Published in final edited form as:

Cortex. 2019 February ; 111: 286–302. doi:10.1016/j.cortex.2018.11.009.

Brain activity in struggling readers before intervention relates to future reading gains

Tehila Nugiel^{1,*}, Mary Abbe Roe^{1,*}, W. Patrick Taylor², Paul T. Cirino², Sharon R. Vaughn³, Jack M. Fletcher², Jenifer Juranek⁴, and Jessica A. Church^{1,5}

¹Department of Psychology, The University of Texas at Austin, 108 E Dean Keeton St, Austin, TX 78712

²Department of Psychology, The University of Houston, 695 Cullen Boulevard Room 126, Houston, TX 77204

³Meadows Center for Prevention of Educational Risk, The University of Texas at Austin, 1912 Speedway D4900, Austin, TX 78712

⁴Department of Pediatrics, The University of Texas Health Science Center at Houston, 7000 Fannin St #1200, Houston, TX 77030

⁵Imaging Research Center, The University of Texas at Austin, 100 E 24th St, Austin, TX 78712

Abstract

Neural markers for reading-related changes in response to intervention may represent biomarkers that could inform intervention plans as a potential index of the malleability of the reading network in struggling readers. Particularly interesting is the role of activation outside the reading network, especially in executive control networks important for reading comprehension. However, it is unclear whether any intervention-related executive control changes in the brain are specific to reading tasks or reflect more domain general changes. Brain changes associated with reading gains over time were compared for a sentence comprehension task as well as for a non-lexical executive control task (a behavioral inhibition task) in upper-elementary struggling readers, and in grade-matched non-struggling readers. Functional MRI scans were conducted before and after 16 weeks of reading intervention. Participants were grouped as improvers and non-improvers based on the consistency and size of post-intervention gains across multiple post-test measures. Engagement of the right fusiform during the reading task, both before and after intervention, was related to gains from remediation. Additionally, pre-intervention activation in regions that are part of the default-mode network (precuneus) and the fronto-parietal network (right posterior middle temporal gyrus) separated improvers and non-improvers from non-struggling readers. None of these differences were observed during the non-lexical inhibitory control task, indicating that the

Correspondence: Tehila Nugiel, 108 E. Dean Keeton Stop A8000, Austin, TX 78712, tnugiel@utexas.edu.
*co first authors

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of interest: none

brain changes seen related to intervention outcome in struggling readers were specific to the reading process.

Keywords

Development; reading comprehension; intervention; fMRI; executive control

1. Introduction

A fundamental question in reading research concerns the basis for individual differences in instructional response, especially in an intervention context (Fletcher, Lyon, Fuchs, & Barnes, In Press). Cognitive studies have shown that the magnitude of response to instruction exists on a continuum (Vellutino, Scanlon, & Fanuele, 2006). The best predictors of reading intervention response have been assessments of baseline reading skills (Stuebing et al., 2015); more severely impaired readers are more likely to show inadequate instructional response, especially when successful response is defined as exceeding an outcome threshold (e.g., > 25th percentile). In addition, cognitive studies have shown that the best predictors of instructional response e.g., (phonological awareness, rapid naming, listening comprehension, vocabulary) also predict reading proficiency and failure, and generally align with the severity of reading difficulties (RD) (Al Otaiba & Fuchs, 2002; Cho et al., 2015; Fletcher et al., 2011; Miciak et al., 2014).

The neural correlates of instructional response in children are not well understood. A growing body of research examines the neural profiles of struggling readers who undergo a variety of reading interventions (e.g., letter-sound associations, timed reading, word study, and phoneme-grapheme correspondence). Many of the same brain regions that differentiate those with RD from non-struggling readers change in activation after remediation (Barquero, Davis, & Cutting, 2014). The most consistent findings from functional magnetic resonance imaging (fMRI) studies prior to intervention are under-activations in children with RD in the dorsal and ventral components of the reading network, spanning ventral fusiform, temporo-parietal cortex, and inferior frontal gyrus (Aylward et al., 2003; Hoeft et al., 2006; Price, 2012; Roe et al., In press), sometimes including subcortical structures (Meyler et al., 2007; Temple et al., 2003). After remediation, reading-related brain activation in these regions often ‘normalizes’, more closely resembling that of children without RD in studies using fMRI (Meyler, Keller, Cherkassky, Gabrieli, & Just, 2008; Odegard, Ring, Smith, Biggan, & Black, 2008b; Shaywitz et al., 2004) and magnetic source imaging (Rezaie et al., 2011; Simos et al., 2002). In addition to changes in activation levels during reading tasks, changes in functional connectivity and cortical thickness in these areas have also been reported post-intervention (Richards & Berninger, 2008; Romeo et al., 2017), suggesting that structural and functional evidence supports changes in reading-related regions after intervention.

There is increasing evidence of changes in additional brain regions that may reflect compensatory mechanisms or cognitive functions associated with intervention response. One possible set of these additional brain regions is linked to executive control processing (Hoeft et al., 2007a; Power & Petersen, 2013). Executive control includes domain general control

processes, or executive functions, like working memory, cognitive flexibility, and inhibition (Diamond, 2013; Mahone et al., 2002; Miyake et al., 2000). Executive control deficits have been linked to poor reading comprehension, lower academic achievement, and life success (Berninger, Abbott, Cook, & Nagy, 2017; Best, Miller, & Naglieri, 2011; Engelhardt, Church, Harden, & Tucker-Drob, *In Press*; Locascio, Mahone, Eason, & Cutting, 2010). Although behavioral studies using executive control measures have not identified robust predictors of reading intervention response (Miciak et al., 2014; Stuebing et al., 2015), Cirino et al. (in revision) found executive control was uniquely (but weakly) related to reading skills, even when controlling for demographic and language-related covariates. Additionally, for reading comprehension, executive control interacted with decoding and listening comprehension such that stronger executive control partially compensated for weak decoding. Taken together, executive control processes may play a supportive role in breaking apart and reconstructing text for comprehension and therefore are candidates for further study of how best to improve reading skills.

So far, few studies have examined the relation between intervention outcome and brain activity outside of the reading network (Aboud, Barquero, & Cutting, 2018; Horowitz-Kraus, Toro-Serey, & DiFrancesco, 2015b; Horowitz-Kraus, Vannest, Gozdasand, & Holland, 2014). Further, it is unclear if any such observed brain activity differences in these regions reflect domain general changes in processing for struggling readers, or if the differences are specific to reading-related tasks. Studies of children with RD have typically employed only reading-related tasks or investigated brain connectivity in reading-related regions at rest (Farris et al., 2011; Koyama et al., 2013; Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan, 2012; Temple et al., 2003). In this study, we test for pre- and post-intervention differences in a heterogeneous group of struggling readers. Unlike studies of struggling readers defined as dyslexic based on word reading accuracy and fluency deficits in which comprehension is expected to be impaired (Meyler et al., 2008; Shaywitz et al., 2004), our group of struggling readers included individuals with both basic reading (fluency) and comprehension deficits. We employ both a sentence comprehension task, and a non-lexical, but executive control-demanding response inhibition task (a variant of the classic Stop-Signal task) to test for the consistency of any control-related group differences across reading and non-reading tasks.

Inhibition, often considered a component of EF, is specifically important for resisting competing or distracting information during reading, as well as for controlling the retention and recall of relevant information for successful comprehension (Borella, Carretti, & Pelegrina, 2010). Inhibition has been broadly related to decoding, comprehension and overall academic achievement (Arrington, Kulesz, Francis, Fletcher, & Barnes, 2014; Borella et al., 2010; St. Clair-Thompson & Gathercole, 2006). Using an inhibition task in a reading intervention context provides an opportunity to examine how inhibition relates to comprehension, as well as to see whether there are brain changes related to reading intervention that extend beyond reading tasks.

In order to examine factors like executive control in reading improvement, it is also necessary to carefully consider how improvement in studies of reading is defined. There is no strong consensus about what constitutes sufficient change or growth to be considered

‘improvement’ after remediation. Group definitions of improvers/non-improvers are thus highly variable (Frijters, Lovett, Sevcik, & Morris, 2012). Some studies impose a cutoff where individuals who score above a certain threshold on a reading measure are considered improvers (Farris et al., 2011; Odegard, Ring, Smith, Biggan, & Black, 2008a). Others have done a median split of the group’s reading scores and label the top half improvers and bottom half non-improvers (Aboud et al., 2018; Davis et al., 2011; Hoeft et al., 2007b). Another method of defining improvement is to base the definition on an *individual’s* change in score relative to the typical amount of change that would be expected over a certain time period such as a year of school (Simos et al., 2007). This method is less common than other aforementioned definitions but it captures growth at the individual level irrespective of where an individual started or how the group changed overall. While there is rationale for each method, it is unclear how different ways of defining improvement impact study outcomes. Here, we define improvement based on the growth within an individual over time, similar to (Simos et al., 2007). However, as a post-hoc test of our “gain” definition, we examine the consistencies and differences in our results to two other common group approaches (median split and cut score) used to define improvement.

We addressed reading intervention-related brain changes through three main analyses. First, we tested for brain differences *before* intervention that were associated with whether or not a child with RD experiences future gains in reading. Second, we tested *after* intervention for any brain differences that related to individual change in reading scores. Third, we tested whether any observed differences during reading were seen during a non-lexical but executive control-demanding task. Finally, we conducted a post-hoc sensitivity analysis of how well our individual gain definition of improvement compared to whole group cut score and median split approaches.

