Reading intervention and neuroplasticity: A systematic review and meta-analysis of brain changes associated with reading intervention

Meaghan V. Perdue*[a,b,c], Kelly Mahaffy[a,b], Katherine Vlahcevic[a,b], Emma Wolfman[a], Florina Erbeli[d], Fabio Richlan[e], & Nicole Landi[a,b]

a. Dept. of Psychological Sciences, University of Connecticut, Storrs, CT, USA
b. Haskins Laboratories, New Haven, CT, USA
c. Dept. of Radiology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
d. Dept. of Educational Psychology, Texas A&M University, College Station, TX, USA
e. Centre for Cognitive Neuroscience & Department of Psychology, University of Salzburg, Salzburg, Austria

*Corresponding Author:

Meaghan Perdue
Alberta Children's Hospital
28 Oki Drive NW
Calgary, AB
Canada

E-mail address: meaghan.perdue@ucalgary.ca

Key words: reading intervention, reading disability, dyslexia, neuroimaging, fMRI, meta-analysis, systematic review

Abstract

Behavioral research supports the efficacy of intervention for reading disability, but the brain mechanisms underlying improvement in reading are not well understood. Here, we review 39 neuroimaging studies of reading intervention to characterize links between reading improvement and changes in the brain. We report evidence of changes in activation, connectivity, and structure within the reading network, and right hemisphere, frontal and sub-cortical regions. Our meta-analysis of changes in brain activation from pre- to post- reading intervention in eight studies did not yield any significant effects. Methodological heterogeneity among studies may contribute to the lack of significant meta-analytic findings. Based on our qualitative synthesis, we propose that brain changes in response to intervention should be considered in terms of interactions among distributed cognitive, linguistic and sensory systems, rather than via a "normalized" vs. "compensatory" dichotomy. Further empirical research is needed to identify effects of moderating factors such as features of intervention programs, neuroimaging tasks, and individual differences among participants.

INTRODUCTION

Learning to read is instrumental for academic success and day-to-day activities, yet a striking nine percent or more of school-age children experience severe and persistent difficulties in accurate and/or fluent word recognition, here referred to as reading disability (RD, also known as developmental dyslexia; Pennington & Bishop, 2009; Peterson & Pennington, 2012; Lyon et al., 1995). Research over the past several decades reveals a pattern of atypical brain activation in groups with RD, most consistently characterized by atypical activation during reading-related tasks in the posterior hubs of the typical reading network: left superior temporal gyrus/sulcus (STG/STS), inferior parietal lobe (IPL), occipito-temporal cortex (OT), and inferior frontal gyrus (IFG) (Maisog et al., 2008; Paulesu et al., 2014; Richlan et al., 2009).

Importantly, research has demonstrated that reading intervention is helpful for individuals with RD; indeed, recent meta-analyses of the effects of reading intervention show moderate effect sizes related to growth in reading ability (Wanzek et al., 2018, 2016, 2013). Reading intervention programs vary in the skills targeted, total hours or weeks of intervention, intensity (number of hours per week), group size (one-on-one vs. small group), and modality (in person vs. computerized). Many reading intervention programs focus on a single pre-reading skill or a small set of pre-reading skills, such as phonological awareness and/or grapheme-phoneme correspondence (e.g., Heim et al., 2015; Brem at el., 2010; Partanen et al., 2019). Some programs also incorporate domain-general skills that support reading, such as executive function or attention training (Horowitz-Kraus et al., 2019; Horowitz-Kraus et al., 2014; Heim et al., 2015). While heterogeneity in approaches and individual differences among participants make it difficult to draw specific conclusions with regard to which aspects of intervention programs are most effective, one consistent finding is that explicit phonics instruction is useful for many people with

RD, and is therefore considered a gold-standard in intervention programming (Galuschka et al., 2014).

