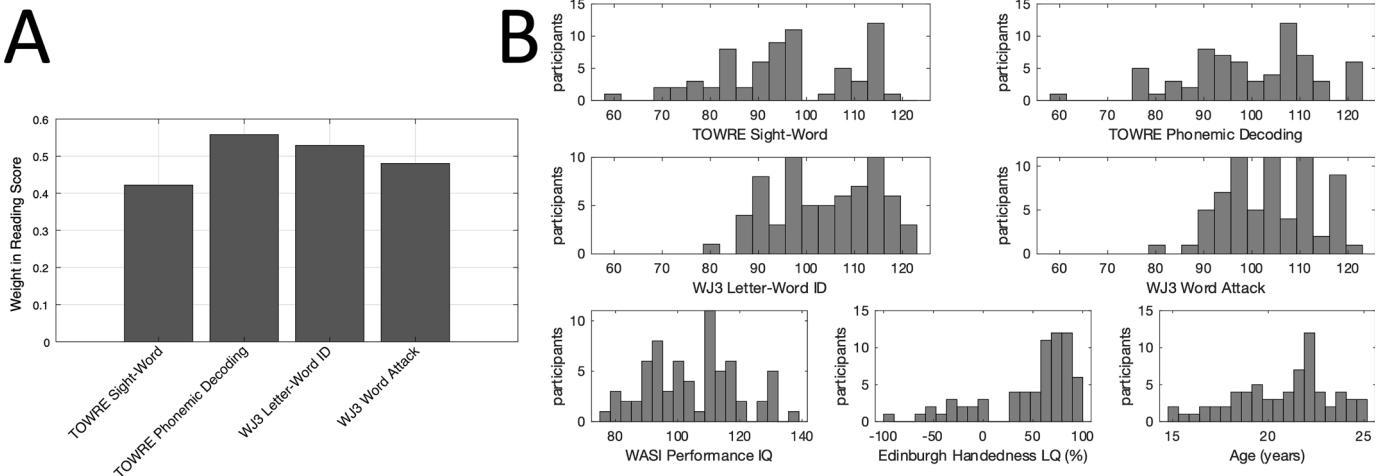


Fig. 1. Distribution of behavioral scores. A: Weights across reading tests used to produce the composite “reading score”. After normalizing each score to mean zero and standard deviation one across subjects, support vector decomposition was used to find the first principal component of the scores. The composite reading score is produced by multiplying each normalized score by its weight and taking the sum. Standard scores, corrected for age, were used. B: Histograms showing the distributions of each behavioral score within the full cohort ($n = 68$).

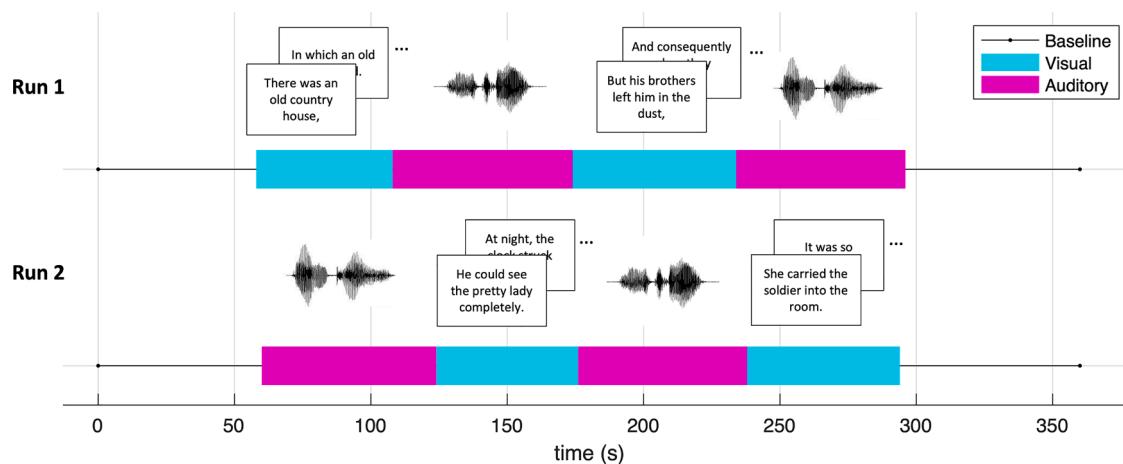


Fig. 2. Task illustration. Two stories were presented in blocks that alternated visual (each phrase presented for 2 s) and auditory (a narrator reading) blocks. All participants saw the same stories in the same order to allow for inter-subject correlation analysis. Each story was preceded and followed by a baseline block with a blank screen. Each block was 48–74 s long according to the length of the passage or to keep the scan duration consistent.

duration “naturalistic” because it occurs frequently in real life and, in past studies, the vast majority of brain regions that responded consistently to a 7 min audio or silent-film narrative retained that consistency when the narrative was scrambled into “paragraphs” of mean 38 s (Lerner et al., 2011; Hasson et al., 2008). Participants were asked to pay attention to the story so that they could answer comprehension questions afterwards. No responses were required of the subject during the task. After each story (i.e., run), participants were asked a comprehension question to ensure that they were paying attention to the task. Unfortunately, the responses to these questions were not recorded.

2.4. MRI preprocessing

Processing was carried out with custom scripts using AFNI (Cox, 1996) and MATLAB (The MathWorks Inc., Natick, MA). We also used Freesurfer’s standard pipeline (recon-all) to produce segmentation maps of the ventricles and white matter (Fischl et al., 2012). Using AFNI, the first 6 volumes in each run were discarded to allow the scanner to reach steady state. Despiking, slice-time correction, motion correction and nonlinear registration to the MNI 152 nonlinear asymmetric atlas (MNI152_T1_2009c), blurring at 6 mm full-width half maximum (FWHM), masking to in-brain voxels, and rescaling to % BOLD signal change were also performed in AFNI. TRs with excessive motion (>0.3 mm) or outliers ($>10\%$ of voxels) were censored. Censored time points were set to zero rather than removed from analyses. This choice was important to preserve the temporal structure across participants in the ISC analyses. Finally, we regressed out nuisance parameters including motion, motion derivatives, and the first three principal components of activity found in the ventricles. We also used AFNI’s ANATCOR capability (Jo et al., 2010) to produce a local white-matter-based noise regressor for each voxel based on the activity of white matter voxels weighted by a three-dimensional gaussian with 30 mm FWHM.

2.5. Exclusion criteria and cohort description

Two subjects were inadvertently presented with the second story first. These two subjects’ runs were switched to align their timecourses with those of the other subjects. One subject was inadvertently presented with the same story twice. Data from this subject was removed from further analysis. Data from six subjects with excessive motion or outliers ($>15\%$ TRs censored) were also excluded from further analysis.

The final cohort of 68 subjects had a mean \pm std age of 20.8 ± 2.5 age (range 14–25). 57% were male, and 7.4% were left-handed (Edinburgh handedness inventory laterality quotient $< -40\%$). They had a mean \pm std performance IQ of 105 ± 14.9 (range 75–139). Because

studies suggest that the neural basis of phonological deficits is independent of IQ (Tanaka et al., 2011; Simos et al., 2014) and response to intervention does not depend on IQ (Morris et al., 2012; Stuebing et al., 2009), we chose not to exclude any participants on the basis of IQ. This cohort had mean age-corrected scores on the reading subtests described above that were between 95 and 105, with standard deviations between 9 and 14. The minimum and maximum subtest scores for individuals were 58 and 123. Four participants had received a clinical dyslexia diagnosis at the time of participation. Distributions of reading subtests and demographic scores are displayed in Fig. 1 and summarized in Table 1.

2.6. MRI analyses

For standard block analyses, we added boxcar regressors of amplitude 1 spanning the duration of the visual blocks and auditory blocks to the general linear model (GLM) in the regression step described above. For each participant, we then calculated contrasts for auditory and visual blocks. The group-level results were modeled as a weighted combination of each participant’s z-scored reading score, motion (quantified as mean framewise displacement), and IQ (WASI performance IQ), with stimulus modality (auditory or visual) included as a within-subject factor. The weights for each factor were estimated and statistically evaluated using AFNI’s 3dMVM function, which uses multivariate modeling (MVM) to include quantitative covariates (like reading score) as well as within-subject factors (like stimulus modality) (Chen et al., 2014).

ISC analysis was used to find the correlation between voxel-wise timecourses in each distinct pair of subjects (Supplementary Fig. S1). We used the AFNI script 3dTCorrelate to calculate the Pearson correlation of each subject’s timecourse in a voxel with the same voxel’s

Table 1
Reading test subscores and demographics in the unmatched sample ($n = 68$).

Measure	M (SD)	Range
TOWRE Sight-Word	95.4 (13.5)	61–119
TOWRE Phonemic Decoding	99.6 (13.4)	58–120
WJ3 Letter-Word ID	104 (10.7)	80–123
WJ3 Word Attack	104 (9.34)	79–120
WASI Performance IQ	105 (14.9)	75–139
Age (years)	20.8 (2.54)	14.8–25.2
Demographic	N (%)	
Male	39 (57.4%)	
Left-Handed	5 (7.4%)	

timecourses in every other subject. Single-group and individual-difference measures were calculated and tested for significance using the AFNI program 3dISC, which fits a linear mixed-effects (LME) model with a crossed random-effects formulation to accurately account for the correlation structure embedded in the ISC data (Chen et al., 2017).