2. Materials and Methods

2.1. Recruitment and classification

2.1.1. Participant selection variables—Recruitment for this neuroimaging study was done in conjunction with an in-school reading intervention study administered as part of the Texas Center for Learning Disabilities (TCLD) 2012–2017 project in Houston, and Austin, Texas. Each wave of the TCLD intervention project screened entire grade cohorts (3rd–5th) of school districts in the Houston and Austin areas. For this study, we report data from Austin site participants only because we did not want to combine across different scanners when testing for subtle effects in subgroups of struggling readers. Also, for this study we required struggling readers to have both a fluency and a comprehension difficulty. For the 2012–2014 cohort, children were classified as RD if they had a reading comprehension standardized score 85 or below on the Gates-MacGinitie Reading Test (GMRT) (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) and a score below 90 on either the Sight Word Efficiency subtest of the Test of Word Reading Efficiency (TOWRE-2), a fluency measure (Torgesen, Wagner, & Rashotte, 1999), or the Test of Silent Reading Efficiency and Comprehension (TOSREC), a measure of both comprehension and fluency (Wagner, Torgesen, Rashotte, & Pearson, 2010). For the 2014–2015 and 2015–2016 cohorts, the criteria for struggling readers was a score below 90 on the Test of Silent Reading Efficiency and Comprehension (TOSREC), a

measure of both comprehension and fluency (Wagner et al., 2010). The struggling readers were then randomized to business-as-usual (BAU) school treatment or the TCLD-designed intervention (See Supplementary section S1.5 Intervention). Non-struggling readers in 3rd-5th grade were recruited from the same school districts and the greater Austin community. Criteria for non-struggling readers for this analysis was a standard score above 90 on the TOSREC.

To maximize sample size, recruitment of struggling readers for neuroimaging involved participants from both treatment and BAU groups, as this analysis focused on brain changes related to reading changes, regardless of how the reading changes occurred. We collapsed across our intervention ($N = 29$) and BAU participants ($N = 24$) because all children in the BAU group received classroom reading instruction and many received additional school intervention. This allowed us to focus on any brain differences in individuals who consistently gained across at least two of the four standardized reading tests ('improvers') versus those who did not ('non-improvers'), irrespective of *how* they improved (i.e., the instruction type) (see definition of improvement below). Results from struggling readers include only participants from pre-intervention and post-intervention neuroimaging sessions who also had the pre- and post-intervention behavioral testing data necessary to identify reading improvement. However, we include imaging data from one struggling reader who was included in the school BAU intervention and was deemed a non-improver using scores available from post intervention.

Participants were excluded if they had MR scanner contraindications such as a non-removable metal implant or vision that could not be corrected with MR compatible glasses. Participants were also excluded if they were diagnosed with or taking medication for any disorders (non-struggling readers) or any disorders other than Attention-Deficit/Hyperactivity Disorder (ADHD) (struggling readers), since RD and ADHD are highly comorbid (Germano, Gagliano, & Curatolo, 2010) yet appear to have distinct effects on information processing (Dhar, Been, Minderaa, & Althaus, 2008; Gooch, Snowling, & Hulme, 2011). Non-struggling and struggling readers followed the same Scan 1 imaging protocol, and only struggling readers returned for the post-intervention (Scan 2) protocol. Participants completed the Scan 1 and 2 protocols at the Imaging Research Center at the University of Texas in Austin, Texas. See Table 1 and Table 2 for participant groups for each in-scanner task (for exclusions see Supplementary section 1: Participant groups).

2.1.2. Improver status: Individual gain approach—In addition to the above selection criteria, participants underwent a large battery of pre- and post-tests. We used four of these measures (the tests more consistently assessed in our neuroimaging sample) to define improvement in our struggling readers. These four measures include two subtests of the Woodcock-Johnson III (Woodcock, McGrew, & Mather, 2001) that measure decoding and comprehension, respectively: Letter Word Identification (WJIII-LWID) and Passage Comprehension (WJIII-PC); the TOWRE-2, a measure of word reading fluency; and the GMRT, a measure of reading comprehension. Reading progress over the school year for struggling readers was measured by calculating gain scores from the fall evaluation to the spring evaluation. Gains in standard scores were separately calculated for the GMRT-2, TOWRE-2, WJIII-LWID, and WJIII-PC. Improver status for each test was determined by

comparing a participant's gain score to their benchmark expected age-based gain score. Expected gains were calculated using benchmark effect sizes for academic gain within a certain grade, as reported by Scammacca, Fall & Roberts, 2015. Standard deviations for each test were multiplied by expected gain effect sizes to create expected gain scores for each school year grade and for each test. If the actual gain surpassed the expected gain for the academic year, the participant was determined to be an 'improver' for that test. Participants who were improvers on *two or more* of the four tests were determined to be overall '*improvers*', while those who improved on one or no tests were designated as '*non-improvers*'. The struggling readers from pre-intervention analyses were then back classified into groups that we refer to as '*future improvers*' (struggling readers before intervention who went on to be deemed 'improvers') or '*future non-improvers*' (struggling readers before intervention who went on to be deemed 'non-improvers') based on their subsequent spring behavioral evaluations (see Tables 1 & 2 for improver group characteristics). Additionally, we averaged across the four reading scores to generate a composite reading score. The composite was as an index of overall reading skill for each individual before and after reading instruction (Table 1). This score was used in correlational analyses with neural data (see 2.2.7 Post intervention correlate and ROI analyses and 2.2.8. Relating change in ROI activation to change in reading skill).

2.2. fMRI protocol

The Institutional Review Board of the University of Texas Health Science Center at Houston approved the current study for both Austin and Houston sites. Parents gave written consent for their child to participate in the research and children gave their informed assent. Participation was voluntary and families were informed they could terminate participation at any time. At the first session, participants and their parents went through the informed consent process, which included a description of the study parameters. Children underwent any academic and IQ testing that hadn't been previously conducted in the schools as part of the larger TCLD project and were exposed to the scanning environment in a mock scanner. At the same visit or a subsequent one as the testing and/or mock visit, children participated in an MRI session. The MRI visit(s) occurred prior to the bulk of any reading instruction related to the experiment (pre-intervention scan, Scan 1).

Participants were compensated for their time. Scanning protocols lasted roughly 90 minutes and consisted of a high resolution T1 and T2 anatomical scan, two resting-state fMRI scans, three Sentence Comprehension scans, two Stop-Signal task (SST) scans, and a diffusion-weighted imaging scan. This analysis reports results from the fMRI Sentence Comprehension and Stop-Signal tasks.

2.2.1. Sentence Comprehension task—The sentence comprehension (SC) task was adapted from Meyler and colleagues ((Meyler et al., 2007) Figure 1b). Each of three scans of the SC task was 7'06" and had 32 sentences. Each task scan began with the probe "Makes sense?" on screen for 2 seconds (s). Sentences were presented for 8 s, during which participants were instructed to make a choice regarding whether the sentence was sensible (e.g., 'A mouse chewed on the cheese') or non-sensible (e.g., 'The sock listened to the pig'), followed by 2 s of a blank screen inter-stimulus interval (ISI) and interspersed with jittered

blank screen ranging from 0–8 s. Sentences were in either active ('A father cooked the eggs') or passive voice ('The car was chased by a dog') and thus were one of four categories: active sensible, active non-sensible, passive sensible, and passive non-sensible. Half of the sentences were sensible ("Yes") and half were non-sensible ("No"). The words "Yes" and "No" appeared to the left or right below each sentence to remind the participants of which button had been assigned to which response; these reminders remained unchanged throughout the MRI session and were counterbalanced across participants.

Participants used their thumbs on a button box controller (Current Designs) to make a choice for each sentence. Participants practiced with a separate set of stimuli on a laptop before the scan session to familiarize themselves with the task. MRI task runs with less than 60% accuracy were excluded from the analyses. Our whole brain MRI results are discussed for all correct trials relative to a resting, task-absent baseline (*correct sentences vs. baseline*), collapsed across voice and sensibility types.

2.2.2 Stop-Signal task—The visual SST paradigm was adapted from a classic SST auditory paradigm (de Jong et al., 2009; Rubia, Smith, Brammer, & Taylor, 2003)(Figure 1c) to avoid auditory issues related to scanner noise. Arrows were presented pointing left or right on the screen and participants were instructed to quickly respond to the direction the arrow was pointing with a button press ("Go" trials) and to try to not respond to the arrow if a red X appeared over the arrow ("Stop" trials). "Go" trials consisted of an arrow presented for 1000 milliseconds (ms) followed by 1000 ms of a blank screen ISI with a 0–4 s jitter. "Stop" trials consisted of an arrow on the screen for (initially) 250 ms (stop signal delay, SSD), then a red X (stop signal) appeared over the arrow and remained for the rest of the trial. If participants correctly stopped, the SSD on the next stop trial was increased by 50 ms (SSD = 300 ms). If the participant was unable to stop and made an incorrect button press, the SSD on the next stop trial decreased by 50 ms (SSD = 200 ms). This staircasing procedure continued throughout the duration of the task. Two scans of the SST task were collected, each scan consisting of 96 "Go" trials and 32 "Stop" trials (6'0" runs, 180 frames each). Task runs were excluded if there was < 70% "Go" accuracy or > 10% go errors, if "Stop" accuracy was < 25% or > 75% or if stop signal reaction time (SSRT) was < 50 ms. To calculate SSRT the mean time between the presentation of the arrow and the appearance of an X (SSD) was subtracted from the mean response time to a "Go" trial (Congdon et al., 2012). To examine the specificity of the engagement of executive control during reading relating to improvement, group differences were examined for the *correct stop vs. correct go* contrast using whole brain and applied control network regions (see 2.2.5 Pre-intervention group whole brain and ROI analyses).

2.2.3. Image acquisition—MRI scanning was conducted at The University of Texas at Austin Imaging Research Center on a Siemens Skyra 3T scanner with a 32-channel head coil. High-resolution anatomical images were collected with T₁-weighted structural images (TR = 2530 ms, TE = 3.37 ms, FOV = 256, 1×1×1 mm voxels) and a turbo-spin echo sequence to collect T₂-weighted structural images (TR = 3200 ms, TE = 412 ms, FOV = 250, 1×1×1 mm voxels). For all functional imaging, a multi-band echo-planar sequence was used with the following parameters: multiband factor of 2, TR = 2000 ms, TE = 30 ms, flip

angle = 60 degrees, 48 axial slices, 2×2×2 mm voxels, and a base resolution of 128×128. All stimuli were presented using PsychoPy software (Peirce, 2007). Participants used a mirror attached to the head coil to view stimuli projected onto a screen located behind the scanner with a PROPixx projector at a resolution of 1920×1080. Optoacoustics (OptoACTIVE Optical MRI Communication System with Active Noise Control) headphones and a microphone were provided to block excessive scanner noise, to allow the participants to hear a movie during the structural sequences, and to enable communication with the researchers throughout the session.