Despite clear evidence that individuals with RD can benefit from intervention, the neurobiological mechanisms that support improvement in reading ability are not well understood. Extant literature on this topic tends to focus on two putative mechanisms by which individuals improve at the neurobiological level: compensation and normalization (Barquero et al., 2014; D'Mello & Gabrieli, 2018; Koyama et al., 2013; Simos et al., 2002). Researchers have suggested that increased activation during reading-related tasks in regions associated with domain-general cognitive processing, including right hemisphere and frontal and subcortical structures, reflects compensation for dysfunction of the left-hemisphere reading system (D'Mello & Gabrieli, 2018; Pugh et al., 2001; Shaywitz et al., 2002). Putative compensatory processing may take several forms, such as increased reliance on working memory, attention, articulatory mechanisms, and/or declarative memory to overcome reading difficulties (D'Mello & Gabrieli, 2018; Hancock et al., 2017; Pugh et al., 2001; Shaywitz et al., 2002; Ullman et al., 2020; Yu et al., 2018). Normalization on the other hand, is typically inferred following increased activation in the "typical" reading network, which is thought to indicate engagement of typical reading strategies via phonological decoding and/or rapid word recognition (Barquero et al., 2014; D'Mello & Gabrieli, 2018; Simos et al., 2002). A growing body of research has begun to reveal intervention-related, changes in gray matter volume, cortical thickness and white matter properties, providing evidence that functional changes are accompanied by changes in brain structure (Davis et al., 2010; Huber et al., 2018; Keller & Just, 2009; Krafnick et al., 2011; Richards et al., 2017, 2018; Romeo et al., 2017).

Interestingly, these patterns do not seem to be all or none as some studies provide evidence of activation increases in both the typical reading network and regions outside this network within the same samples (e.g., Horowitz-Kraus, et al. 2014; Temple, et al. 2003). A previous meta-analysis of functional brain differences following reading intervention included 8 studies and showed effects in the left thalamus, right insula/IFG, left IFG, right posterior cingulate, and left middle occipital gyrus (Barquero et al., 2014). This set of regions supports reading sub-skills such as phonological processing as well as broader cognitive functions such as attention and memory. Thus, the meta-analytic findings show that intervention-related changes occur both within and outside the typical reading network. Importantly, the cognitive and sensory mechanisms underlying improvement in reading ability cannot be inferred based on the locations of functional changes alone; rather, regional changes provide hints about the possible mechanisms, and carefully designed empirical studies are needed to investigate these pathways. Moreover, methodological factors such as features of the intervention program and/or fMRI task may influence the patterns of changes in the brain, and these factors must be considered when interpreting brain changes in response to intervention.

We suggest that framing the neural mechanisms associated with reading intervention as “normalized” versus “compensatory” is an over-simplified dichotomy, and a dynamic, network-based interpretation should be considered. As the understanding of human brain function moves towards a network approach, emerging research on the neural mechanisms of reading and reading intervention similarly reveal that the changes in the connections among brain regions are as important, if not more important, than the changes in local activation. Indeed, distinct patterns of resting state functional connectivity have been observed among typically developing children and children with a history of RD who were grouped based on profiles of remediation of reading and spelling abilities (Koyama, et al., 2013). Findings from reading intervention studies show that retuning of the connections among brain regions involves enhancement of some connections, and reductions of others (e.g., Richards et al., 2016). These types of changes could underlie the mixed effects observed in task-based functional activation studies and may be more fully explored by examining changes in functional connectivity associated with intervention. While these approaches have been applied less frequently, we include available studies in our systematic review that have used such methods.

We present a systematic review and meta-analysis of neuroimaging studies of reading intervention to identify convergent intervention-related effects in the current literature. In our systematic review, we consider changes in brain function and structure associated with reading intervention, with particular attention to links between patterns of brain activation and responsiveness to intervention. We aim to provide a comprehensive review of this topic, so we additionally include studies of neural predictors of response to intervention and studies of brain changes with intervention in pre-readers at risk of RD.

Our meta-analysis focuses more narrowly on intervention-related changes in brain activation during reading-related tasks and includes 8 studies. A prior meta-analysis on this topic provided a broad overview of differences in brain activation following reading intervention, including studies of pre-post intervention change and post-intervention group comparisons (Barquero et al., 2014). Here, we sought to provide a more targeted analysis by only including results of pre-post intervention change. This allowed us to evaluate effects across a more homogeneous set of findings, focusing on intervention-related *changes* in brain function. In addition, we used an alternate meta-analytic method that accounts for effect sizes and signs of peaks reported in primary studies and accounts for sample sizes of the primary studies to weigh contributions to the meta-analytic results. Based on previous findings, we hypothesized that reading intervention would be related to changes in brain activation both within the left-hemisphere hubs of the reading network (e.g., Shaywitz, et al., 2004) and in homotopic "compensatory" regions of the right hemisphere (e.g., Aylward, et al., 2003). We conclude by offering suggestions for future empirical research to advance the understanding of neural mechanisms underlying reading remediation.