We modeled each participant pair's ISC as a weighted combination of their combined reading score, IQ, and motion. To generate a combined score for each pair, we calculated the sum of the two individuals' z-scored reading scores, IQ, and motion. For consistency with block analyses, ISC analyses were not restricted to task blocks: all TRs were used except during modality-specific (e.g., visual or auditory) analyses.

Statistical maps were thresholded at $p < 0.002$ and separated into contiguous clusters (the stringent threshold of 0.002 was used rather than the more common 0.05 or 0.01 to guard against false positives in cluster correction, as in Finn et al. (2018)). Cluster thresholds to control for family-wise error were calculated using AFNI's 3dClustSim command with a mixed-model auto-correlation function. This function found that clusters of >38 voxels could be considered significant at a level of $\alpha < 0.05$.

In post-hoc tests to quantify the ISC induced by each modality separately, we cropped the preprocessed data to include only the time-points collected during each block type. To avoid the transients induced by block onsets, we excluded the first 5 TRs of each block, and we did not include TRs after the block ended. The first 5 TRs of each block (except for the first block) were classified as transition periods. All the blocks for a given modality were concatenated before the ISC was calculated. This approach resulted in 178 s (89 TRs) of visual blocks, 214 s (107 TRs) of auditory blocks, and 100 s (50 TRs) of transition periods.

2.7. IQ control analyses

In considering our results, we investigated participant motion and IQ as potential confounds. Motion and IQ both correlated significantly with reading score (Supplementary Fig. S2). We therefore performed additional analyses to characterize their influence on our results, first visualizing the voxelwise impact of motion or IQ instead of reading score on ISC on the LME model.

To further disentangle IQ and reading ability, we performed a control analysis with a matched subset of subjects. To find this subset, we noted that each reading-IQ group pairing (e.g., reading score above median, IQ above median) had at least 10 subjects. We therefore randomly chose 10,000 different subsets of 40 subjects with 10 in each group pairing. We then selected the subset with the lowest correlation between reading ability and IQ ($r = 0.122$, $p = 0.452$). Distributions of reading subtests and demographic scores in this new subset are summarized in Supplementary Table S1.

Next, we performed the same ISC analyses (i.e., an LME model including reading score, IQ, and motion) on this matched cohort that we had on the original cohort. Cluster thresholds were calculated in the same way on this smaller cohort, and they again found a cluster threshold of 38 voxels corresponding to $\alpha < 0.05$.

2.8. Cluster ISC analyses

We next investigated the pairwise ISC in clusters that showed a significant impact of reading score on ISC in the full cohort. Where we found a single large cluster (>6000 voxels) overlapping with many known language regions, we examined three ROIs where this cluster overlapped with language regions derived from AFNI's "CA_ML_18_MNI" atlas, which contains Eickhoff-Zilles macro labels from N27 in MNI space. The regions were "Left Superior Temporal Gyrus (code 81)", "between Left Fusiform Gyrus (code 55)", and "Left Angular Gyrus (code 65)".

To calculate the cluster-wise ISC values for each pair of subjects, we performed a Fisher Z transform on the correlation coefficient between that pair of subjects in each voxel. We then averaged these values across

all voxels in the region of interest (ROI). Finally, we performed an inverse Fisher Z transform to convert back to an r value.

To statistically assess whether "worse" readers (defined as those with reading scores at or below the median) had higher ISC with "better" readers (defined as those with reading scores above the median) or with other worse readers, we converted Pearson correlation r values to z values (using a Fisher Z transform) and calculated the mean ISC between pairs of better readers, pairs of worse readers, and "mixed" reader pairs (one better reader, one worse reader). We then randomly permuted the participants' group ("better" or "worse") labels 10,000 times to calculate a null distribution of z values for better, worse, and mixed reader pairs. The differences between (a) better and mixed reader pairs, (b) mixed and worse reader pairs, and (c) better and worse reader pairs were considered significant if they were larger than that from over 95% of permutations in almost all of the ROIs (i.e., an uncorrected $p < 0.05$).

Because the statistical test just described is somewhat susceptible to selection bias (since clusters were selected based on better pairs' ISC $>$ worse pairs' ISC), we also confirmed its trend using a whole-brain, dimensional analysis as described in the next section.

2.9. Representational similarity analysis

To test the hypothesis (described in Section 3.5) that worse readers had more "idiosyncratic" responses in ROIs across the whole brain, we performed a dimensional analysis, adopting the inter-subject representational similarity analysis (IS-RSA) approach described by Finn et al. (2020). In this approach, the behavioral scores of each subject were used to define a "behavioral similarity" matrix that predicts how synchronized each pair's neural responses will be based solely on their behavioral scores. We then calculated a Spearman rank correlation between the unique (i.e., upper triangular) elements of that matrix and the same values in the ISC matrix for a given ROI. If that correlation is high, it implies that behavioral model was a good predictor for the observed ISC in that region. To investigate these findings in ROIs across the brain without preliminary analyses biasing our results, we used all ROIs in the Shen Atlas, a 268-region parcellation of the whole brain (including cerebellum and brainstem) (Shen et al., 2013).

To generate a null distribution for significance testing, we permuted the rows and columns of the behavioral similarity matrix 10,000 times and recomputed the RSA correlation value for each ROI. This builds up a distribution of correlation coefficients where the behavioral similarity score's structure is the same, but it is based on meaningless values rather than the participant's real reading scores. The p value was calculated as the fraction of correlation coefficients in the null distribution that exceeded the true correlation coefficient value. These p values were calculated for all ROIs, then corrected for multiple comparisons using false discovery rate (FDR) correction.

We investigated two competing models for the behavioral similarity matrix. First, the "Anna Karenina" (AnnaK) model predicts that worse readers will have idiosyncratic responses (mirroring the statement in Leo Tolstoy's novel Anna Karenina that "every unhappy family is unhappy in its own way") (Finn et al., 2020). In this model, the predicted synchronization is quantified as the mean reading score of the two participants. (While we hypothesized more consistency among better readers and more variability among worse readers, note that this model is equally powered to detect effects in the opposite direction—i.e., more consistency among worse readers and more idiosyncrasy among better readers—this would be indicated by negative [rather than positive] representational similarity r values, suggesting that as a pair's mean reading score decreases, their neural synchrony increases.) Second, the "Nearest Neighbor" (NN) model predicts that participants with similar reading scores will have similar responses, regardless of whether those scores are high or low. This model could be more consistent with the idea of compensatory mechanisms for struggling readers. In this model, the predicted synchronization is quantified as the opposite of the absolute value of the difference between the two participants' reading

scores (i.e., $-\text{abs}(\text{score}_A - \text{score}_B)$). The predicted synchronization matrices of each model can be seen in Supplementary Fig. S3 (top).

2.10. Time-resolved BOLD and ISC analyses

We next investigated the timecourse of BOLD activity in regions that showed significant impacts of reading score on ISC in the full cohort ($n = 68$). We used the timecourse with nuisance variables – but not task blocks – regressed out (i.e., the same timecourse used for ISC analyses). For each subject, we calculated the mean timecourse of activity across all voxels in an ROI. We then calculated the mean and standard error across all subjects in a group. We used a median split of reading score to group participants into “better” and “worse” readers for plotting.

Finally, we examined the timecourse of ISC in these same regions using a sliding-window approach. To do this, we calculated the Pearson correlation coefficient between (for a given voxel and pair of subjects) samples in a 15-sample (30 s) window. We slid the window forward one sample and repeated the process until we reached the end. Each time, we performed a Fisher Z transform on the correlation coefficient, averaged the values across all voxels in the ROI and all participants in a group, and performed an inverse Fisher Z transform to convert back to an r value for plotting.

3. Results

3.1. Single-group analysis

Standard GLM block analysis (any task vs. baseline) showed that narrative stimuli induced activation in regions that closely resemble the canonical “reading circuit” (Supplementary Fig. S4A; all whole-brain fMRI analyses used $p < 0.002$ and $\alpha < 0.05$). Inter-subject correlation analysis of the whole session revealed even more regions where the task induced consistent activation patterns (Supplementary Fig. S4B). In fact, nearly all of the brain exhibited some degree of significant positive ISC. Note that ISC is positive even in regions where the task causes deactivation: synchronized deactivations will still present as a positive correlation across subjects. The greatest magnitude of ISC was found in visual and auditory cortices, which is not surprising given the large activations seen in the block analysis of these areas.

3.2. Impact of modality

Visual and auditory blocks evoked different patterns of activation and ISC in some regions and similar patterns in others. In our standard block analysis, early sensory cortices showed the expected modality preferences (Supplementary Fig. S5A, B). The stimulus modality also had an effect on ISC in some areas (Supplementary Fig. S5C, D). Medial occipital cortex had greater ISC during visual blocks, matching block activation patterns. More surprisingly, parts of the middle occipital gyrus had greater ISC during auditory blocks, and superior temporal cortex had greater ISC during visual blocks. But most of the parietal cortex, middle temporal cortex, and many frontal regions had significant ISC in both block types that was not significantly different between blocks. The smaller extent of significant ISC in this modality-specific analysis is not surprising: by focusing on one modality, this analysis uses roughly 1/3 of the data for each subject.