2.2.4. Imaging processing—Processing and analysis of neuroimaging data was executed with version 5.0.2 of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl; (Smith et al., 2004). fMRI data processing was carried out using FEAT (FMRIB Expert Analysis Tool) Version 6.00. High-resolution T₁-weighted structural images were skull stripped and extracted using Freesurfer 5.3.0 (Reuter, Schmansky, Rosas, & Fischl, 2012) or BET (Brain Extraction Tool; (Smith, 2002)). We used the Boundary Based Registration algorithm to register the functional data to the high-resolution structural image (Greve & Fischl, 2009). Registration of the high resolution structural to standard space was executed using FMRIB's Linear Image Registration Tool (FLIRT; (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001)). Images were spatially smoothed using a Gaussian kernel of FWHM 5 mm; the 4D dataset was grand-mean intensity normalized by a single multiplicative factor; high pass temporal filtering (Gaussian-weighted least-squares, straight line fitting, with sigma = 50.0 s). A double-gamma HRF time-series model was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). The model included 6 motion regressors, as well as temporal derivatives for each motion regressor. Since head motion in young populations can be a source of imaging confounds (Engelhardt et al., 2017; Power, Schlaggar, & Petersen, 2015), we included nuisance regressors to model out single TR's flagged for excessive motion according to a Framewise Displacement (FD) > 0.9 mm (Siegel et al., 2014). Runs with at least 60% of frames (at least 127 out of 212 for SC, and 108 out of 180 for SST) remaining were included in further analyses. Additionally, we ensured there were no significant differences between any two groups before or after motion censoring in mean FD (all group t-tests $p > .1$).

Second-level modeling, which averages contrast estimates over runs within subject, was carried out using FMRIB's Local Analysis of Mixed Effects (FLAME; (Beckmann, Jenkinson, & Smith, 2003; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004)), forcing random effects variance to zero.

2.2.5. Pre-intervention group whole brain and ROI analyses—To examine differences between future improvers, future non-improvers, and non-struggling readers before intervention, third-level group analysis was also carried out using FLAME stage 1 (Beckmann et al., 2003; Woolrich, 2008; Woolrich et al., 2004). Group-level Z (Gaussianised T/F) statistic images from the Sentence Comprehension (*all correct vs. baseline*) and Stop-Signal Task (*correct stop vs. correct go*) contrasts of interest were thresholded to correct for multiple comparisons. A $Z > 3.1$ was used to define contiguous clusters with a cluster probability of $p < .05$ (Eklund, Nichols, & Knutsson, 2016). Gaussian

random field theory was used for whole-brain multiple comparison corrections (Worsley, 2001). All cortical regions are reported in MNI coordinates and identified using the Harvard-Oxford Atlas in the FMRIB FSL-view software. Data were projected onto inflated brains maps for visualization purposes using Caret software (Van Essen, 2012).

Region of interest (ROI) analyses were carried out using FSL and R version 3.2.1 (R Development Core Team, 2014). To examine group differences between improvers and non-improvers as well as improver groups and non-struggling readers, 25 literature-derived 8 mm radius ROIs (Rao & Singh, 2015) were created using the T1 MNI152 2 mm mask in FSL. Ten reading ROIs were chosen from the reading literature (Cohen & Dehaene, 2004; Rao & Singh, 2015; Richardson, Seghier, Leff, Thomas, & Price, 2011; Vogel et al., 2013) and 15 executive control (10 fronto-parietal and 5 cingulo-opercular) ROIs were identified from resting state fMRI literature (Dosenbach et al., 2007b; Greene et al., 2014) (Supplementary Table 1). All 25 regions were tested in the Sentence Comprehension *all correct vs. baseline* contrast. To test for group differences in control engagement in a non-reading inhibition task, the 15 executive control ROIs were applied to the Stop-Signal *correct stop vs. correct go* contrast. Mean percent (%) BOLD signal change was calculated for each ROI for each individual. Group ROI differences were calculated using the Welch's t-test adjustment for unequal variances and unequal sample sizes in R. Since alphas did not survive correction for multiple comparisons, reported alphas are uncorrected (see Supplementary section 3). A larger sample size would have provided more power for the post hoc tests, but adding additional participants was not feasible, given the limitations of school-based recruitment.

2.2.6 Sensitivity Analysis: Testing alternative improvement definitions—We wished to test the robustness of our “individual gain” approach to improver definition, and thus performed a post-hoc sensitivity analysis of our group definition by trying two other popular approaches. We then tested how those group definitions impacted the outcome of our whole brain group comparison models. For two additional sets of improver/non-improver groups (see below) we ran the same set of whole brain group comparisons for the SC and SST contrast of interest (improver vs. non-improver, and each vs. non-struggling) and assessed the degree of overlap among the results from each analysis. For details on the number of individuals in each group and degree of overlap see Table 3.

The first alternative method was a median split of the group based on the post intervention composite reading score (Median Split approach). This composite reading score was an average of the four standardized reading measures used to define growth (GMRT-2, WJIII-LWID, WJII-PC and TOWRE). Struggling readers in the top half of the group after intervention were labeled improvers ($N=23$) and the bottom half were non-improvers ($N=23$).

The second alternative definition was to implement a cutoff score of 90 or above on a standardized reading measure, post-intervention TOWRE-2, where those below 90 were labeled non-improvers ($N=23$) and those with a 90 or greater were labeled improvers ($N=23$) (Cut Score approach). By exploring these alternative but common methods of defining treatment improvement, we were able to empirically test the impact of defining

improvement from individual gain scores and to test the specificity or generality of our gain score definition.

2.2.7 Post intervention correlate and ROI analyses—Our post-intervention sample was half the size of our fall dataset. Therefore, we do not report the whole brain group difference analysis data to avoid type II errors from unreliable small group estimates. Rather than split into groups, we examined individual differences in the change in composite reading score across all of our struggling readers through a whole brain regression on the SC contrast of interest (*all correct vs. baseline*) and *SST contrast of interest (correct stop vs. correct go)*. Mean centered change in composite reading score (post-composite score – pre-composite score) for each participant was added as a third level correlate using FLAME stage 1 ($z > 3.1$, $p < .05$). We tested for group differences between improvers ($N = 15$) and non-struggling readers ($N = 28$) (our two largest subgroups) in our literature-derived ROI-set.

2.2.8. Relating change in ROI activation to change in reading skill—A subset of struggling reader participants had eligible data from both scan periods ($N = 16$; 9 females, 11 improvers). We examined whether change in reading skill related to change in brain activation at the ROI level. The change in composite reading score (post-pre) was tested for a relationship to change in % BOLD signal in a given ROI over scan session (post-pre) during the SC task. Pearson correlations between an individual's difference in composite reading score and difference in MR signal were carried out in R for the set of 25 ROIs described above.

3. Results

3.1. Group differences in reading assessments

When we tested for differences in composite reading score after instruction across all study participants, improvers had a higher composite reading score than non-improvers at scan 2 ($t(48.39) = -2.078$, $p = 0.04$). However, there were no differences between future improvers and future non-improvers at Scan 1 ($t(47.69) = .48$, $p = .63$; Figure 1a). As expected from our definition of improvement by individual gain over time, we also found a significant difference in average composite score gain over time (the difference between composite score between Scan 1 and Scan 2) between pre-intervention struggling reader groups ($t(36.41) = 4.59$, $p < .001$), as well as between post-intervention improvers and non-improvers ($t(20.45) = 4.38$, $p < .01$; Table 1).

3.2. fMRI task behavioral performance

There were no significant differences in accuracy or reaction time on the SC or SST tasks between future improvers and future non-improver before intervention, or improvers and non-improvers after intervention (all $ps > 0.05$). Struggling readers as a combined group were significantly slower and less accurate than non-struggling readers for the SC task, and the SST go accuracy measure. For a detailed description of task performance across groups see Supplementary section 2.

3.3. Neuroimaging: Pre-intervention Sentence Comprehension task

Scan 1 (Fall): Whole brain analysis.—Future improvers had more positive activation in right ventral fusiform cortex and right lingual gyrus cortex than future non-improvers during sentence processing (Figure 2a). Future improvers also had less activity relative to baseline than non-struggling readers in the precuneus, a region considered part of the default mode network (Fox & Raichle, 2007; Raichle et al., 2001) Figure 2b).

Future non-improvers had more decreased activity relative to baseline than non-struggling readers in right posterior middle temporal gyrus (MTG) and left post central gyrus white matter Figure 2c). Comparison with Neurosynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) suggests activity in this aspect of the right MTG is most often part of the fronto-parietal network at rest. For peak activation coordinates and cluster sizes see Table 4.

3.4. Neuroimaging: Pre-intervention Stop-Signal task

Before intervention, there were no whole brain differences between future improvers and future non-improvers or between either group and non-struggling readers for the *correct stop* vs. *correct go* contrast. Similarly, there were no significant group differences in ROI % BOLD signal change between improver groups or between improver groups and non-struggling readers, although all 15 of the applied task control ROIs were significantly active during the task across participants.

3.5. Neuroimaging: Pre-intervention ROI analyses Sentence Comprehension and Stop-Signal Task

We found intervention response group differences in some of our task-control ROIs using our SC task pre-intervention data. However, these effects did not survive correction for multiple comparisons. There were no differences between improver groups in any of the ROIs during the SST. The uncorrected results of these analyses can be found in the Supplementary section 3: ROI analyses results.