2. METHOD

2.1 Literature search & screening

2.1.1. Eligibility Criteria

The following criteria were set for inclusion in our systematic review: (1) Primary research studies including peer-reviewed, published journal articles, in press articles, in prep articles, conference proceedings, conference presentations, dissertations; (2) Article full-text must have been available in English; (3) Studies must have included participants with or at-risk[1] for developmental reading disability (i.e. dyslexia); acquired forms of reading disability excluded (e.g. resulting from trauma); (4) Studies must have included reading related instruction/intervention; (5) Studies must have included pre- and/or post-intervention neuroimaging in structural magnetic resonance imaging (MRI), functional MRI, or magnetoencephalography (MEG) modality using a reading or reading-related task (e.g. phonological processing, orthographic processing); (6) case studies were excluded. Additional criteria were applied for inclusion in the quantitative meta-analysis: (7) Neuroimaging acquired using fMRI modality; (8) Neuroimaging acquired at both pre- and post-intervention time points; (9) Whole-brain voxel-wise analysis must have been used.

[1] We acknowledge that risk status is not a perfect indicator of later reading outcomes, but we chose to include studies of children at-risk of RD based on low performance on assessments of pre-literacy skills and/or familial risk in order to obtain the full scope of reading intervention literature. Evaluation of behavioral and neural outcomes of early interventions is needed to inform educational practice, and research in young children relies on risk status to classify groups.

2.1.2. Retrieval of records

We conducted a literature search in a set of databases that include research related to psychology, education, and neuroimaging (PsychInfo, ERIC, Academic Search Ultimate, MedLine, EBSCOhost eBook Collection, PubMed). Search terms are reported in the supplementary materials. In addition, we sent calls for grey literature to the mailing lists of relevant scientific societies: the Society for the Scientific Study of Reading, the Cognitive Development Society, and the Society for the Neurobiology of Language. The initial literature search and calls for grey literature were conducted in March 2020. We also screened the references of previously published reviews of neuroimaging studies of reading intervention to identify additional relevant articles (Barquero et al., 2014; Richards et al., 2006). In addition, we conducted a search for articles that have cited the articles that met criteria for our systematic review using the "search within citing articles" function in Google Scholar (search terms available in supplementary materials).

2.1.3. Screening of records

Records were screened according to the eligibility criteria listed in section 2.1.1. The screening was tracked using the PRISMA Flow Diagram (Fig. 1; Page et al., 2020). Our initial database search yielded 787 records before removal of duplicates. After removal of duplicates, 571 records remained for initial screening. Five additional records were identified through screening the reference lists of prior reviews. Screening of titles and abstracts was divided among 4 authors (MP, KM, KV, & EW) and tracked using the Rayyan QCRI web application. Following initial record screening, 42 full-text articles were screened for inclusion in the systematic review and meta-analysis. Information about participants, intervention procedures and imaging measures was coded from the full text articles. Two authors independently screened and coded data from the full-text articles (KM & KV; agreement among raters was 96.54%); discrepancies were resolved by a third author blind to the initial coder (MP). Coding procedures and the full list of items coded is available in the Supplementary Materials (Supplementary Table 1 and online at https://osf.io/eyt5h/?view_only=2c0933bfd57b459aaaf05a8db785427c). 31 studies met criteria for inclusion in the systematic review, and the reference lists of these articles were screened to identify additional articles for inclusion. 12 additional full-text articles were screened, and 4 of these met criteria for inclusion in the systematic review. Based on citation searching in Google Scholar, 19 additional full-text articles were screened, and 4 of these met criteria for inclusion in our qualitative review. In total, 39 articles met criteria for inclusion in the qualitative review. Of these, 8 met criteria for inclusion in the quantitative meta-analysis. Reasons for exclusion of 31 articles included in the systematic review but not in the meta-analysis are noted in the flowchart (Fig. 1).