3.3. Dimensional reading ability analysis

In identifying an impact of individual differences in reading ability on responses to the extended narrative, ISC had sensitivity superior to the block analysis. Using standard block analysis, no parts of the brain exhibited a significant effect of reading score. But ISC analysis showed that higher reading scores were associated with significantly more consistent responses in a widespread set of brain areas, including areas of overlap with reading-circuit areas left angular gyrus (LAG), left

fusiform (lFus), and left inferior frontal gyrus (lIFG) (Fig. 3, Table 2). The right-hemisphere analogs of each of these regions except IFG also showed a significantly positive impact of reading ability on ISC, though the effect did not appear to be as strong on the right.

Other regions showing an effect of reading ability spanned a wide range of functions, most falling within a large contiguous cluster. Auditory and semantic regions included bilateral superior temporal gyrus (STG), temporal pole (TP), and middle temporal gyrus (MidTG). Visuospatial regions included bilateral calcarine gyrus (CalG), middle occipital gyrus (MidOG), superior occipital gyrus (SOG), precuneus (Precun), and intraparietal area. Subcortical regions included the thalamus and hippocampus. Frontal regions included bilateral frontal eye fields (FEF), anterior cingulate cortex (ACC) and middle cingulate cortex (MidCC), and left middle frontal gyrus (lMidFG). Cerebellar regions included the vermis, though our field of view excluded the inferior half of the cerebellum. This wide range of regions suggests that reading ability and/or its correlates intersect with a wide array of brain functions.

ISC model effect sizes can be interpreted as follows. Because we used z-scored behavioral scores in the model, a value of 0.05 on a surface plot in Fig. 3 means that increasing the reading score of either participant by 1 standard deviation tended to increase the ISC (i.e., the Fisher-transformed correlation coefficient) by $z = 0.05$. At the low correlation values observed in our ISC data (Supplementary Fig. S4B), the correlation coefficient r and its Fisher transform z are nearly equivalent. For the mean voxel with a significant effect of reading score, the average pair’s correlation coefficient was $r = 0.077$ and the reading score effect size was $r = 0.0140$. This means that increasing the reading score of either participant by 1 standard deviation tended to increase the ISC in this mean voxel from $r = 0.077$ to $r = 0.091$ (an 18% increase).

3.4. Motion and IQ analyses

To test whether motion might be driving the results, we calculated the mean framewise displacement for each participant and examined the relationship between our composite reading score and participant motion (Supplementary Fig. S2A). Results showed that the two were significantly anticorrelated (Pearson correlation, $r^2 = 0.0827$, $p = 0.0174$). We examined the impact of motion on ISC by examining the impact of motion instead of reading score in the same LME model (Supplementary Fig. S6A). Results showed that 37.7% of voxels in the significant reading score clusters were also present in the significant motion clusters. In all regions except lMidOG, motion exhibited a negative relationship with ISC. Given motion’s anticorrelation with reading score, this suggests that motion’s inclusion in the model reduced the ISC attributed to reading score.

To test the specificity of our findings to reading ability, we examined the relationship between our composite reading score and IQ (Supplementary Fig. S2B). IQ correlated significantly with reading score (Pearson correlation, $r^2 = 0.316$, $p = 6.13 \times 10^{-7}$). We examined the impact of IQ on ISC by examining the impact of IQ instead of reading score in the same LME model (Supplementary Fig. S6B). Results showed that 21.6% of voxels in the significant reading-score clusters were also present in the significant IQ clusters. In some regions (including lMidOG and lSTG), IQ exhibited a negative relationship with ISC. Given IQ’s positive correlation with reading score, this presented a plausible risk that IQ was inflating the ISC attributable to reading score in the model.

We therefore examined the group differences between a subset of 40 participants in which the reading and IQ median-split “groups” (e.g., reading score above the median, IQ above the median) were each matched for the other factor (Supplementary Fig. S7; see Section 2.7 for a description). As before, we used dimensional LME modeling to examine the individual differences in ISC attributable to each factor. These results are shown in Supplementary Fig. S8 (with the original 68-subject reading score clusters superimposed) and Supplementary Table S2. Even with this stringent IQ correction, the model found that

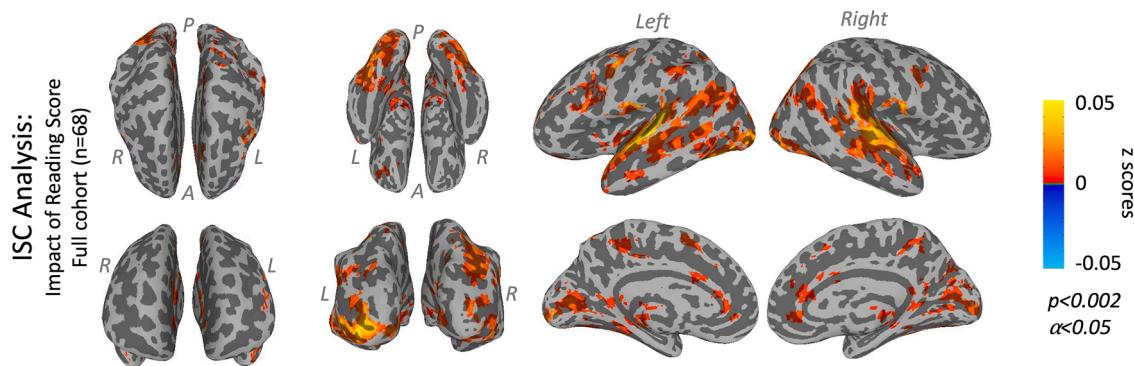


Fig. 3. LME ISC results. Clusters showing a significant impact of reading score on ISC in the full cohort ($n = 68$). Block GLM analyses (not shown) showed no significant group differences. In each sub-figure, results are projected onto the surface of an inflated brain as seen from 8 different viewpoints. Light gray represents gyri and dark gray represents sulci. The views (listed from left to right, top to bottom) are top, bottom, lateral left, lateral right, front, back, medial left, medial right.

Table 2

Clusters showing a significant impact of reading score on ISC in the LME modeling analysis of the full cohort of participants ($n = 68$). Volume is in voxels; peak coordinates are RAI.

Cluster	Label	Volume	Peak (x,y,z)
1	Language	6606	(67.5,-13.5,4.5)
2	lPreCG	204	(52.5,1.5,46.5)
3	lIFG-pOper-pTri	199	(52.5,-13.5,28.5)
4	ACC	171	(-7.5,-40.5,-4.5)
5	SMA	169	(10.5,-4.5,58.5)
6	lIPL-lSMG	155	(61.5,34.5,49.5)
7	Precun	140	(-10.5,40.5,49.5)
8	rPreCG	75	(-52.5,-4.5,37.5)
9	lMidFG	58	(46.5,-43.5,28.5)
10	lSPL-lIPL	54	(28.5,55.5,52.5)
11	lPrecun-lSPL	51	(10.5,61.5,61.5)
12	CerVer	46	(1.5,55.5,-37.5)
13	lIFG-pOrb	44	(34.5,-34.5,-10.5)

reading ability was associated with higher ISC in widespread regions which, as in the full cohort, included both canonical reading regions and frontal regions. Overall, this analysis suggests that much of the original pattern of results—showing more consistent activation for better readers—stands even when controlling more strongly for IQ.

3.5. Cluster ISC analyses

We next plotted the pairwise ISC between each pair of subjects in the clusters where reading score had a significant impact on ISC in the full cohort ($n = 68$) to see whether the observed results could be attributed to an alternative response pattern. Because the largest cluster covered many language-related regions, we examined three ROIs (lSTG, lFus, and lAG) where this cluster overlapped with language regions derived from an atlas.

In general, our group of “worse readers” (those having reading scores below the median) had higher ISC with “better readers” (those above the median) readers than with other worse readers. This pattern is illustrated using representative ROIs in Fig. 4; the full set of clusters can be seen in Supplementary Fig. S9. This observation was assessed statistically using the permutation tests described in Section 2.8 and was significant in most of the ROIs (uncorrected $p < 0.05$). This suggests that in these regions, worse readers tend to have “idiosyncratic” response timecourses that are more similar to better readers than to other worse readers. (If the BOLD response at each time point were a dimension, one might visualize this as a cluster of better readers in the middle with a wider cloud of worse readers around it.) Of course, these regions were selected for having a strong impact of reading ability on ISC, so this pattern might plausibly have emerged by chance. We therefore tested

this theory across the whole brain using a dimensional approach.

3.6. Representational similarity analysis

To test the hypothesis of idiosyncratic responses against an alternative hypothesis of similar behavior suggesting similar neural responses, we implemented the IS-RSA approach described in Section 2.9 with two alternative models. First, the AnnaK model predicts that better readers will share a consistent response, while worse readers will have idiosyncratic responses. Second, the “Nearest Neighbor” (NN) model predicts that participants with similar reading scores will have similar responses, regardless of whether those scores are high or low. This model could be more consistent with the idea of compensatory mechanisms for worse readers.