3.6. Neuroimaging: Sensitivity analysis of Improver definition

We tested two additional methods for defining our improver/non-improvers groups in our pre-intervention data. Using a composite score median split or a TOWRE 90 > cut score approach of defining improvement did not reveal the same right fusiform or lingual gyrus that were seen when comparing improvers and non-improvers using our primary definition: the individual gain method (Figure 2a). In fact, the median split and cut score approaches revealed no significant differences between future improvers and future non-improvers. However, we found that the group models comparing future improver to non-struggling, and future non-improvers to non-struggling, remained stable across the three methods of defining improvement. Replicating our individual gain method, the median composite split and TOWRE cut score approaches also found a precuneus region for future improvers vs. non-struggling contrast, and right MTG and left post central white matter differences in future non-improvers vs. non-struggling (Figure 3).

For the SST, consistent our original individual gain definition, the median composite split and TOWRE cut score definitions also found no group differences between future improvers and non-improvers or either group and non-struggling readers.

3.7. Neuroimaging: Post-intervention correlate analysis

After remediation there was a positive relationship between change in composite reading score over time, and neural activity during the SC task. Greater activity in the right fusiform during the post-intervention performance of the SC task was related to greater gains in composite reading score (Figure 4, Table 4). This activity partially overlapped with activity differentiating groups of future improvers and future non-improvers before intervention (Figure 2a).

There was no relationship between change in composite reading score and neural activity during the SST task.

3.8. Neuroimaging: Post-intervention ROI analyses of Improvers vs. Non-struggling

We found a pattern of intervention response group differences in some of both reading and task-control ROIs during the sentence comprehension task, similar to pre-intervention data. However, these effects did not survive correction for multiple comparisons. After intervention there were no differences during the SST task between improvers and non-struggling readers in any of the task-control ROIs. The uncorrected results of these analyses can be found in the Supplementary section 3: ROI analyses results.

3.9. Neuroimaging: Relating change in ROI activation during reading to change in reading skill

We found that change in activity from pre to post intervention during the SC task related to change in reading scores in the VWFA and left posterior fusiform ROIs ($p < .05$). However, these effects did not survive correction for multiple comparisons. More details and the uncorrected results of these analyses can be found in Supplementary section 3: ROI analyses results.

4. Discussion

Our study addressed three main questions. First, were there brain activation patterns *before* intervention characteristic of struggling readers who go on to substantially improve their reading? We found differential activation for sentence comprehension in right fusiform for those who go on to improve compared to those who do not go on to improve as much, and greater depression of a default mode network region (precuneus) compared to non-struggling readers. Second, after intervention, were there any regions where brain activation related to the degree of improvement in struggling readers? We found that activity in the right fusiform continued to distinguish for struggling readers who had larger gains in reading skill after remediation. Third, do more general executive function processes explain any differences between these reader groups? We found that right middle temporal gyrus activity, a putative area of the fronto-parietal executive control network, was suppressed in future non-improvers during sentence comprehension. We also found some marginal ROI

results in executive function regions from the literature that differed in improvers before and after intervention; these control-related group differences were not seen in the control-demanding but non-lexical inhibition task.

Our findings reveal not only differences related to skill change at a neural level both before and after intervention, but also suggest that these changes are somewhat reading-specific and not indicative of alterations in general task control processing. If engagement of executive control is related to propensity for change in reading skill, it may be specifically in the service of reading, and not in a broad, domain general capacity.

Before intervention, neural activation during the Sentence Comprehension task related to future gains

Future improvers, relative to future non-improvers, showed positive activation in the right fusiform and lingual gyrus, while non-improvers showed less activation relative to baseline in these regions. Previous studies of reading intervention have also found that activity in right fusiform predicted later response to intervention (Hoeft et al., 2007b). There is extensive evidence that struggling readers more generally show under-activation in occipital and temporal regions during reading tasks in the left (Maisog et al., 2008; Meyler et al., 2008; Pugh et al., 2001; Richlan, Kronbichler, & Wimmer, 2011; Shaywitz et al., 1998; Temple et al., 2001) and right hemisphere (Hoeft et al., 2006; Maisog et al., 2008; Meyler et al., 2008; Rimrodt et al., 2009; Rumsey et al., 1997). Here, we find that struggling readers who go on to improve are having larger activations in the right hemisphere ventral cortex relative to future non-improvers, extending previous work and suggesting these deactivations are counterproductive to modifying the reading process.

We found future non-improvers had decreased activation in right posterior middle temporal cortex, a region functionally correlated with the fronto-parietal network at rest (Gordon et al., 2016; Yarkoni et al., 2011). The fronto-parietal executive control brain network was originally identified through adult meta-analyses and resting-state analyses (Dosenbach et al., 2007a; Power et al., 2011) but also shows engagement in children during executive control demanding tasks (Church, Bunge, Petersen, & Schlaggar, 2017) and is thought to support rapid control adjustments (Dosenbach, Fair, Cohen, Schlaggar, & Petersen, 2008; Dosenbach et al., 2007a). This region is also commonly implicated in reading studies (Gebauer et al., 2012; Simos et al., 2007). The posterior-temporal cortex may serve as a critical juncture linking reading and executive control. In comparison to adult readers, children learning to read have shown less engagement of dorsolateral frontal and superior parietal regions that also appear to overlap with the fronto-parietal control network (Brown et al., 2005; Schlaggar & Church, 2009). Our findings are in line with the idea that skilled, fluent reading involves engagement of executive control processes as greater deactivation in a node of the executive control network was related to smaller gains and less response to reading remediation over a school year.

Future improvers had larger decreases in precuneus activity when compared to non-struggling readers during sentence comprehension. Deactivation of this part of the posterior midline is consistent with its membership in the default mode network, which is a set of regions that typically decrease in activation during most fMRI tasks relative to fixation

baseline (Fransson, 2005; Raichle et al., 2001; Raichle & Snyder, 2007). Furthermore, when engaging in a task, deactivation of the posterior cingulate is proportional to the difficulty of the task and the amount of attentional resources recruited for the task (Leech, Kamourieh, Beckmann, & Sharp, 2011; McKiernan, Kaufman, Kucera-Thompson, & Binder, 2003; Pearson, Heilbronner, Barack, Hayden, & Platt, 2011). Our findings suggest that greater suppression of the precuneus during a reading task, potentially a marker of greater attentional focus, relates to better future gains in reading ability amongst struggling readers.

After instruction, individual differences in growth related to activation in right fusiform

When we examined our group of struggling readers after intervention, we found that change in the composite reading score from pre-intervention to post-intervention was related to brain activity in the right fusiform. We also saw activity differences in right fusiform comparing future improvers to future non-improvers, as discussed earlier. Brain activity in this region has been shown to change after intervention (Simos et al., 2007) and the region is posited as facilitative for the reading process in early readers and struggling readers (Hoef et al., 2007b; Simos et al., 2007). Our current findings support and expand this view as activity in the right fusiform both before and after intervention in our study was found to relate to gains in reading skill.

Our data are uniquely positioned to test both before and after intervention, and to use continuous measures of growth over time. Others have related measures of phonology and word reading to brain task-related activation (Bach, Richardson, Brandeis, Martin, & Brem, 2013; Hoef et al., 2007b), functional connectivity (Horowitz-Kraus, DiFrancesco, Kay, Wang, & Holland, 2015a), and structural connectivity (Saygin et al., 2013). These neural measures however, were either collected only from struggling readers after intervention (Bach et al., 2013; Horowitz-Kraus et al., 2015a) or from children who did not undergo an intervention (Saygin et al., 2013). The current study benefited from our ability to measure skill change within struggling readers, relating neural data to out-of-scanner measures of reading ability after reading instruction. By doing so, we found a consistent result after instruction to what we see prior to instruction distinguishing future improvers from non-improvers. Further, within the small sample of children for which we had both pre-scan and post-scan quality data, the only significant (but uncorrected) results for testing change in BOLD to change in reading ability were in the left hemisphere, also along the fusiform (*See Supplementary section 3*). Our work adds to a large body of literature identifying the fusiform as a critical region for reading (Price, 2012). The current findings expand the literature by suggesting that bilateral engagement of the fusiform region in struggling readers is correlated both to potential for change, and subsequent changes in reading skill.

Reader group brain differences were specific to reading

Studies in the educational and behavioral realm have established that reading comprehension is related to executive control for attention focus, information selection, and successful recall (Borella et al., 2010; Cirino et al., in press; Cirino et al., *under review*; Locascio et al., 2010). The relation between executive control and reading comprehension at the neural level is still largely unexplored and remains an important avenue in understanding atypical reading development and remediation. One important question is whether any executive

control differences found between reader groups relate specifically to reading or reflect more general, widespread executive control group differences. So far, studies relating executive control and reading brain activity have only focused on neural data collected while performing a reading-related task (Hoefl et al., 2011) or at rest (Aboud et al., 2018; Horowitz-Kraus et al., 2015a). The current sample uniquely includes both an in-scanner reading comprehension and a non-lexical inhibition task from the same group of struggling readers (also reported in (Roe et al., In press)). This dataset can begin to directly address the specific vs. general nature of an executive control differences observed between reader groups. During the reading comprehension task, we found whole brain differences in non-reading network regions such as the precuneus and right MTG (regions that are part of the default mode and fronto-parietal networks, respectively) that distinguished improvers and non-improvers from non-struggling readers before intervention. Critically, these group differences were limited to the reading task and we found no differences between any reader groups during the non-lexical inhibition task, despite strong engagement of executive control regions across all individuals.

There is little evidence that training programs focused on EF promote reading outcomes (Jacob & Parkinson, 2015) - but few studies have even evaluated this question. Where EF has been evaluated as a cognitive training method, it has typically been done via working memory training, seeking to boost the underlying cognitive skill (memory) rather than the academic outcome. Given the purpose of many such studies, academic outcomes are not always measured, though where they are, EF training studies show little transfer to academic skill (Melby-Lervåg & Hulme, 2013). Our results are consistent with this growing set of results, as our lack of results in our general inhibition task tentatively suggest that struggling readers do not have a widespread problem with executive control. We found moderate differences between reader groups in executive control ROIs (e.g., left inferior parietal lobe and a right frontal region) during the reading task, but we did not see these results during the stop-signal task. We suggest that our data indicate that the (perhaps subtle) pattern of greater engagement of executive control networks in readers who go on to improve, or who have improved, are largely specific to the reading process itself. Importantly, the same set of task control ROIs were engaged by all participants during the inhibition task, though there were no differences in control engagement across reader groups. Thus, general executive control activity alone was not a marker of future remediation response. Our results suggest that any reading intervention that seeks to add an executive control-related component may be most beneficial if the executive control aspects are tailored to the reading process. These results indicate that executive control engagement may be tuned specifically for different skill sets, but this idea is in need of further research.