---------------------------*Insert Figure 1 here*--------------------------------

2.2. Meta-analysis procedure

2.2.1 Data extraction and preparation

Quantitative meta-analysis was conducted using seed-based d mapping (SDM; Albajes-Eizagirre, Solanes, Vieta, & Radua, 2019). Our main analysis tested for effects related to patterns of brain activation in children with/at-risk for reading disability following reading intervention. Contrasts of interest included (1 per study, in preferred order): RD post-intervention vs. pre-intervention (Gebauer et al., 2012; Heim et al., 2015; Richards, et al., 2006; Temple et al., 2003; Yamada et al., 2011), RD follow-up vs. pre-intervention (Shaywitz et al., 2004), or group-by-time interaction between RD and TD participants from pre-to-post-intervention (Eden et al., 2004; Partanen et al., 2019). For each study, we extracted the peak coordinates and *T*, *Z*, or *F* statistics reported for the contrast of interest. *Z* statistics were converted to *T* statistics, and *T*-values were estimated from *F*-statistics for two studies that reported only interaction effects, and not within-group main effects of time (Eden, et al., 2004; Partanen et al., 2019). Peak coordinates, *T*-statistics and significance thresholds for all the remaining studies were entered into SDM for meta-analysis. Peak coordinate files and the SDM table can be accessed at: https://osf.io/eyt5h/?view_only=2c0933bfd57b459aaaf05a8db785427c).

2.2.3. Meta-analysis using SDM-PSI

Voxel-based meta-analysis was conducted via permutation of subject images for seed-based d mapping following the procedures recommended by the authors of the software (SDM-PSI; Albajes-Eizagirre et al., 2019). This algorithm has been successfully implemented in recent meta-analyses, including studies of treatment-related change in brain activation (Chan et al., 2021; Lv et al., 2021).

Preprocessing consisted of estimating the lower and upper bounds of possible effect sizes for each voxel and creating image maps of these values for each study. This step was conducted with default values for fMRI: correlation template=gray matter, anisotropy = 1.00, Isotropic FWHM = 20mm, mask = gray matter, voxel size = 2mm. Next, the most likely effect size and its standard error was estimated using maximum likelihood estimation with 50 iterations. A set of imputed effect-size image maps were created for each study within the range of possible effect sizes estimated during preprocessing. To facilitate permutation testing, subject images that realistically represent local spatial covariance were imputed for each study and adapted to the different imputed study images. Subject-based permutation testing with 1000 permutations was applied to the imputed data for each study to control the family wise error rate. Maximum statistic tests were conducted using threshold free cluster enhancement (TFCE) statistics (Smith & Nichols, 2009). Thresholding of results is computed with a pair of one-tailed tests, so the probability threshold was set to $p < .025$ to control the false positive rate.

2.3. Method for Assessing Risk to Internal Validity

To assess risk to internal validity of primary studies, all publications included in the systematic review were rated on 13 domains that may introduce bias into controlled intervention studies. We rated studies based on the criteria proposed by the National Institutes of Health to assess quality of controlled intervention studies (e.g., randomization, attrition; https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools). A list of domains and the specific criteria on which raters made their judgements is available in the supplementary materials (Supplementary Table 2). The publications were divided among two raters (KV & KM) for this assessment. In order to evaluate inter-rater reliability, an overlapping sample of 8 of the papers (20.5%) were coded by both raters, and their inter-rater reliability was 92.71% agreement.

Once all publications had been coded, the web-based version of *robvis,* a risk-of-bias visualization tool, was used to create graphic summaries of each article's respective ratings (McGuinness and Higgins, 2020). We highlight concerns related to the overall risk of bias in the reviewed literature in the discussion.

3. RESULTS

3.1. Systematic Review

We systematically identified 39 studies that met criteria for our qualitative review. Study characteristics, including information about samples, intervention methods, and neuroimaging methods, along with key results, are presented in tables 1-4. A qualitative review and synthesis of findings across these studies is presented in the following sections. We begin with a review of neural changes in response to intervention in MEG, functional MRI, functional connectivity, and structural MRI modalities (i.e., white and gray matter structure). These sections are followed by a review of neural predictors of response to intervention.

3.1.1. Functional neural changes in response to intervention: MEG studies

Some of the earliest studies on this topic applied MEG methods to measure brain activation changes associated with reading intervention in. In their first report, Simos and colleagues observed normalization of neural activation during a visual pseudoword rhyme matching task following intervention (Simos et al., 2002). Prior to intervention, children with RD had lower activation than TD children in the left posterior STG, and rightward lateralization of activation in the posterior STG. Following intervention, phonological decoding scores were improved to the average range and lateralization of activation in the STG was shifted to a left-dominant pattern, driven by increased activation in the left and modestly decreased activation in the right. In a later study, the authors compared changes in brain activation during pseudoword reading over a 16-week intervention period and found that children who improved showed increased activation in left STG and IPL (Simos et al., 2007b). In contrast, those who did not improve showed increased activation in right STG and bilateral IFG, indicating that recruitment of these regions was not beneficial to reading performance at the group level.