Results showed that the AnnaK model produced significantly better fits across the brain than the NN model (Wilcoxon signed-rank test, $z = 12.9$, $p = 7.48 \times 10^{-38}$). 96 of the 268 regions in the atlas, spread widely across the brain, showed significant IS-RSA with the AnnaK model at an FDR-corrected $q < 0.05$ (Fig. 5). This lends support to the LME results and supports the interpretation that worse readers tend to have more idiosyncratic responses.

In follow-up analyses, we found that AnnaK models based on individual reading subscores also produced better fits than NN models. The WJ3 Letter-Word ID subscore and the TOWRE Phonetic Decoding subscore predicted ISC in regions closely matching those in Fig. 6. The WJ3 Word Attack subscore significantly predicted ISC in only a few regions near the canonical reading circuit, and the TOWRE sight-word subscore did not significantly predict ISC in any regions. Results of these analyses are shown in Supplementary Fig. S11.

3.7. Cluster timecourse analysis

We next turned our attention to the timecourse of ISC (in the clusters showing a significant impact of reading score on ISC in the full cohort), to explore whether results might be driven by auditory blocks, visual blocks, or transitions between blocks. The timecourses of activity in these clusters provide insight into why ISC detects consistent activity that block analysis misses (Fig. 6 and Supplementary Fig. S10, subplots labeled “BOLD”). In each region, there is consistent activity that is clearly not captured by the block structure. Much of this activity is consistent across the averages of better readers and worse readers, but it is distinct from each of the other regions.

Timecourses of sliding-window ISC (Fig. 6 and Supplementary Fig. S10, subplots labeled “30 s ISC”) were similarly distinct between brain regions. Several areas exhibited high ISC differences during the transitions between blocks. The lSTG showed its greatest ISC differences at the onset and offset of auditory blocks. The lFus, which contains the VWFA, showed greater ISC group differences at the onset and offset of

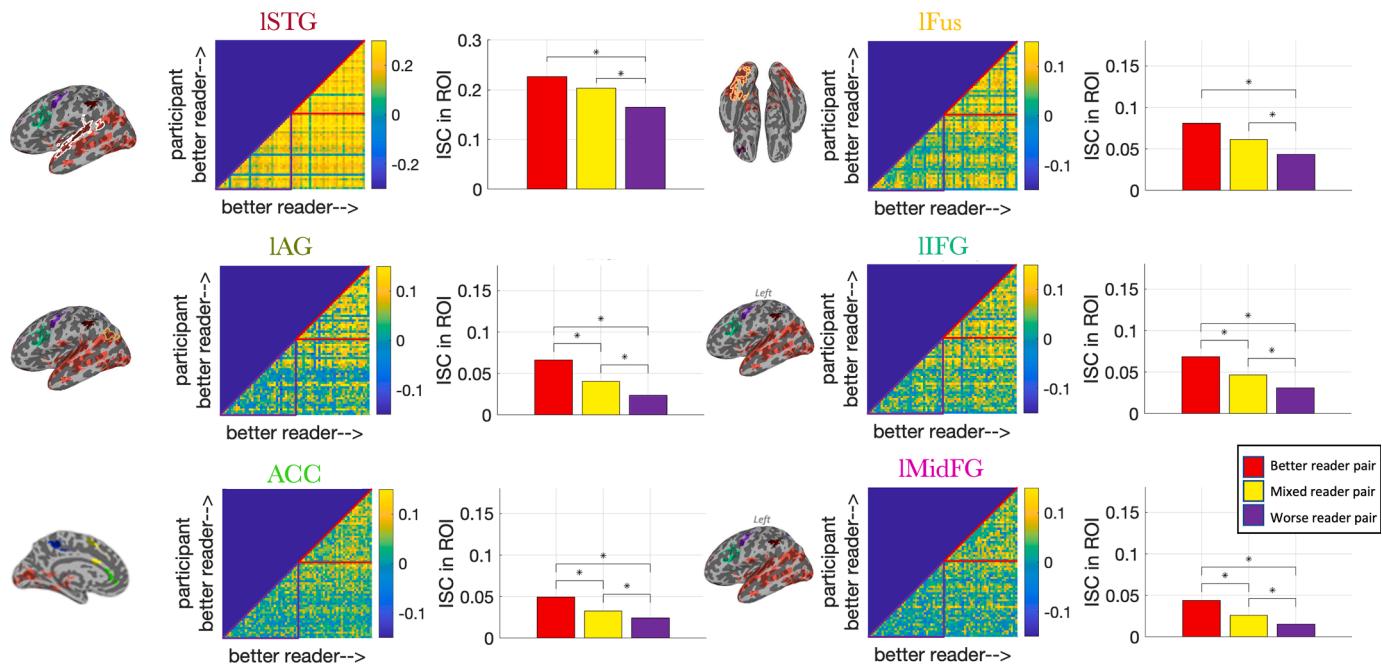


Fig. 4. Pairwise ISC in select clusters. A subset of the clusters showing a significant impact of reading score on ISC in our full cohort ($n = 68$) are shown here. Each cluster has a pair of plots with a title whose color matches the corresponding cluster seen on the brain to its left. Outlines indicate the largest cluster's overlap with a language region derived from an atlas. Middle: pairwise ISC (r value) between all subjects in these clusters. Each row and column is a participant, and the color in row i and column j represents the mean Pearson correlation between participants i and j across all the voxels in the ROI. Subjects are sorted according to their composite reading score, such that better readers are closer to the top and farther to the right. Colored triangles mark the median-split groupings into better reader pairs (red) and worse reader pairs (purple). Right: mean pairwise ISC among two better readers (“better”), one better reader and one worse reader (“mixed”), and two worse readers (“worse”). Stars indicate group differences were significant in permutation tests at $p < 0.05$ (uncorrected).

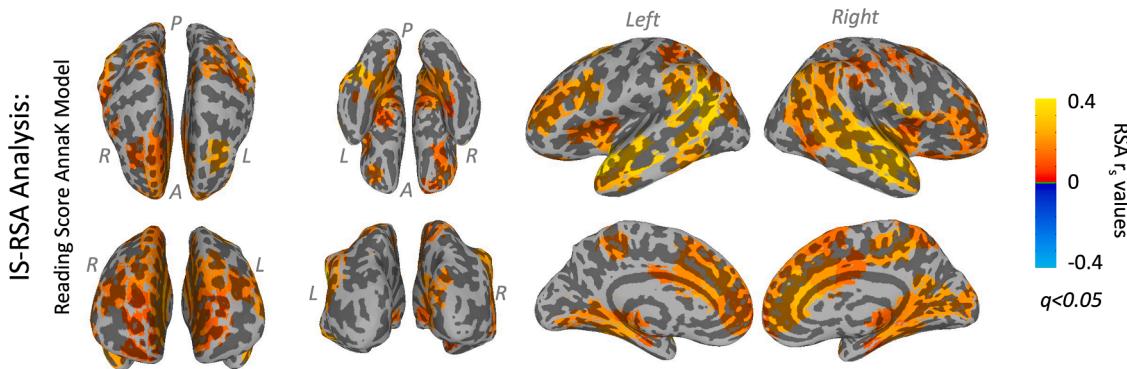


Fig. 5. Whole-brain IS-RSA analysis. IS-RSA Spearman r values from each ROI of the 268-node Shen atlas in which RSA reached a threshold of FDR-corrected $q < 0.05$, projected back onto the surface of the brain. The pairwise ISC matrix was compared with a behavioral similarity based on the “AnnaK” model, in which the predicted synchronization is based on the mean reading score across the pair.

visual blocks. The IAG showed ISC group differences within both block types. The IIFG showed high ISC differences at most transitions in and out of baseline. Other frontal regions, like the ACC and lMidFG clusters, showed patterns of ISC whose relationship to the stimulus blocks and BOLD responses were difficult to discern.

To quantitatively check whether the regions showing greater impacts of reading ability on ISC had a preference for one stimulus modality, we performed contrasts between the 1-group ISC during visual blocks vs. auditory blocks (excluding the 10 s of transition at the start of each block). Results are shown in Supplementary Fig. S5D. Both visual and auditory cortices showed significantly greater ISC in visual blocks than in auditory ones, but bilateral MidOG showed the opposite. Other regions whose ISC exhibited a stimulus type preference showed little overlap with the reading ability clusters. These results suggest that the regions in which ISC varied with reading ability did not generally have

greater ISC during visual stimuli (as one might expect for an ability defined by reading words on a page). This is consistent with a body of work which has demonstrated that reading skill is supported by oral language skills (e.g., vocabulary, phonological processing) and that there is significant overlap in reading skill and oral language skills (Hogan et al., 2014). It is also consistent with Ryherd et al. (2018), who, with this same task, found that reading skill-linked patterns of functional activation were similar for the visual and auditory blocks.

4. Discussion

In this study, we used ISC analysis to reveal that the temporal patterns of activation in many brain regions are significantly more consistent across subjects with better reading ability. These inter-subject differences were not found using more typical block-design analyses.