Methods of defining improvement impacted study outcomes

We defined improvement based on a struggling reader's individual growth on outside-of-scanner reading measures, regardless of the group's average scores. As this is a relatively novel approach, we post-hoc examined two other methods of defining improvement. Defining improvement based on a cut score (> 90 standard score on the TOWRE-2) or based on a median group split (of our post-intervention composite score), replicated our struggling reader improver group differences relative to non-struggling readers. However, these group

definitions were not sensitive to the fusiform results found in our individual gain model comparison between struggling reader subgroups. We believe this is evidence that the definition of improvement is a critical decision that shapes study outcomes. Further, we suggest that defining improvement at the individual level based on growth, as opposed to relative group-level outcomes, may reveal remediation effects that other methods miss. We prioritized the size of change in an individual over time, and also required large gains on at least two standardized measures to minimize single measurement noise or error. Our current study confirms that group membership decisions about intervention response can have a significant impact on the study results.

The young, but growing subfield of neuroimaging research that assesses educational interventions (Aboud et al., 2018; Huber, Donnelly, Rokem, & Yeatman, 2018) is complicated by variable definitions of how to measure improvement or growth, mainly using behavioral tests alone (Aboud et al., 2018; Hoeft et al., 2007b; Simos et al.). Information gathering about different definitional approaches, such as those in this paper, could help the field come to a greater consensus about how to define improvement. Brain data also provide an opportunity to characterize remediation response and achievement beyond observable behavior (Colvin, 2016). The hope is that at some point, neural measures can inform the design and implementation of more individualized interventions (Tandon & Singh, 2016).

Limitations

While this study contributes novel findings to the intervention literature, there are some limitations. It is difficult to recruit participants for neuroimaging studies from schools that are part of intervention studies. Many parents were not willing to travel to the imaging site and even less likely to do so more than once. By splitting the struggling readers into two groups at each scan, our group sizes were small, yielding only mild statistical differences. Thus, our results should be interpreted with caution, and only as first steps to answering these complex questions relating brain changes to academic skill changes. Additionally, it is important to consider differences between struggling readers who have both fluency and comprehension deficits and those who only have comprehension deficits, which is a distinction that is beyond the scope of this paper. In future studies, greater numbers of both pre- and post-intervention scans from the same individuals will boost power and allow for more in-depth comparisons. Lastly, our stop-signal task only targets one facet of executive function (inhibition) and employed a staircasing procedure which impacts task difficulty. There is still much research that remains to understand how different facets of executive function and different types of control-demanding tasks relate to reading skill.

Conclusions

Engagement of right fusiform, as well as posterior middle temporal cortex and precuneus during a reading task in struggling readers was related to reading improvement over the course of reading remediation in late elementary students. These findings were reading-specific and not present in a control-demanding non-lexical task. Continued examination of neural profiles for skill change will generate a richer context for not only identifying struggling readers, but also for creating more tailored and successful reading interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We would like to thank the Imaging Research Center at UT Austin, and Jeanette Mumford for experimental design and statistical assistance. We would also like to thank the UT Meadows Center for Prevention of Educational Risk, especially Gregory Roberts, Garrett J Roberts, and Stephanie Stillman, and Jeremy Miciak at the University of Houston for their critical roles in the intervention and assessment arms of the TCLD project. We thank Annie Zheng, Leonel Olmedo, Joel Martinez, Lauren Deschner, and Laura Engelhardt for help with data collection. The authors declare no conflicts of interest.

FUNDING

This research was supported by grant P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). The opinions expressed in the paper do not necessarily reflect the opinions of the NICHD or the National Institutes of Health.

References

- About KS, Barquero LA, & Cutting LE (2018). Prefrontal mediation of the reading network predicts intervention response in dyslexia. *Cortex*, 101, 96–106. doi:10.1016/j.cortex.2018.01.009 [PubMed: 29459284]
- Al Otaiba S, & Fuchs D (2002). Characteristics of children who are unresponsive to early literacy intervention. *Remedial and Special Education*, 23(5), 300–316. doi: 10.1177/0741932502030050501
- Arrington CN, Kulesz PA, Francis DJ, Fletcher JM, & Barnes MA (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading*, 18(5), 325–346. doi:10.1080/10888438.2014.902461
- Aylward EH, Richards TL, Berninger VW, Nagy WE, Field KM, Grimme AC, ... Cramer SC (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology*, 61(2), 212–219. doi:10.1212/01.wnl.0000068363.05974.64 [PubMed: 12874401]
- Bach S, Richardson U, Brandeis D, Martin E, & Brem S (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *NeuroImage*, 82, 605–615. doi:10.1016/j.neuroimage.2013.05.062 [PubMed: 23727320]
- Barquero LA, Davis N, & Cutting LE (2014). Neuroimaging of reading intervention: a systematic review and activation likelihood estimate meta-analysis. *PLOS ONE*, 9(1), e83668. doi:10.1371/journal.pone.0083668 [PubMed: 24427278]
- Beckmann CF, Jenkinson M, & Smith SM (2003). General multi-level linear modelling for group analysis in FMRI. *NeuroImage*, 20(2), 1052–1063. doi:10.1016/S1053-8119(03)00435-X [PubMed: 14568475]
- Berninger V, Abbott R, Cook CR, & Nagy W (2017). Relationships of attention and executive functions to oral language, reading, and writing skills and systems in middle childhood and early adolescence. *Journal of Learning Disabilities*, 50(4), 434–449. doi:10.1177/0022219415617167 [PubMed: 26746315]
- Best J, Miller P, & Naglieri J (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, 21, 327–336. [PubMed: 21845021]
- Borella E, Carretti B, & Pelegrina S (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, 43(6), 541–552. doi: 10.1177/0022219410371676 [PubMed: 20606207]
- Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, & Schlaggar BL (2005). Developmental changes in human cerebral functional organization for word generation. *Cerebral cortex* (New York, NY : 1991), 15(3), 275–290. doi:10.1093/cercor/bhh129

- Cho E, Roberts GJ, Capin P, Roberts G, Miciak J, & Vaughn S (2015). Cognitive attributes, attention, and self-efficacy of adequate and inadequate responders in a fourth grade reading intervention. *Learning Disabilities Research & Practice*, 30(4), 159–170. doi:10.1111/ldrp.12088 [PubMed: 26997755]
- Church JA, Bunge SA, Petersen SE, & Schlaggar BL (2017). Preparatory engagement of cognitive control networks increases late in childhood. *Cerebral Cortex*, 27(3), 2139–2153. doi:10.1093/cercor/bhw046 [PubMed: 26972753]
- Cirino PT, Ahmed Y, Miciak J, Taylor WP, Gerst EH, & Barnes MA (in press). A framework for executive function in the late elementary years. *Neuropsychology*
- Cirino PT, Miciak J, Ahmed Y, Barnes MA, Taylor WP, & Gerst EH (under review). Executive function: Association with multiple reading skills
- Cohen L, & Dehaene S (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage*, 22(1), 466–476. doi:10.1016/j.neuroimage.2003.12.049 [PubMed: 15110040]
- Colvin R (2016). Optimising, generalising and integrating educational practice using neuroscience. *npj Science of Learning*, 1(1). doi:10.1038/npjscilearn.2016.12
- Congdon E, Mumford J, Cohen JR, Galvan A, Canli T, & Poldrack RA (2012). Measurement and reliability of response inhibition. *Frontiers in Psychology*, 3(37), 1–10. doi:10.3389/fpsyg.2012.00037 [PubMed: 22279440]
- Davis N, Barquero L, Compton DL, Fuchs LS, Fuchs D, Gore JC, & Anderson AW (2011). Functional correlates of children's responsiveness to intervention. *Developmental Neuropsychology*, 36(3), 288–301. doi:10.1080/87565641.2010.549875 [PubMed: 21462008]
- de Jong CG, Van De Voorde S, Roeyers H, Raymaekers R, Oosterlaan J, & Sergeant J (2009). How distinctive are ADHD and RD? Results of a double dissociation study. *Journal of Child Psychology and Psychiatry*, 50(10), 1007–1017. doi:10.1007/s10802-009-9328-y
- Dhar M, Been PH, Minderaa RB, & Althaus M (2008). Distinct information processing characteristics in dyslexia and ADHD during a covert orienting task: an event-related potential study. *Clinical Neurophysiology*, 119(9), 2011–2025. doi:10.1016/j.clinph.2008.05.027 [PubMed: 18621578]
- Diamond A (2013). Executive functions. *Annual Review of Psychology*, 64(1), 135–168. doi:10.1146/annurev-psych-113011-143750
- Dosenbach NUF, Fair DA, Cohen AL, Schlaggar BL, & Petersen SE (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences*, 12(3), 99–105. doi:10.1016/j.tics.2008.01.001 [PubMed: 18262825]
- Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, ... Petersen SE (2007a). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 104(26), 11073–11078. doi:10.1073/pnas.0704320104 [PubMed: 17576922]
- Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, ... Petersen SE (2007b). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, 104(26), 11073–11078. doi:10.1073/pnas.0704320104
- Eklund A, Nichols T, & Knutsson H (2016). Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, 113(33), E4929. doi:10.1073/pnas.1612033113 [PubMed: 27503879]
- Engelhardt LE, Church JA, Harden KP, & Tucker-Drob EM (In Press). Accounting for the shared environment in cognitive abilities and academic achievement with measured socioecological contexts
- Engelhardt LE, Roe MA, Juranek J, DeMaster D, Harden KP, Tucker-Drob EM, & Church JA (2017). Children's head motion during fMRI tasks is heritable and stable over time. *Developmental Cognitive Neuroscience*, 25, 58–68. doi:10.1016/j.dcn.2017.01.011 [PubMed: 28223034]
- Farris EA, Odegard TN, Miller HL, Ring J, Allen G, & Black J (2011). Functional connectivity between the left and right inferior frontal lobes in a small sample of children with and without reading difficulties. *Neurocase*, 17(5), 425–439. doi:10.1080/13554794.2010.532141 [PubMed: 21590585]