In addition, these authors examined brain activation during word reading in a small sample of children with severe RD who had persistent difficulties despite quality reading instruction in first grade (Simos et al., 2007a). MEG scans were acquired before and after intervention. Overall, children improved their reading skills across the intervention period. Increased activation in bilateral posterior middle temporal gyri was observed following intervention, and activation in the left region was positively associated with in-scanner word reading performance, showing that increased activation in this left posterior temporal region may facilitate better word reading outcomes. Decreased onset latency of activity, which is thought to reflect increased efficiency, was observed in left middle temporal and right OT regions, and was associated with in-scanner reading performance, suggesting that greater neural efficiency in these regions supports word reading accuracy. The degree of change in onset latency of activation in the left premotor cortex was inversely related to word reading performance in scanner, such that earlier onset of activity in this region was associated with poorer word reading performance, and the authors interpret this effect as a shift away from a compensatory engagement of the premotor cortex in children who had better reading outcomes.

Altogether, this body of MEG research supports the hypothesis that increased engagement of left hemisphere reading network hubs facilitates remediation of reading difficulties. The ability to probe latency effects using MEG also provides insight to the mechanisms that may be associated with more efficient processing as children improve their reading skills.

3.1.2. Functional neural changes in response to intervention: fMRI studies

The majority of studies of the neural correlates of reading intervention have used fMRI. This section is organized based on the type of fMRI task applied: single word/pseudoword reading, orthographic-phonological mapping (e.g., letter-sound matching), phonological processing (auditory stimuli only), and sentence comprehension.

Single word/pseudoword reading

Many of the studies of intervention-related changes in brain activation have used word and pseudoword reading tasks. In this section, we review studies using tasks with visually presented single word and/or pseudoword stimuli with or without a lexical or phonological judgement.

Several early studies employed tasks with single word/pseudoword stimuli including phonological judgement, spelling judgement, and/or morphological semantic judgement (Aylward et al., 2003; Richards, Aylward, Berninger, et al., 2006; Richards, Aylward, Field, et al., 2006). Aylward and colleagues (2003) acquired fMRI scans from children performing phonological and morphological tasks before and after an intensive 2-week intervention. Children improved in pseudoword decoding, morphological awareness, and oral reading accuracy over the course of the intervention. Prior to intervention, children with RD showed lower activation than the control group in the left inferior/middle frontal gyrus and bilateral superior parietal cortex during phonological judgement. These group differences were no longer present after intervention due to both increased activation in children with RD and decreased activation in the control group. Interestingly, the control group showed decreased activation in the right, but not the left, superior parietal cortex, which may reflect developmentally appropriate leftward lateralization for reading driven by disengagement of the right parietal cortex in this group. In contrast, the RD group showed increased activation bilaterally in the parietal cortex to a level that no longer differed from controls. During the morphological judgement task at baseline, children with RD showed reduced activation relative to controls in the right OT, right superior parietal cortex, bilateral occipital-parietal junction and left middle frontal gyrus. Following intervention, the only significant difference between groups during the morphological task was in the right visual cortex. Increased activation in the RD group accounted for elimination of the group difference in right OT; elimination of group differences in parietal and frontal regions were driven by effects in the TD group, and/or small changes in opposite directions in both groups. Notably, although brain differences between the RD and TD groups diminished and the RD group improved substantially in reading and related skills, the RD group did not show complete behavioral remediation and still performed below average on behavioral assessments.

Richards and colleagues conducted several studies examining brain activation changes following intervention targeting skills in various reading and language domains (i.e. phonology, morphology, orthography). In one study, children with RD had reduced activation during orthographic judgement in right parietal and IFG regions relative to TD children at pre-intervention scanning (Richards, Aylward, Berninger, et al., 2006); following an intervention targeting orthographic skills, children in the RD group increased activation to levels that no longer differed from the TD group, though their spelling skills remained below average. In another study, children received either phonological or morphological treatment and both groups showed improvement in word level reading and reading comprehension (Richards, Aylward, Field, et al., 2006). At the brain level, increases in activation following treatment were observed in left OT and left posterior insula during a phonological judgement task; increases in activation during morphological judgement were observed in bilateral STG, superior frontal gyri and left anterior insula. In a third study by Richards and colleagues (2007), activation during pseudoword reading was examined before and after groups with RD received either phonological or non-phonological (nonverbal problem solving) intervention. Both groups improved in reading skills, and the non-phonological treatment group showed increased activation in superior-lateral occipital cortex, but the phonological treatment group did not show any significant change during this task.