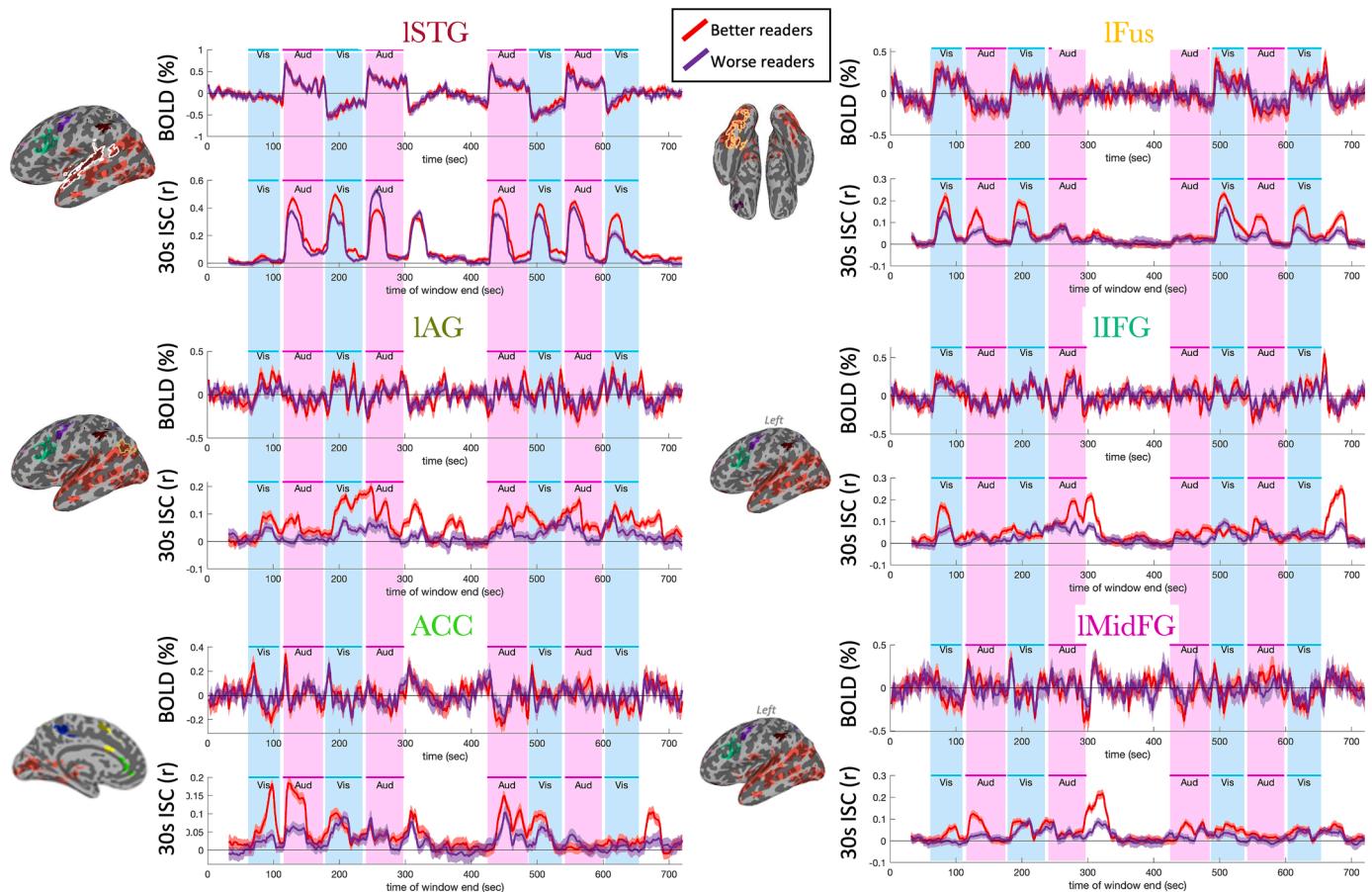


Fig. 6. Time-courses of activity and ISC in select clusters with significant impacts of reading score on ISC in the full cohort ($n = 68$). Each cluster has a pair of plots with a title whose color matches the corresponding cluster seen on the brain to its left. Outlines indicate the largest cluster's overlap with a language region derived from an atlas. Below each title, we see at the top the mean BOLD timecourse across better readers (red) and worse readers (purple) in the cluster. Below this is the mean pairwise ISC in a 30-second window ending at each time point (i.e., the value at each time is the ISC in the previous 30 s). Patches represent standard error of the mean. The magenta and cyan bars behind each plot show the times when auditory and visual stimuli were being presented (these times are offset 6 s to account for typical hemodynamic delays).

The spatial distribution of these areas includes several regions not typically implicated in reading studies, suggesting an expansion of the current view of where and how reading ability is reflected in the brain.

4.1. ISC offers increased sensitivity

ISC lends itself well to stimuli with an extended narrative where the neural response model is unknown, and reading certainly matches this description. ISC's naturalistic approach allows stimulus-locked analyses with fewer assumptions than traditional block or event-related designs (Finn et al., 2020). In this dataset, we used extended narratives and ISC together, providing an increase in both ecological validity and sensitivity. In single-group analyses, ISC analysis revealed more brain regions that were responding to the task than did traditional block-design analyses (Supplementary Fig. S1). When studying group differences, block-design analyses failed to find any significant differences, but ISC analyses showed greater sensitivity and revealed widespread response differences between groups (Fig. 3).

One might argue that ISC has greater sensitivity in our analyses only because our task was designed for ISC analysis and not block analysis. However, a previous study found that this is not true when the roles are reversed: when analyzing conventional stimuli designed for GLMs, ISC and GLMs produced similar results (Pajula et al., 2012). ISC's superior sensitivity is further supported by the wider set of brain regions it identifies in this study as reading-ability-related compared to previous studies designed for GLM analysis. This suggests that the naturalistic

stimuli and analyses in our study offer increased sensitivity to reading-ability-related neural activity over the conventional stimuli and analyses used in past studies.

The likely reason for ISC's increased sensitivity can be seen in the ROI timecourses of BOLD activity (Supplementary Fig. S10). The time-courses of activity in each ROI contained patterns that were consistent between independent groups but were not captured by the task's block design. These patterns provide a wealth of information on participant similarities and differences. Because ISC uses one subject's brain activity as a model for another's, it is more sensitive to these brain activity patterns than block analysis is, as previous studies have illustrated (Hasson et al., 2010; Pajula et al., 2012). The increased sensitivity of ISC to group differences is less well studied. While several studies have used two-group ISC formulations to study group differences (Hasson et al., 2009; Salmi et al., 2013; Byrge et al., 2015; Finn et al., 2018; Gruskin and Patel, 2022), few have done so with dimensional analyses (Chen et al., 2017), and this study is the first to compare its sensitivity directly to that of a standard block analysis.

Previous research has used fMRI ISC to investigate inter-subject variability by presenting stimuli relevant to the trait being investigated. Participants with autism spectrum disorder were found to have lower ISC with other subjects than healthy controls did when watching movies of social interactions (Hasson et al., 2009; Salmi et al., 2013; Byrge et al., 2015). Patients with a melancholic subtype of depression had reduced ISC when watching emotional films (Guo et al., 2015), and those with first-episode psychosis had reduced ISC during a film

(Mäntylä et al., 2018). (In a notable exception to the link between symptom severity and decreased ISC, participants with higher trait paranoia had *higher* ISC when listening to an ambiguous story that aroused suspicion (Finn et al., 2018)). Our study extends this method of fMRI ISC individual differences to study differences in reading ability in adolescents and young adults. Like these previous studies, we find that individuals with greater symptoms of dyslexia have altered ISC in regions related to that condition. In our case, this means reduced ISC in lfus, lAG, and lIFG (Fig. 3), part of the canonical “reading circuit” regions Oc-T, TMP, and IFG (Pugh et al., 2001). However, worse readers in our analysis also showed reduced ISC in numerous other regions, expanding the picture of reading ability’s correlates beyond the canonical reading circuit.

4.2. ISC expands our view of reading in the brain

Worse readers showed reduced ISC in widespread regions (Fig. 3). While this is a departure from more conventional studies of reading skill, it is perhaps not surprising given that much of the brain responds in some way to semantic information (Huth et al., 2016). If worse readers are absorbing less of the semantic content of a story, it stands to reason that activity in these widespread regions will be less stimulus-driven and, therefore, less synchronized with other readers. We note that in our single-group results, primary sensory cortices show the highest ISC and reading-ability-related ISC effects. But many of the regions showing reading ability impacts on ISC were higher-level regions that respond to semantic meaning. This suggests that the regions whose ISC is most related to reading ability may extend both “above” and “below” any neural deficit(s) in the semantic processing hierarchy.

Brain areas showing significant impacts of reading ability on ISC included frontal regions that previous studies have not implicated, particularly the ACC and lMidFG. These regions are anterior to the canonical reading circuit region lIFG, where block-design studies of dyslexia have found group differences (Temple et al., 2001). The ACC has been implicated in varied functions including the control of continuous behavior in dynamic contexts, where it integrates information across multiple timescales (Monosov et al., 2020). Our lMidFG cluster may be part of the dorsolateral prefrontal cortex (DLPFC), which is often implicated in executive control and task switching, but recent work suggests it also plays a role in sentence (as opposed to single-word) processing, predictive top-down speech processing, and functionally connecting the language network with other functional networks (Herrlich et al., 2021). It plays a role in the fronto-parietal network identified as a deficit in dyslexia in a PET/fMRI meta-analysis (Paulescu et al., 2014). Such frontal regions tend to encode information on a longer timescale than primary sensory cortices, requiring narratives longer than a sentence (Lerner et al., 2011; Hasson et al., 2008). This may help explain why our study found effects in these regions where some previous studies using only brief, randomized word or phrase stimuli have not.