- Fletcher JM, Lyon GR, Fuchs LS, & Barnes MA (In Press). Learning disabilities: From identification to intervention New York: Guilford Press.
- Fletcher JM, Stuebing KK, Batrth AE, Denton CA, Cirino PT, Francis DJ, & Vaughn S (2011). Cognitive correlates of inadequate response to reading intervention. *School Psychology Review*, 40(1), 3–22. [PubMed: 23125475]
- Fox MD, & Raichle ME (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9), 700–711. doi:10.1038/nrn2201 [PubMed: 17704812]
- Fransson P (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping*, 26(1), 15–29. doi:10.1002/hbm.20113 [PubMed: 15852468]
- Frijters JC, Lovett MW, Sevcik RA, & Morris RD (2012). Four methods of identifying change in the context of a multiple component reading intervention for struggling middle school readers. *Reading and Writing*, 26(4), 539–563. doi:10.1007/s11145-012-9418-z
- Gebauer D, Fink A, Kargl R, Reishofer G, Koschutnig K, Purgstaller C, ... Enzinger C (2012). Differences in brain function and changes with intervention in children with poor spelling and reading abilities. *PloS One*, 7(5), e38201. doi:10.1371/journal.pone.0038201 [PubMed: 22693600]
- Germano E, Gagliano A, & Curatolo P (2010). Comorbidity of ADHD and Dyslexia. *Developmental Neuropsychology*, 35(5), 475–493. doi:10.1080/875656412010494748 [PubMed: 20721770]
- Gooch D, Snowling M, & Hulme C (2011). Time perception, phonological skills and executive function in children with dyslexia and/or ADHD symptoms. *Journal of child psychology and psychiatry*, 52(2), 195–203. doi:10.1111/j.1469-7610.2010.02312.x [PubMed: 20860755]
- Gordon EM, Laumann TO, Adeyemo B, Gilmore AW, Nelson SM, Dosenbach NUF, & Petersen SE (2016). Individual-specific features of brain systems identified with resting state functional correlations. *Neuroimage*, 1–22. doi:10.1016/j.neuroimage.2016.08.032
- Greene DJ, Laumann TO, Dubis JW, Ihnen SK, Neta M, Power JD, ... Schlaggar L (2014). Developmental changes in the organization of functional connections between the basal ganglia and cerebral cortex. *Journal of Neuroscience*, 34(17), 5842–5854. doi:10.1523/JNEUROSCI.3069-13.2014 [PubMed: 24760844]
- Greve DN, & Fischl B (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1), 63–72. doi:10.1016/j.neuroimage.2009.06.060 [PubMed: 19573611]
- Hoeft F, Hernandez A, McMillon G, Taylor-Hill H, Martindale JL, Meyler A, ... Gabrieli JDE (2006). Neural basis of dyslexia: a comparison between dyslexic and nondyslexic children equated for reading ability. *The Journal of Neuroscience*, 26(42), 10700–10708. doi:10.1523/JNEUROSCI.4931-05.2006 [PubMed: 17050709]
- Hoeft F, McCandliss BD, Black JM, Gantman A, Zakerani N, Hulme C, ... Gabrieli JDE (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), 361–366. doi:10.1073/pnas.1008950108 [PubMed: 21173250]
- Hoeft F, Meyler A, Hernandez A, Juel C, Taylor-Hill H, Martindale JL, ... Gabrieli JDE (2007a). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences of the United States of America*, 104(10), 4234–4239. doi: 10.1073/pnas.0609399104 [PubMed: 17360506]
- Hoeft F, Ueno T, Reiss AL, Meyler A, Whitfield-Gabrieli S, Glover GH, ... Gabrieli JDE (2007b). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience*, 121(3), 602–613. doi:10.1037/0735-7044.121.3.602 [PubMed: 17592952]
- Horowitz-Kraus T, DiFrancesco M, Kay B, Wang Y, & Holland SK (2015a). Increased resting-state functional connectivity of visual- and cognitive-control brain networks after training in children with reading difficulties. *Neuroimage: Clinical*, 8, 619–630. doi:10.1016/j.nicl.2015.06.010 [PubMed: 26199874]
- Horowitz-Kraus T, Toro-Serey C, & DiFrancesco M (2015b). Increased Resting-State Functional Connectivity in the Cingulo-Opercular Cognitive-Control Network after Intervention in Children

with Reading Difficulties. *PloS One*, 10(7), e0133762. doi:10.1371/journal.pone.0133762.s001 [PubMed: 26197049]

Horowitz-Kraus T, Vannest J, Gozdasand E, & Holland S (2014). Greater utilization of neural-circuits related to executive functions is associated with better reading: a longitudinal fMRI study using the verb generation task. *Frontiers in Human Neuroscience*, 8. doi:10.3389/fnhum.2014.00447

Huber E, Donnelly PM, Rokem A, & Yeatman JD (2018). Rapid and widespread white matter plasticity during an intensive reading intervention. *Nat Commun*, 9(1), 2260. doi:10.1038/s41467-018-04627-5 [PubMed: 29884784]

Jacob R, & Parkinson J (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Review of Educational Research*, 85(4), 512–552. doi:10.3102/0034654314561338

Jenkinson M, Bannister P, Brady M, & Smith S (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841. doi:10.1006/nimg.2002.1132 [PubMed: 12377157]

Jenkinson M, & Smith S (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2), 143–156. doi:10.1016/S1361-8415(01)00036-6 [PubMed: 11516708]

Koyama MS, Di Martino A, Kelly C, Jutagir DR, Sunshine J, Schwartz S, ... Milham M (2013). Cortical signatures of dyslexia and remediation: an intrinsic functional connectivity approach. *PloS One*, 8(2). doi:10.1371/journal.pone.0055454

Leech R, Kamourieh S, Beckmann CF, & Sharp DJ (2011). Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *The Journal of Neuroscience*, 31(9), 3217–3224. doi:10.1523/JNEUROSCI.5626-10.2011 [PubMed: 21368033]

Locascio G, Mahone EM, Eason SH, & Cutting LE (2010). Executive dysfunction among children with reading comprehension deficits. *Journal of Learning Disabilities*, 43(5), 441–454. doi:10.1177/0022219409355476 [PubMed: 20375294]

MacGinitie WH, MacGinitie RK, Maria K, & Dreyer LG (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside.

Mahone EM, Cirino PT, Cutting LE, Cerrone PM, Hagelthorn KM, Hiemenz JR, ... Denckla MB (2002). Validity of the behavior rating inventory of executive function in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, 17(7), 643–662. Retrieved from http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6VDJ-46WM6S8-7&_user=4275&_coverDate=10%2F31%2F2002&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000000938&_version=1&_urlVersion=0&_userid=4275&md5=a624bc2863e5ff984240cale711ad555 [PubMed: 14591848]

Maisog JM, Einbinder ER, Flowers DL, Turkeltaub PE, & Eden GF (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences*, 1145(1), 237–259. doi:10.1196/annals.1416.024 [PubMed: 19076401]

McKiernan KA, Kaufman JN, Kucera-Thompson J, & Binder JR (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of cognitive neuroscience*, 15(3), 394–408. doi:10.1162/089892903321593117 [PubMed: 12729491]

Melby-Lervåg M, & Hulme C (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270–291. doi:10.1037/a0028228 [PubMed: 22612437]

Meyler A, Keller TA, Cherkassky VL, Gabrieli JD, & Just MA (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. doi:10.1016/j.neuropsychologia.2008.03.012 [PubMed: 18495180]

Meyler A, Keller TA, Cherkassky VL, Lee D, Hoeft F, Whitfield-Gabrieli S, ... Just MA (2007). Brain activation during sentence comprehension among good and poor readers. *Cerebral Cortex*, 17(12), 2780–2787. doi:10.1093/cercor/bhm006 [PubMed: 17317678]

Miciak J, Stuebing KK, Vaughn S, Roberts G, Barth AE, & Fletcher JM (2014). Cognitive attributes of adequate and inadequate responders to reading intervention in middle school. *School Psychology Review*, 43(4), 407–427. doi:10.17105/SPR-13-0052.1 [PubMed: 28579668]