In a study of children's response to intervention programs targeting either phonological, attention, or word recognition skills, Heim and colleagues (2015) reported reading improvement in all groups as well as domain-specific gains in the phonological and attention treatment groups. With respect to brain activation during overt word and pseudoword reading, the RD group showed a greater increase in activation across the intervention period in the left inferior OT relative to the TD group, regardless of treatment group. Effects specific to the treatment groups showed that both phonological and reading groups had greater increases from pre- to post-intervention in bilateral parietal activation relative to the attention group; the attention group had greater increases in left STG/STS. Children with lower reading scores at pre-intervention and those with a larger discrepancy between phonological awareness and attention scores at pre-intervention showed greater increases in left OT activation over the intervention period.

Gebauer and colleagues (2012) tested the neural effects of a 5-week morpheme-based intervention in children with poor spelling and reading skills using a lexical decision fMRI paradigm. Children with spelling difficulties were assigned to an intervention group or a waiting control group, and a group of TD children who did not receive intervention was included for comparison. Notably, children with poor spelling had significantly poorer reading skills compared to the TD control group, but many performed within the average range on standardized reading assessments. The intervention group improved substantially in spelling and modestly in reading comprehension, and all groups improved in reading speed. Prior to intervention, children with reading/spelling difficulties had lower activation in left OT regions, hippocampus, and cerebellum than the TD group, along with higher activation in precuneus, right posterior paracingulate, medial frontal, right frontal, and right temporal regions. The intervention group showed increased activation specific to pseudoword processing in right posterior cingulate, left inferior/middle temporal gyrus, and left hippocampus/parahippocampus regions. Direct comparison of intervention group and waiting group on change over time showed training-specific effects in bilateral parahippocampal gyri and cerebellum during processing of misspelled words. Meanwhile, the waiting group had activation increases in precuneus, cerebellum, left frontal pole, right lateral occipital, and right parieto-temporal regions during correctly spelled and misspelled word processing. Importantly, increased activation in right occipital and temporal regions, left precentral gyrus, and bilateral cerebellum was associated with less behavioral improvement in the intervention group, indicating that engagement of these regions may be detrimental for spelling ability.

Horowitz-Kraus and colleagues (2014) have conducted research to evaluate the behavioral and neural effects of a computer-based intervention program that targets reading fluency and executive functioning skills. Along with behavioral improvement in reading accuracy, speed, and comprehension, they reported increased activation after intervention for word relative to pseudoword reading in left inferior occipital gyrus and left STG across groups and in right anterior cingulate cortex in the RD group. Group differences following intervention showed that the RD group had greater activation in right IFG than the TD group. Within the RD group, gain in contextual reading rate positively correlated with post-intervention activation of left anterior middle frontal gyrus, and gain in word/pseudoword reading fluency positively correlated with activation in left middle frontal gyrus, left OT, and left inferior occipital gyrus. In the TD group, positive correlations were observed between reading gains and activation in right IFG, middle frontal and IPL regions.

Koen and colleagues (2018) reported a set of trending effects related to a computer-based intervention program targeting reading fluency. The intervention group showed more clusters of activation at post-intervention than pre-intervention, and more right hemisphere than left hemisphere clusters of activation at the post intervention time point during a pseudoword phonological judgement task. Together with the findings from Horowitz-Kraus et al. (2014), this study provides evidence in support of computer-based fluency intervention to effect changes in brain activation, though further research in larger samples is needed to further support the findings related to the Koen et al. program.