Our group difference results expand upon recent magnetoencephalography (MEG) ISC findings on the response of dyslexic Finnish adults to spoken narratives (Thiede et al., 2020). This analysis broke down responses into various frequency bands and identified ISC differences between dyslexic and control groups in the envelope of each band. These results diverge significantly from ours, finding multiple regions where dyslexic participants’ ISC exceeded that of the control group. The different ages, diagnoses, and stimuli used in our study could contribute to this disagreement, and the imaging method is likely a strong contributing factor. In a comparison study, fMRI exhibited greater intra- and inter-subject ISC than MEG, and the two modalities generally did not agree on timecourses outside of primary sensory regions (Lankinen et al., 2018).

4.3. Dissociating reading ability and IQ

Because reading ability was correlated with IQ in our sample, we took care to investigate and control for it as a potential confound. In addition to including this potential confound in our LME model, we observed that a subset of subjects matched for IQ continued to show significant impacts of reading score on ISC in a widespread set of brain regions. This matching procedure is a conservative step that many studies would omit because it produces overcorrected findings (Dennis et al., 2009). We include it here to illustrate that the full-cohort findings cannot be explained by IQ alone, and that overall findings are robust even to this stringent correction.

4.4. Lack of consistent compensatory mechanism

The lack of regions in which worse readers had significantly *greater* ISC than better readers suggests that our ISC results were not driven by a “compensatory mechanism” in which worse readers consistently rely on alternate brain regions for reading comprehension (Yu et al., 2018). Indeed, the heterogeneity of behavioral deficits and cognitive profiles associated with dyslexia makes it unlikely that a common compensatory mechanism would be adopted among dyslexic participants (Zoubinetsky et al., 2014).

Our IS-RSA analysis with different reading test subscores did not promote this compensatory mechanism notion (Supplementary Fig. S11). If distinct deficits yielded distinct compensatory mechanisms, we might expect some regions to show a negative correlation between ISC and subtest scores (i.e., participants who struggle with a particular skill use a consistent region to compensate), but this was not observed. In different regions, distinct sets of subjects do appear to have low ISC with the others (Fig. 4, middle column). But we were not able to discern whether these distinctions are linked to measurable behavioral differences.

We also used IS-RSA analysis to investigate an alternative model that might illuminate compensatory mechanisms (Fig. 5). This dimensional analysis confirmed that ROIs across the brain had patterns of ISC that matched an “Anna Karenina” model, in which higher reading scores predicted better synchronization with all other participants. This model strongly outperformed a “Nearest Neighbor” model, in which similar reading scores predicted better synchronization (Supplementary Fig. S3). This suggests that worse readers had “idiosyncratic” responses that were more different from each other than they were from those of better readers.

4.5. Impact of stimulus modality

Our use of both visual and auditory blocks was unusual, but our single-group ISC results largely agreed with previous studies. The extended narrative induced patterns of activity in many brain regions that were more consistent across participants than would be expected by chance, including frontal cortices in addition to primary sensory areas. This result agrees with previous studies contrasting the ISC across written and spoken narratives (Regev et al., 2013). Our modality-specific analyses (Supplementary Fig. S5) disagree, however: Regev et al. found that primary auditory cortex had higher ISC during auditory narratives, but in our analyses, both primary auditory and visual cortices exhibited greater ISC during visual blocks than during auditory ones. This may be related to the fact that when separating by modality we excluded transition periods, which in some regions induced greater ISC than the interior of either auditory or visual blocks (Supplementary Fig. S10). In general, our results reinforce previous findings that reading skill affects semantic processing regardless of input modality (Gibson et al., 2006).

4.6. Limitations

The original study for which these data were collected was not designed for the ISC analyses in this manuscript, so the behavioral test choices were not optimized for it. The WJ3 and TOWRE tests used to assess reading ability are word-level reading tests. Contextual reading/listening tasks most often take much longer than word level reading tasks, and the original study had to prioritize the word-level tests to accommodate the time constraints of the child. It is possible that a different set of brain regions would have ISC that correlates with contextual reading ability.

The paradigm used in this study includes unnatural transitions between visual and auditory delivery of the narrative. Although this allowed us to explore the influence of modality (Supplementary Fig. S5), we cannot be sure how a more naturalistic reading experience would change the results. We also note that the complete set of visual blocks had a slightly longer duration (214 s) than auditory blocks (178 s). This might make the task-vs.-baseline results slightly more representative of the effects of visual blocks than of auditory blocks. The regions most likely to be affected by this imbalance can be seen in Supplementary Fig. S5D. Most regions outside of primary sensory cortices did not show a significant ISC difference between block types.

This study focused on adolescents and young adults, an age range with benefits and challenges in reading research. Encompassing high school and college years when reading skills are relatively mature and crucial to academic success, this age range is often analyzed together in neuroimaging and wider medical research (Nielson et al., 2016; Torre and Eden, 2019; Sawyer et al., 2018). Still, the adolescent brain is undergoing significant changes and can vary systematically with age (Simmonds et al., 2014; Paulesu et al., 2014). We took several steps to address this concern. First, behavioral scores were normed for age before being used to sort participants. Second, we performed a dimensional IS-RSA analysis and found that no regions in the Shen atlas had ISC patterns that correlated significantly with the mean ages of each pair. Together, this suggests that age was not a major factor modifying ISC in our sample. Nevertheless, our results may differ from those of an early-childhood or adult sample.

In almost all regions examined, worse readers had greater ISC with better readers than they did with other worse readers (Fig. 4, right), and worse readers tended to have more idiosyncratic responses (Fig. 5). This pattern could result from group differences in sustained attention or mind-wandering (i.e., worse readers were less engaged with the task), and we are unable to use behavior to assess whether both groups attended to and comprehended the stories equally. However, several pieces of evidence argue against this sustained attention explanation in our results. The distinct temporal profiles of sliding-window ISC across different ROIs (Supplementary Fig. S10) suggests that the findings are not driven by global differences in physiology, sustained attention, engagement, or arousal. The fact that ISC differences are not restricted to the beginning or end of the run provides additional evidence that group differences are not attributable to onset or offset artifacts or differences in susceptibility to fatigue. In the regions with different group ISC, worse readers do not typically have smaller average BOLD responses at the block level, as we might expect (Supplementary Fig. S10, top subplots in each pair). During the task, worse readers do not exhibit significantly lower BOLD activity in visual and auditory cortices, or significantly higher BOLD activity in default mode regions, as we would expect in readers who ignore the reading and instead retrieve autobiographical memories as is common in mind-wandering (Zhang et al., 2022). These results support our preferred explanation that worse readers' ISC reductions are not driven by reductions in attention and arousal, but rather by idiosyncratic responses that begin with deficits low in the semantic hierarchy and continue at regions higher in the hierarchy that rely on them for input.

We were unable to replicate previous GLM-based findings reported in meta-analyses of reading and dyslexia, such as reduced

temporoparietal and/or occipitotemporal activation in worse readers (Shaywitz et al., 2001; Pugh et al., 2001; Martin et al., 2015; Zhang and Peng, 2022). We see three factors that may have contributed to these null findings. First, our task is an extended narrative rather than many brief reading events, possibly leading to this difference. Second, our participants were recruited from a community sample of the population and examined using a dimensional MVM analysis rather than recruited from dyslexic and typically reading populations grouped by diagnosis. Third, it is possible that previous group difference findings are spurious, as the significant publication bias towards positive findings can lead to false positives in small samples being reported and true negatives being omitted from publication (Ioannidis, 2011; Button et al., 2013; Jednorog et al., 2015). To address these factors, we hope that the promising results in the current study will motivate future work utilizing naturalistic stimuli to differentiate clinical samples from typical development in large samples.

5. Conclusion

This study combined extended auditory and visual narratives with inter-subject correlation analysis to reveal how stimulus-locked activity across the brain differs with reading skill in a group of adolescents and young adults. The results identified multiple regions beyond the canonical reading circuit that were affected by reading skill. We hope that future work will include targeted experiments to elucidate the role of these regions in narrative processing. Future longitudinal data collection may also allow this ISC method to be used to predict the outcome of reading interventions before they begin (similar to Aboud et al. (2018)). Such results could be used to develop personalized interventions and select the best treatment for an individual in advance (Gabrieli et al., 2015).

CRediT authorship contribution statement

David C. Jangraw: Data curation, Formal analysis, Methodology, Software, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Emily S. Finn:** Methodology, Software, Writing – review & editing. **Peter A. Bandettini:** Conceptualization, Supervision, Resources, Funding acquisition, Writing – review & editing. **Nicole Landi:** Conceptualization, Investigation, Data curation. **Haorui Sun:** Software, Formal analysis, Visualization. **Fumiko Hoeft:** Conceptualization, Writing – review & editing. **Gang Chen:** Methodology, Software, Writing – review & editing. **Kenneth R. Pugh:** Conceptualization, Supervision, Resources. **Peter J. Molfese:** Conceptualization, Investigation, Data curation, Project administration, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Data are publicly available on the Open Science Framework at <https://osf.io/qk5yd/>. All code used to produce the results in this study is publicly available on GitHub at <https://github.com/djangraw/ReadingAbilityIsc>.