- Miciak J, Taylor WP, Stuebing KK, Fletcher JM, & Vaughn S (2016). Designing Intervention Studies: Selected Populations, Range Restrictions, and Statistical Power. *Journal of Research on Educational Effectiveness*, 9(4), 556–569. doi:10.1080/19345747.2015.1086916 [PubMed: 28479943]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. doi:10.1006/cogp.1999.0734 [PubMed: 10945922]
- Odegard TN, Ring J, Smith S, Biggan J, & Black J (2008a). Differentiating the neural response to intervention in children with developmental dyslexia. *Annals of Dyslexia*, 58(1), 1–14. doi: 10.1007/s11881-008-0014-5 [PubMed: 18483867]
- Odegard TN, Ring J, Smith S, Biggan J, & Black J (2008b). Differentiating the neural response to intervention in children with developmental dyslexia. *Annals of Dyslexia*, 58(1), 1–14. doi: 10.1007/s11881-008-0014-5 [PubMed: 18483867]
- Pearson JM, Heilbronner SR, Barack DL, Hayden BY, & Platt ML (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends in Cognitive Sciences*, 15(4), 143–151. doi: 10.1016/j.tics.2011.02.002 [PubMed: 21420893]
- Peirce JW (2007). PsychoPy--Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13. doi:10.1016/j.jneumeth.2006.11.017 [PubMed: 17254636]
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, ... Petersen SE (2011). Functional network organization of the human brain. *Neuron*, 72(4), 665–678. doi:10.1016/j.neuron.2011.09.006 [PubMed: 22099467]
- Power JD, & Petersen SE (2013). Control-related systems in the human brain. *Current Opinion in Neurobiology*, 23(2), 223–228. doi:10.1016/j.conb.2012.12.009 [PubMed: 23347645]
- Power JD, Schlaggar BL, & Petersen SE (2015). Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage*, 105, 536–551. doi:10.1016/j.neuroimage.2014.10.044 [PubMed: 25462692]
- Price C (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. doi:10.1016/j.neuroimage.2012.04.062 [PubMed: 22584224]
- Pugh KR, Mencl WE, Jenner AR, Katz L, Frost SJ, Lee JR, ... Shaywitz BA (2001). Neurobiological studies of reading and reading disability. *Journal of communication disorders*, 34(6), 479–492. doi: 10.1016/S0021-9924(01)00060-0 [PubMed: 11725860]
- R Development Core Team. (2014). R: A language and environment for statistical computing Vienna Austria: R Foundation for Statistical Computing Retrieved from <http://www.R-project.org>
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, & Shulman GL (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), 676–682. doi:10.1073/pnas.98.2.676 [PubMed: 11209064]
- Raichle ME, & Snyder AZ (2007). A default mode of brain function: A brief history of an evolving idea. *NeuroImage*, 37(4), 1083–1090. doi:10.1016/j.neuroimage.2007.02.041 [PubMed: 17719799]
- Rao C, & Singh NC (2015). Visuospatial complexity modulates reading in the brain. *Brain and Language*, 141, 50–61. doi:10.1016/j.bandl.2014.11.010
- Reuter M, Schmansky NJ, Rosas HD, & Fischl B (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4), 1402–1418. doi:10.1016/j.neuroimage.2012.02.084 [PubMed: 22430496]
- Rezaie R, Simos PG, Fletcher JM, Cirino PT, Vaughn S, & Papanicolaou AC (2011). Temporo-parietal brain activity as a longitudinal predictor of response to educational interventions among middle school struggling readers. *Journal of the International Neuropsychological Society*, 17(5), 875–885. doi:10.1017/S1355617711000890 [PubMed: 21740612]
- Richards TL, & Berninger VW (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: before but not after treatment. *Journal of Neurolinguistics*, 21(4), 294–304. doi: 10.1016/j.jneuroling.2007.07.002 [PubMed: 19079567]

- Richardson FM, Seghier ML, Leff AP, Thomas MS, & Price CJ (2011). Multiple routes from occipital to temporal cortices during reading. *Journal of Neuroscience*, 31(22), 8239–8247. doi:10.1523/JNEUROSCI.6519-10.2011 [PubMed: 21632945]
- Richlan F (2012). Developmental dyslexia: dysfunction of a left hemisphere reading network. *Frontiers in Human Neuroscience*, 6, 120. doi:10.3389/fnhum.2012.00120 [PubMed: 22557962]
- Richlan F, Kronbichler M, & Wimmer H (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage*, 56(3), 1735–1742. doi:10.1016/j.neuroimage.2011.02.040 [PubMed: 21338695]
- Rimrodt SL, Clements-Stephens AM, Pugh KR, Courtney SM, Gaur P, Pekar JJ, & Cutting LE (2009). Functional MRI of sentence comprehension in children with dyslexia: beyond word recognition. *Cerebral Cortex*, 19(2), 402–413. doi:10.1093/cercor/bhn092 [PubMed: 18515796]
- Roe MA, Martinez JE, Mumford JA, Taylor WP, Cirino PT, Fletcher JM, ... Church JA (In press). Control engagement during sentence and inhibition fMRI tasks in children with reading difficulties
- Romeo RR, Christodoulou JA, Halverson KK, Murtagh J, Cyr AB, Schimmel C, ... Gabrieli JDE (2017). Socioeconomic status and reading disability: neuroanatomy and plasticity in response to intervention. *Cerebral Cortex*, 1–16. doi:10.1093/cercor/bhx131 [PubMed: 28365777]
- Rubia K, Smith AB, Brammer MJ, & Taylor E (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *NeuroImage*, 20(1), 351–358. doi:10.1016/S1053-8119(03)00275-1 [PubMed: 14527595]
- Rumsey JM, Horwitz B, Donohue BC, Nace K, Maisog JM, & Andreason P (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain : a journal of neurology*, 120 (Pt 5), 739–759. [PubMed: 9183247]
- Saygin ZM, Norton ES, Osher DE, Beach SD, Cyr AB, Ozernov-Palchik O, ... Gabrieli JDE (2013). Tracking the roots of reading ability: white matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children. *The Journal of Neuroscience*, 33(33), 13251–13258. doi:10.1523/JNEUROSCI.4383-12.2013 [PubMed: 23946384]
- Schlaggar BL, & Church JA (2009). Functional Neuroimaging Insights Into the Development of Skilled Reading. *Current Directions in Psychological Science*, 18(1), 21–26. doi:10.1111/j.1467-8721.2009.01599.x [PubMed: 19750204]
- Shaywitz BA, Shaywitz SE, Blachman BA, Pugh KR, Fulbright RK, Skudlarski P, ... Gore JC (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, 55(9), 926–933. doi:10.1016/j.biopsych.2003.12.019 [PubMed: 15110736]
- Shaywitz SE, Shaywitz BA, Pugh KR, Fulbright RK, Constable RT, Mencl WE, ... Gore JC (1998). Functional disruption in the organization of the brain for reading in dyslexia Paper presented at the Proceedings of the National Academy of Sciences of the United States of America.
- Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, & Petersen SE (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Human Brain Mapping*, 35(5), 1981–1996. doi:10.1002/hbm.22307 [PubMed: 23861343]
- Simos PG, Breier JJ, Fletcher JM, Foorman BR, Castillo EM, & Papanicolaou AC (2002). Brain mechanisms for reading words and pseudowords: an integrated approach. *Cerebral Cortex*, 12(3), 297–305. doi:10.1093/cercor/12.3.297 [PubMed: 11839603]
- Simos PG, Fletcher JM, Sarkari S, Billingsley RL, Denton C, & Papanicolaou AC (2007). Altering the Brain Circuits for Reading Through Intervention: A Magnetic Source Imaging Study. *Neuropsychology*, 21(4), 485–496. doi:10.1037/0894-4105.21.4.485 [PubMed: 17605581]
- Smith SM (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155. doi:10.1002/hbm.10062 [PubMed: 12391568]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, ... Matthews PM (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1, S208–219. doi:10.1016/j.neuroimage.2004.07.051 [PubMed: 15501092]
- Solis M, Scammacca NK, Barth AE, & Roberts G (2017). Text-based Vocabulary Intervention Training Study: Supporting Fourth Graders with Low Reading Comprehension and Learning

- Disabilities. *Learning Disabilities*, 15(1), 103–115. doi:10.1007/s11145-012-9418-z [PubMed: 29093765]
- St. Clair-Thompson HL, & Gathercole SE (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, 59(4), 745–759. [PubMed: 16707360]
- Stuebing KK, Barth AE, Trahan LH, Reddy RR, Miciak J, & Fletcher JM (2015). Are child cognitive characteristics strong predictors of responses to intervention? a meta-analysis. *Review of Educational Research*, 85(3), 395–429. doi:10.3102/0034654314555996 [PubMed: 26535015]
- Tandon PN, & Singh NC (2016). Educational Neuroscience: Challenges and Opportunities. *Ann Neurosci*, 23(2), 63–65. doi:10.1159/000443560. [PubMed: 27647954]
- Temple E, Deutsch GK, Poldrack RA, Miller SL, Tallal P, Merzenich MM, & Gabrieli JDE (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 100(5), 2860–2865. doi:10.1073/pnas.0030098100 [PubMed: 12604786]
- Temple E, Poldrack RA, Salidis J, Deutsch GK, Tallal P, Merzenich MM, & Gabrieli JDE (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: an fMRI study. *Neuroreport*, 12(2), 299–307. Retrieved from <http://meta.wkhealth.com/pt/pt-core/template-journal/lwwgateway/media/landingpage.htm?issn=0959-4965&volume=12&issue=2&spage=299>, <http://graphics.tx.ovid.com/ovftpdfs/FPDDNCIBDDJDIA00/fs004/ovft/live/gv006/00001756/00001756-200102120-00024.pdf> [PubMed: 11209939]
- Torgesen JK, Wagner RK, & Rashotte CA (1999). *Test of word reading efficiency (TOWRE)* Austin, TX: Pro-Ed.
- Van Essen DC (2012). Cortical cartography and Caret software. *NeuroImage*, 62(2), 757–764. doi: 10.1016/j.neuroimage.2011.10.077 [PubMed: 22062192]
- Vaughn S, Solis M, Miciak J, Taylor WP, & Fletcher JM (2016). Effects from a randomized control trial comparing researcher and school-implemented treatments with fourth graders with significant reading difficulties. *Journal of Research on Educational Effectiveness*, 9(Suppl 1), 23–44. doi:10.1080/19345747.2015.1126386 [PubMed: 28491206]
- Vellutino FR, Scanlon DM, & Fanuele DP (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: evidence for the role of kindergarten and first-grade interventions. *Journal of learning disabilities*, 39(2), 157–169. doi: 10.1177/00222194060390020401 [PubMed: 16583795]
- Vogel AC, Church JA, Power JD, Miezin FM, Petersen SE, & Schlaggar BL (2013). Functional network architecture of reading-related regions across development. *Brain and Language*, 125(2), 231–243. doi:10.1016/j.bandl.2012.12.016 [PubMed: 23506969]
- Wagner RK, Torgesen JK, Rashotte CA, & Pearson NA (2010). *Test of Silent Reading Efficiency and Comprehension (TOSREC)* examiner's manual Austin, TX: Pro-Ed.
- Woodcock RW, McGrew KS, & Mather N (2001). *Woodcock-Johnson III Tests of Achievement* Itasca, IL: Riverside.
- Woolrich MW (2008). Robust group analysis using outlier inference. *Neuroimage*, 41(2), 286–301. doi:10.1016/j.neuroimage.2008.02.042 [PubMed: 18407525]
- Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, & Smith SM (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, 21(4), 1732–1747. doi:10.1016/j.neuroimage.2003.12.023 [PubMed: 15050594]
- Woolrich MW, Ripley BD, Brady M, & Smith SM (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, 14(6), 1370–1386. doi:10.1006/nimg.2001.0931 [PubMed: 11707093]
- Worsley KJ (2001). Statistical analysis of activation images. In Jezzard P, Matthews P, & Smith S (Eds.), *Functional MRI: An introduction to methods* (1st ed., Vol. 14, pp. 251–270). Oxford (GB): Oxford Medical Publications.
- Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, & Wager TD (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8(8), 665–670. doi:10.1038/nmeth.1635 [PubMed: 21706013]