In a recent study, Partanen and colleagues (2019) examined intervention-related changes in brain activation during a printed word rhyming task and a spelling judgement task. Children with RD showed greater activation following school-based intervention than at pre-intervention scanning during the printed word rhyming task in bilateral insula and IFG (Partanen et al., 2019). Following intervention, poor readers had greater activation than good readers in right parietal cortex for easy (more frequent) words, and activation in that region was positively associated with improvement in non-word decoding, supporting a compensatory role of the right parietal cortex in phonological aspects of reading. These findings appear to be specific to the phonological aspects of reading because no such effect was observed during the spelling judgement task, and the functional changes were associated with non-word decoding, but not word recognition. No intervention-specific effects were observed for the spelling task, but a main effect of time showed increased activation following intervention in bilateral cerebellum and right IPL in both good and poor reader groups. A group-specific effect showed increased activity in the right STG/STS in good readers, but not poor readers. Thus, the skills gained in this intervention appeared to primarily affect phonological skills and associated neural circuitry and may not generalize to orthographic knowledge and recognition of irregularly spelled words.

Together, these studies reveal intervention-related changes in word/pseudoword reading in a widespread set of regions in the left and right hemispheres including reading network hubs, as well as cingulate cortex, hippocampus, and cerebellum, but do not converge on a specific set of functional brain changes associated with reading intervention. Though these studies used similar types of stimuli (words and pseudowords) in their fMRI tasks, they differ in the contrasts reported (e.g. pseudoword-specific processing [Gebauer et al., 2012] vs. word-specific processing [Horowitz-Kraus et al., 2014]) and in whether they required a lexical (Gebauer et al., 2012; Horowitz-Kraus et al., 2014), phonological (Koen et al., 2018; Partanen et al., 2019; Richards, Aylward, Field, et al., 2006), orthographic (Aylward et al., 2003; Richards, Aylward, Berninger, et al., 2006), semantic (Aylward et al., 2003; Richards, Aylward, Field, et al., 2006) judgement, or no judgement (Heim et al., 2015). Moreover, differences between activation patterns elicited by different tasks (single-word reading versus lexical decision; words versus pseudowords) have been reported (Murphy et al., 2019; Taylor et al., 2013), so these methodological differences along with intervention-related factors may account (at least in part) for the lack of convergence among studies.

Orthographic-phonological mapping tasks

The ability to integrate orthographic and phonological information is a fundamental component of reading, and several studies of reading intervention have examined changes in activation during orthographic and phonological integration. In this section, we review studies that used tasks involving letter-sound matching and/or letter rhyming.

One such study showed that children who received an acoustic-based intervention targeting phonological processing showed group-level improvements in word reading, comprehension, oral language, and rapid naming, accompanied by increased activation during letter rhyming in a broad bilateral network spanning frontal, cingulate, middle temporal, and parieto-occipital regions as well as anterior thalamus (Temple et al., 2003). The increase in right IFG activation from pre-to-post intervention was positively associated with improvement in phonological processing performance (blending words), indicating that this region could be engaged to support phonological aspects of reading. In addition, the authors reported a positive correlation between increase in left TP activation and improvement in oral language and phonics skills, supporting the role of this region in reading-related skill development.

There is also evidence linking right hemisphere activation during orthographic-phonological mapping in RD to greater gains in reading over the intervention period. Odegard and colleagues (2008) reported that activation of the right IFG during phoneme-grapheme mapping following a comprehensive remediation program distinguished children who responded to treatment from those who did not, with higher activation in those who improved their phonological awareness and decoding abilities to the average range.

Shaywitz and colleagues (2004) examined brain activation during an audio-visual letter identification task in children with RD who received an experimental phonics-based intervention relative to a TD group and to an RD group that received varied community-based intervention/tutoring. Immediately following intervention, children in the experimental intervention group and children in the TD group showed greater increases in activity in left IFG and posterior middle temporal regions relative to the community intervention group; the experimental intervention group showed reduced activation in the right caudate nucleus relative to both control groups. In addition, comparisons of activation at pre-intervention and one-year follow-up in the experimental group showed increases in activation in bilateral IFG, left STS, and left OT. Decreases of activation were reported in the right middle temporal gyrus and caudate nucleus.

Though only a few studies have used orthographic-phonological integration tasks to investigate functional changes related to reading intervention, they provide preliminary evidence to support a role of the right IFG as increased activation in this region was associated with improvement in reading and/or phonological processing abilities in two of the studies (Temple et al., 2003; Odegard et al., 2008). Shaywitz and colleagues (2004) also reported increased activation in the bilateral IFG at follow-up, pointing to long-term effects of intervention in these regions. These findings support a potential role of the right IFG as a compensatory mechanism to overcome impairments in left hemisphere phonological processing pathways.