Acknowledgments

The authors thank Stephen J Frost for his assistance with the manuscript. They thank Airey Lau for her assistance with the manuscript and data collection. Data collection and curation was supported by the National Institute of Child Health and Human Development (NICHD) R01HD065794 (PI: K. Pugh). Analyses were supported by the Intramural

Research Program of the NIMH (annual report ZIAMH002783). Portions of this study used the computational capabilities of the NIH HPC Biowulf cluster (biowulf.nih.gov).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2023.120390](https://doi.org/10.1016/j.neuroimage.2023.120390).

References

- Aboud, K.S., Barquero, L.A., Cutting, L.E., 2018. Prefrontal mediation of the reading network predicts intervention response in dyslexia. *Cortex* 101, 96–106. <https://doi.org/10.1016/J.CORTEX.2018.01.009>.
- Bailey, L.M., McMillan, L.E., Newman, A.J., 2020. A sinister subject: quantifying handedness-based recruitment biases in current neuroimaging research. *Eur. J. Neurosci.* 51 (7), 1642–1656.
- Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14 (5), 365–376.
- Byrge, L., Dubois, J., Tyszka, J.M., Adolphs, R., Kennedy, D.P., 2015. Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. *J. Neurosci.* 35 (14), 5837. <https://doi.org/10.1523/JNEUROSCI.5182-14.2015>.
- Chen, G., Adleman, N.E., Saad, Z.S., Leibenluft, E., Cox, R.W., 2014. Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. *Neuroimage* 99, 571–588.
- Chen, G., Taylor, P.A., Shin, Y.W., Reynolds, R.C., Cox, R.W., 2017. Untangling the relatedness among correlations, Part II: inter-subject correlation group analysis through linear mixed-effects modeling. *Neuroimage* 147, 825–840.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29 (3), 162–173. <https://doi.org/10.1006/cbmr.1996.0014>.
- Dehaene, S., le Clec'h, G., Poline, J.B., le Bihan, D., Cohen, L., 2002. The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport* 13 (3). <https://doi.org/10.1097/00001756-200203040-00015>.
- Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. *Trends Cogn. Sci. (Regul. Ed.)* 15 (6), 254–262.
- Dennis, M., Francis, D.J., Cirino, P.T., Schachar, R., Barnes, M.A., Fletcher, J.M., 2009. Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *J. Int. Neuropsychol. Soc.* 15 (3), 331–343. <https://doi.org/10.1017/S1355617709090481>.
- Ferrer, E., Shaywitz, B.A., Holahan, J.M., Marchione, K., Shaywitz, S.E., 1999. Uncoupling of reading and IQ over time: empirical evidence for a definition of dyslexia. *Psychol. Sci.* 21 (1), 93–101. <https://doi.org/10.1177/0956797609354084>.
- Finn, E.S., Corlett, P.R., Chen, G., Bandettini, P.A., Constable, R.T., 2018. Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. *Nat. Commun.* 9 (1), 2043. <https://doi.org/10.1038/s41467-018-04387-2>.
- Finn, E.S., Glerean, E., Khojandi, A.Y., Nielson, D., Molfese, P.J., Handwerker, D.A., Bandettini, P.A., 2020. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. *Neuroimage* 215. <https://doi.org/10.1016/J.NEuroIMAGE.2020.116828>.
- Fischl, B., 2012. FreeSurfer. *Neuroimage* 62 (2), 774–781. <https://doi.org/10.1016/J.NEuroIMAGE.2012.01.021>.
- Gabrieli, J.D.E., Ghosh, S.S., Whitfield-Gabrieli, S., 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron* 85 (1), 11–26. <https://doi.org/10.1016/J.NEURON.2014.10.047>.
- Gibson, L.Y., Hogben, J.H., Fletcher, J., 2006. Visual and auditory processing and component reading skills in developmental dyslexia. *Cogn. Neuropsychol.* 23 (4), 621–642. <https://doi.org/10.1080/02643290500412545>.
- Goldston, D.B., Walsh, A., Mayfield Arnold, E., Reboussin, B., Sergeant Daniel, S., Erkanli, A., Nutter, D., Hickman, E., Palmes, G., Snider, E., Wood, F.B., 2007. Reading problems, psychiatric disorders, and functional impairment from mid- to late adolescence. *J. Am. Acad. Child Adolesc. Psychiatry* 46 (1), 25–32. <https://doi.org/10.1097/01.CHI.0000242241.77302.F4>.
- Gollo, L.L., Zalesky, A., Hutchison, R.M., Van Den Heuvel, M., Breakspear, M., 2015. Dwelling quietly in the rich club: brain network determinants of slow cortical fluctuations. *Philos. Trans. R. Soc. B Biol. Sci.* 370 (1668), 20140165.
- Gruskin, D.C., Patel, G.H., 2022. Brain connectivity at rest predicts individual differences in normative activity during movie watching. *Neuroimage* 253, 119100. <https://doi.org/10.1016/J.NEuroIMAGE.2022.119100>.
- Guo, C.C., Nguyen, V.T., Hyett, M.P., Parker, G.B., Breakspear, M.J., 2015. Out-of-Sync: Disrupted Neural Activity in Emotional Circuitry During Film Viewing in Melancholic Depression. Nature Publishing Group. <https://doi.org/10.1038/srep11605>.
- Hasson, U., Avidan, G., Gelbard, H., Vallines, I., Harel, M., Minshew, N., Behrmann, M., 2009. Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Res.* 2 (4), 220–231. <https://doi.org/10.1002/aur.89>.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. *Science* 303 (5664), 1634–1640.
- Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N., 2008. A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* 28 (10), 2539–2550. <https://doi.org/10.1523/JNEUROSCI.5487-07.2008>.
- Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* 14 (1), 40–48. <https://doi.org/10.1016/J.TICS.2009.10.011>.
- Hertrich, I., Dietrich, S., Blum, C., Ackermann, H., 2021. The role of the dorsolateral prefrontal cortex for speech and language processing. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2021.645209>.
- Hogan, T.P., Adlof, S.M., Alonso, C.N., 2014. On the importance of listening comprehension. *Int. J. Speech Lang. Pathol.* 16 (3), 199–207.
- Huth, A.G., de Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532 (7600), 453–458. <https://doi.org/10.1038/nature17637>.
- Ioannidis, J.P., 2011. Excess significance bias in the literature on brain volume abnormalities. *Arch. Gen. Psychiatry* 68 (8), 773–780.
- Jednoróg, K., Marchewka, A., Altarelli, I., Monzalvo Lopez, A.K., van Ermingen-Marbach, M., Grande, M., Grabowska, A., Heim, S., Ramus, F., 2015. How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? Insights from a large-scale voxel-based morphometry study. *Hum. Brain Mapp.* 36 (5), 1741–1754. <https://doi.org/10.1002/hbm.22734>.
- Jo, H.J., Saad, Z.S., Simmons, W.K., Milbury, L.A., Cox, R.W., 2010. Mapping sources of correlation in resting state fMRI, with artifact detection and removal. *Neuroimage* 52 (2), 571–582. <https://doi.org/10.1016/J.NEuroIMAGE.2010.04.246>.
- Kearns, D.M., Hancock, R., Hoeft, F., Pugh, K.R., Frost, S.J., 2018. The neurobiology of dyslexia. *Teach. Except. Child.* 51 (3), 175–188. <https://doi.org/10.1177/0040059918820051>.
- Lankinen, K., Saari, J., Hluchchuk, Y., Tikka, P., Parkkonen, L., Hari, R., Koskinen, M., 2018. Consistency and similarity of MEG- and fMRI-signal time courses during movie viewing. *Neuroimage* 173, 361–369. <https://doi.org/10.1016/J.NEuroIMAGE.2018.02.045>.
- Lerner, Y., Honey, C.J., Silbert, L.J., Hasson, U., 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* 31 (8), 2906–2915. <https://doi.org/10.1523/JNEUROSCI.3684-10.2011>.
- Mäntylä, T., Nummenmaa, L., Rikandi, E., Lindgren, M., Kieseppä, T., Hari, R., Suvisaari, J., Raji, T.T., 2018. Aberrant cortical integration in first-episode psychosis during natural audiovisual processing. *Biol. Psychiatry* 84 (9), 655–664. <https://doi.org/10.1016/J.BIOPSYCH.2018.04.014>.
- Martin, A., Schurz, M., Kronbichler, M., Richlan, F., 2015. Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36 (5), 1963–1981. <https://doi.org/10.1002/hbm.22749>.
- Monosov, I.E., Haber, S.N., Leuthardt, E.C., Jezzini, A., 2020. Anterior cingulate cortex and the control of dynamic behavior in primates. *Curr. Biol.* 30 (23), R1442–R1454.
- Morris, R.D., Lovett, M.W., Wolf, M., Sevcik, R.A., Steinbach, K.A., Frijters, J.C., Shapiro, M.B., 2012. Multiple-component remediation for developmental reading disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *J. Learn. Disabil.* 45 (2), 99–127.
- Nastase, S.A., Gazzola, V., Hasson, U., Keysers, C., 2019. Measuring shared responses across subjects using intersubject correlation. In: *Social Cognitive and Affective Neuroscience*, 14. Oxford University Press, pp. 667–685.
- Nielsen, K., Abbott, R., Griffin, W., Lott, J., Raskind, W., Berninger, V.W., 2016. Evidence-based reading and writing assessment for dyslexia in adolescents and young adults. *Learn. Disabil. 21 (1), 38 (Pittsburgh, Pa.)*.
- Norton, E.S., Beach, S.D., Gabrieli, J.D., 2015. Neurobiology of dyslexia. *Curr. Opin. Neurobiol.* 30, 73–78.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Ozernov-Palchik, O., Yu, X., Wang, Y., Gaab, N., 2016. Lessons to be learned: how a comprehensive neurobiological framework of atypical reading development can inform educational practice. *Curr. Opin. Behav. Sci.* 10, 45–58.
- Pajula, J., Kauppi, J.P., Tohka, J., 2012. Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. *PLoS ONE* 7 (8), e41196. <https://doi.org/10.1371/journal.pone.0041196>.
- Paulsesu, E., Danelli, L., Berlingeri, M., Zoccolotti, P., Pernet, C.R., 2014. Reading the Dyslexic Brain: Multiple Dysfunctional Routes Revealed by a New Meta-Analysis of PET and fMRI Activation Studies. <https://doi.org/10.3389/fnhum.2014.00830>.
- Price, C.J., Devlin, J.T., 2011. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci. (Regul. Ed.)* 15 (6), 246–253.
- Price, C.J., 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62 (2), 816–847.
- Pugh, K.R., Mencl, W.E., Jenner, A.R., Lee, J.R., Katz, L., Frost, S.J., Shaywitz, S.E., Shaywitz, B.A., 2001. Neuroimaging studies of reading development and reading disability. *Learn. Disabil. Res. Pract.* 16 (4), 240–249.
- Pugh, K.R., Frost, S.J., Sandak, R.D., Landi, N., Moore, D., della Porta, G., Cornelissen, P.L., Hansen, P.C., Kringselbach, M.L., Pugh, K.R., 2010. Mapping the word reading circuitry in skilled and disabled readers. In: Cornelissen, P., Hansen, P., Kringselbach, M., Pugh, K. (Eds.), *The Neural Basis of Reading*. Oxford University Press, pp. 281–305. <https://doi.org/10.1093/acprof:oso/9780195300369.003.0011>.
- Pugh, K.R., Landi, N., Preston, J.L., Mencl, W.E., Austin, A.C., Sibley, D., Fulbright, R.K., Seidenberg, M.S., Grigorenko, E.L., Constable, R.T., Molfese, P., Frost, S.J., 2013. The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain Lang.* 125 (2), 173–183. <https://doi.org/10.1016/J.BANDL.2012.04.004>.