Figure 1.

a. Mean composite reading score at pre-instruction and post-instruction for all unique SC participants, sorted by improver label (gains in 2 or more reading scores); improvers had significantly higher composite scores than non-improvers after instruction ($t(48.39) = -2.078, p = .04$); NS = Non-struggling. Error bars reflect standard error of the mean. b. SC task adaptation and timing from Meyler and colleagues (Meyler et al., 2007). c. SST task adaptation from de Jong and colleagues (de Jong et al., 2009). Task figures adapted from (Roe et al., In press). Relative size and placement of text altered from experiment for easier viewing.

Figure 2.

Scan 1 (fall, pre-instruction) whole brain SC group differences. a. Future improvers vs. future non-improvers, future improvers had more activation in right fusiform and lingual gyrus; b. Future improvers vs. non-struggling readers, future non-improvers had reduced activity compared to non-struggling readers in right posterior middle temporal gyrus and a left post central / white matter regions; c. Future non-improvers vs. non-struggling readers, future improvers have reduced activity compared to non-struggling readers in the precuneus. Brains are corrected at $z > 3.1$ with a cluster probability of $p < .05$; Peak ROIs were extracted from whole brain contrasts to demonstrate direction of activations only and were not used statistically.

a. Future Improvers vs. Non-struggling**b. Future Non-Improvers vs. Non-struggling** **Figure 3.**

Scan 1 (fall, pre-instruction) overlapping SC results of future improvers/non-improvers vs. non-struggling models reflecting three different methods of defining improvers. a. Future improvers vs. non-struggling readers; b. Future non-improvers vs. non-struggling readers. Brains are corrected for multiple comparisons at $z > 3.1$ with a cluster probability of $p < 0.05$; Individual gain = Our principle method of defining improvement based on above expected gain in 2+ tests; median split = alternative improvement definition 1: a group median split of post intervention composite reading score; > 90 TOWRE = alternative improvement definition 2: improvers defined by scoring 90 or above on TOWRE-2 after intervention (a cut-score approach).

Figure 4.

Scan 2 (spring, post-instruction) whole brain analysis of the SC task correlated with individual gain in composite correlate analysis. Activity in the right fusiform cortex had a positive relationship (38, -50, -16) with change in the composite reading score over time.

Table 1:

Participant characteristics and descriptive statistics for behavioral tests for the Sentence Comprehension (SC) task.

	Non-struggling Readers (<i>N</i> = 28)		Future Improvers Scan 1 (<i>N</i> = 19)		Future Non-improvers Scan 1 (<i>N</i> = 27)		Improvers Scan 2 (<i>N</i> = 15)		Non-improvers Scan 2 (<i>N</i> = 8)	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Gender										
Female	13	46.4	8	42.1	12	44.4	8	53.3	5	62.5
Male	15	53.6	11	57.9	15	55.6	7	46.7	3	37.5
Race										
Black	2	7.1	3	15.7	3	11.1	2	13.3	1	12.5
Hispanic	6	21.4	10	63.1	12	48.1	8	60.0	3	37.5
NH White	19	67.9	4	21.0	11	40.7	4	26.6	3	37.5
Multiracial	1	3.6	2	10.5	1	3.7	1	6.7		-
Native American		-		-		-		-	1	12.5
Free/reduced lunch	2	7.1	10	52.6	14	51.8	7	46.7	6	75.0
ADHD	-	-	2	10.5	3	11.1	1	6.7	-	-
EXP INT	-	-	10	52.6	15	55.5	11	73.3	4	50
	<i>N</i>	<i>M</i> (SD)	<i>N</i>	<i>M</i> (SD)	<i>N</i>	<i>M</i> (SD)	<i>N</i>	<i>M</i> (SD)	<i>N</i>	<i>M</i> (SD)
Age	28	9.8 (.82)	19	10.09 (.55)	27	10.22 (.54)	15	10.60 (.41)	8	10.91 (.42)
K-BIT 2	28	115.64 (11.02)	19	96.16 (8.33)	27	100.92 (10.36)	15	91.13 (10.22)	8	93.12 (12.20)
Verbal	27	111.74 (12.66)	19	91.63 (12.35)	27	97.00 (10.96)	15	90.07 (13.06)	8	90.62 (11.12)
Non-Verbal	28	114.43 (13.96)	19	101.26 (10.68)	27	104.15 (12.42)	15	94.27 (11.94)	8	97.37 (16.90)
COMP	28	107.3 (7.73)	19	90.21 (7.01)	27	91.48 (5.97)	15	93.38(7.62)	8	88.94(7.02)
COMP-DIFF	-	-	19	4.74 (3.55) **	27	0.04 (3.22)	15	5.41 (4.12) **	8	0 (1.77)

NH = Non-Hispanic; EXP INT= individuals in the TCLD experimental intervention (and not business-as-usual struggling controls); KBIT = Kaufman Brief Intelligence Test; COMP = composite reading score derived from the average of four reading scores (see section 2.1.2); COMP-DIFF = difference in composite reading score from pre-intervention to post-intervention *'s denote improver group differences

*
 $p < .05$

**
 $p < .01$.

Across all struggling readers from both time points ($N = 53$), improvers had higher composite reading scores than non-improvers ($t(48.39) = -2.078$ $p = 0.04$, Figure 1a). Welch's t-test adjustment for unequal variances and unequal sample sizes were performed between future improvers/future non-improvers and improvers/non-improvers respectively.

Table 2:

Participant characteristics and descriptive statistics for behavioral tests for the Stop-Signal task (SST).

	Non-struggling Readers (<i>N</i> = 27)		Future Improvers Scan 1 (<i>N</i> = 16)		Future Non-improvers Scan 1 (<i>N</i> = 29)		Improvers Scan 2 (<i>N</i> = 14)		Non-improvers Scan 2 (<i>N</i> = 7)	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Gender										
Female	12	44.4	5	31.2	11	37.9	8	57.1	5	71.4
Male	15	55.6	11	68.8	18	62.1	6	42.9	2	28.6
Race										
Black	2	7.4	1	6.2	3	10.4	1	7.1	1	14.2
Hispanic	6	22.2	9	56.2	13	44.8	8	57.1	3	42.9
NH White	18	66.7	4	25	13	44.8	4	28.6	3	42.9
Multiracial	1	3.7	2	12.5	-	-	1	7.1	-	-
Free/reduced lunch	2	7.4	9	56.2	13	44.8	8	57.1	6	85.7
ADHD	-	-	3	18.7	4	13.8	1	7.1	-	-
EXP INT	-	-	9	56.2	17	58.6	9	64.3	4	57.1
	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)	<i>N</i>	<i>M</i> (<i>SD</i>)
Age	27	9.91 (.8)	16	10.12 (.56)	29	10.30 (.60)	14	10.60 (.42)	7	10.96 (.42)
K-BIT 2	27	116.37 (10.9)	16	97.81 (8)	29	98.72 (10.88)	14	90 (10.08)	7	94.86 (12.08)
Verbal	26	112.58 (12.65)	16	91.69 (12.25)	29	94.70 (10.91)	14	88 (12.04)	7	91.57 (11.66)
Non-Verbal	27	114.85 (13.92)	16	104.12 (11.19)	29	102.72 (12.95)	14	94.43 (12.28)	7	99.43 (17.15)
COMP	27	108.43 (8.45)	16	90.59 (7.53)	29	90.39 (6.5)	14	92.54(7.26)	7	88.5(7.47)
COMP-DIFF	-	-	16	4.77 (3.75) **	29	0.12 (3.07)	14	4.90 (3.98) **	7	0.43 (1.39)

NH = Non-Hispanic; EXP INT = individuals in the TCLD experimental intervention; KBIT = Kaufman Brief Intelligence Test; COMP = composite reading score; COMP-DIFF = difference in composite reading score from pre-intervention to post-intervention *'s denote improver group differences

*
 $p < .05$

**
 $p < .01$.

Welch's t-test adjustment for unequal variances and unequal sample sizes performed between future improvers/future non-improvers and improvers/non-improvers respectively.

Table 3.

Number of improvers / non-improvers in each unique improver definition grouping.

Improver Definition	Individual gain	Median split	TOWRE-2
Improvers SC	19	23	23
Non-improvers SC	27	23	23
% participants shared with gain definition	-	0.61	0.65
% participants shared with median split	0.61	-	0.83

Table 4.

Peak coordinates and size of cluster from whole brain models of SC task: all correct trials vs. baseline

	Area	Peak Coordinates			No. of voxels
Pre-intervention group comparisons		x	y	z	
Future improvers vs future non-improvers	Right fusiform	+36	-62	-12	152
	Right lingual gyrus	+24	-58	+2	
Future improvers vs. non-struggling	precuneus	-6	-60	+26	155
Future non-improvers vs. non-struggling	Right posterior middle temporal gyrus	+62	-52	-10	263
	Left post central white matter	-36	-26	+32	102
Post intervention gain correlate	Right fusiform	+38	-50	-16	99

Cluster corrected for multiple comparisons $z > 3.1$, $ps < .05$.