- Regev, M., Honey, C.J., Simony, E., Hasson, U., 2013. Selective and invariant neural responses to spoken and written narratives. *J. Neurosci.* 33 (40), 15978. <https://doi.org/10.1523/JNEUROSCI.1580-13.2013>.
- Ryherd, K., Jasinska, K., Van Dyke, J.A., Hung, Y.H., Baron, E., Mencl, W.E., Zevin, J., Landi, N., 2018. Cortical regions supporting reading comprehension skill for single words and discourse. *Brain Lang.* 186, 32–43.
- Salmi, J., Roine, U., Glerean, E., Lahnakoski, J., Nieminen-Von Wendt, T., Tani, P., Leppämäki, S., Nummenmaa, L., Jääskeläinen, I.P., Carlson, S., Rintahaka, P., Sams, M., 2013. The brains of high functioning autistic individuals do not synchronize with those of others. *Neuroimage Clin.* 3, 489–497. <https://doi.org/10.1016/J.NICL.2013.10.011>.
- Sawyer, S.M., Azzopardi, P.S., Wickremarathne, D., Patton, G.C., 2018. The age of adolescence. *Lancet Child Adolesc. Health* 2 (3), 223–228.
- Schlaggar, B.L., McCandliss, B.D., 2007. Development of neural systems for reading. *Annu. Rev. Neurosci.* 30 (1), 475–503. <https://doi.org/10.1146/annurev.neuro.28.061604.135645>.
- Schultz, J., Pilz, K.S., 2009. Natural facial motion enhances cortical responses to faces. *Exp. Brain Res.* 194, 465–475.
- Shaywitz, B.A., Shaywitz, S.E., Pugh, K.R., Fulbright, R.K., Mencl, W.E., Constable, R.T., Gore, J.C., 2001. The neurobiology of dyslexia. *Clin. Neurosci. Res.* 1 (4), 291–299.
- Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* 82, 403–415. <https://doi.org/10.1016/j.neuroimage.2013.05.081>.
- Simonds, D.J., Hallquist, M.N., Asato, M., Luna, B., 2014. Developmental stages and sex differences of white matter and behavioral development through adolescence: a longitudinal diffusion tensor imaging (DTI) study. *Neuroimage* 92, 356–368.
- Simos, P.G., Rezaie, R., Papanicolau, A.C., Fletcher, J.M., 2014. Does IQ affect the functional brain network involved in pseudoword reading in students with reading disability? A magnetoencephalography study. *Front. Hum. Neurosci.* 7, 932.
- Sonkusare, S., Breakspear, M., Guo, C., 2019. Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* 23 (8), 699–714.
- Stuebing, K.K., Barth, A.E., Molfese, P.J., Weiss, B., Fletcher, J.M., 2009. IQ is not strongly related to response to reading instruction: a meta-analytic interpretation. *Except. Child.* 76 (1), 31–51. <https://doi.org/10.1177/001440290907600102>. Oct 1PMID: 20224749; PMCID: PMC2836021.
- Szafarski, J.P., Binder, J.R., Possing, E.T., McKiernan, K.A., Ward, B.D., Hammeken, T.A., 2002. Language lateralization in left-handed and ambidextrous people. *Neurology* 59 (2), 238. <https://doi.org/10.1212/WNL.59.2.238>.
- Tanaka, H., Black, J.M., Hulme, C., Stanley, L.M., Kesler, S.R., Whitfield-Gabrieli, S., Hoeft, F., 2011. The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychol. Sci.* 22 (11), 1442–1451.
- Temple, E., Poldrack, R.A., Salidis, J., Deutsch, G.K., Tallal, P., Merzenich, M.M., Gabrieli, J.D.E., 2001. Disrupted neural responses to phonological and orthographic processing in dyslexic children: an fMRI study. *Neuroreport* 12 (2), 299–307.
- Thiede, A., Glerean, E., Kujala, T., Parkkonen, L., 2020. Atypical MEG inter-subject correlation during listening to continuous natural speech in dyslexia. *Neuroimage* 216, 116799. <https://doi.org/10.1016/J.NEUROIMAGE.2020.116799>.
- Toffalini, E., Giofrè, D., Pastore, M., Carretti, B., Fraccadori, F., Szűcs, D., 2021. Dyslexia treatment studies: a systematic review and suggestions on testing treatment efficacy with small effects and small samples. *Behav. Res. Methods* 53, 1954–1972. <https://doi.org/10.3758/s13428-021-01549-x/Published>.
- Torgesen, J.K., Wagner, R.K., Rashotte, C.A., 1999. Test of Word Reading Efficiency (TOWRE). Pro-Ed, Austin, TX.
- Torre, G.A.A., Eden, G.F., 2019. Relationships between gray matter volume and reading ability in typically developing children, adolescents, and young adults. *Dev. Cogn. Neurosci.* 36, 100636.
- Vandermosten, M., Hoeft, F., Norton, E.S., 2016. Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: a review and quantitative meta-analysis. *Curr. Opin. Behav. Sci.* 10, 155–161.
- Wechsler, D. (2011). Wechsler abbreviated scale of intelligence, Second Edition (WASI-II).
- Woodcock, R.W., McGrew, K.S., Mather, N., 2001. Woodcock-Johnson III Test. Riverside Publishing Company, Itasca, IL.
- Yeatman, J.D., White, A.L., 2021. Reading: the confluence of vision and language. *Annu. Rev. Vis. Sci.* 7, 487–517.
- Yu, X., Zuk, J., & Gaab, N. (2018). What factors facilitate resilience in developmental dyslexia? Examining protective and compensatory mechanisms across the neurodevelopmental trajectory. [10.1111/cdep.12293](https://doi.org/10.1111/cdep.12293).
- Zhang, M., Bernhardt, B.C., Wang, X., Varga, D., Krieger-Redwood, K., Royer, J., Rodriguez-Cruces, R., et al., 2022. Perceptual coupling and decoupling of the default mode network during mind-wandering and reading. *eLife* 11, e74011.
- Zhang, Z., Peng, P., 2022. Reading real words versus pseudowords: a meta-analysis of research in developmental dyslexia. *Dev. Psychol.* 58 (6), 1035.
- Zoubrinetsky, R., Bielle, F., Valdois, S., 2014. New insights on developmental dyslexia subtypes: heterogeneity of mixed reading profiles. *PLoS ONE* 9 (6), e99337.