Auditory phonological processing tasks

Eden and colleagues (2004) examined changes in activation during phonological processing following a phonologically based intervention in adults with RD. Participants showed behavioral improvement in phonological processing, pseudoword decoding, and text reading accuracy, along with increased activation in bilateral reading network regions including left OT, right STG/STS, and bilateral parietal cortex during a phonological manipulation task. In another study, Richards et al. (2007) investigated effects of phonological and nonphonological (nonverbal problem solving) interventions on brain activation during an aural pseudoword repetition task. The group that received phonological treatment showed decreased activation from pre-post intervention in the left IPL and left postcentral gyrus, while the non-phonological treatment group showed increased activation in these regions. These findings provide evidence that patterns of change in brain activation differ by the focus of the intervention.

Sentence comprehension tasks

Meyler and colleagues (2008) assessed effects of a six-month intervention on brain activation during sentence comprehension and reported changes in a distributed set of brain regions. Prior to intervention, children with RD showed lower activation in left occipital/angular gyrus, left parietal, and left middle frontal cortex, and right IPL, as well as higher activation in supplementary motor area. Regions of reduced activation in children with RD relative to TD children diminished with remediation, though reduced activation remained in left superior parietal, superior occipital, and middle frontal regions. Diminished group differences were explained by significantly increased activation in left IPL and superior parietal cortex in the RD group and decreased activation in bilateral IPL in the TD group. The RD group additionally showed higher activation than the TD group in left putamen and right insula/IFG after intervention, possibly indicating recruitment of compensatory mechanisms. At 1-year follow-up, the RD group showed hyper-activation in a distributed bilateral set of cortical and subcortical regions, and hypo-activation in the left occipital cortex relative to the TD group.

Another recent study showed that post-intervention activation in the right OT during sentence comprehension positively correlated with individual differences in gains on a composite reading score after remediation over 1-2 school years (Nugiel et al., 2019). In addition, the authors reported that greater increases in activation in the left OT over the intervention period were associated with greater improvement in reading, but these effects did not survive correction for multiple comparisons, and replication of findings is needed.

3.1.3 fMRI studies of intervention in pre-readers

Early intervention has been associated with greater reading gains over the course of treatment and in the years following treatment (Lovett et al., 2017), but only a few studies have investigated the brain mechanisms associated with intervention in pre-readers at risk of RD. In one study, pre-reading children at varying family risk of RD completed eight weeks of computerized training focused on letter-sound correspondences and eight weeks of a nonlinguistic control training on number knowledge (Brem et al., 2010). Training-specific effects showed that increases in activation in bilateral fusiform gyrus and cuneus during implicit print processing were greater over the reading training phase than the control number knowledge training phase. Notably, children with and without family risk of RD were pooled for the analysis, so effects related to typical reading acquisition cannot be distinguished from effects specific to training in children at-risk of RD. A follow-up study was conducted to investigate whether activation during an explicit word reading task that was administered after the reading training was associated with reading outcomes measured in second grade (Bach et al., 2013). Post-intervention activation in left OT correlated positively with changes in lower-case letter knowledge and negatively with reading risk score over the kindergarten intervention period.

Yamada and colleagues (2011) examined brain activation during letter processing in kindergarteners who had on-track or at-risk performance on a literacy screening assessment. The at-risk group showed increases in activation from pre-to-post-intervention in left hemisphere reading network regions and right hemisphere homologues, including left posterior STG and orbitofrontal cortex, and bilateral IPL, IFG and anterior cingulate cortex. In comparisons of post-intervention activation between the at-risk and on-track groups, the at-risk group showed greater activation than the on-track group in right IPL, anterior cingulate cortex, middle frontal gyrus, left IFG and precentral gyrus, and bilateral paracingulate gyrus; the on-track group showed greater activation in left superior lateral occipital cortex.

The findings in pre-readers at risk of RD support the notion that engaging the hubs of the left hemisphere reading network is beneficial for subsequent reading acquisition. The bilateral activation observed in at-risk children is consistent with the observation of more bilaterally distributed processing for reading in groups with RD. However, given the young age of these participants, it is difficult to determine whether right hemisphere activation at this stage reflects compensatory processing, or a delay in shifting toward a left-lateralized reading network.

In our quality analysis, both Brem (2010) and Bach (2013) raised moderate concern centered on a high drop-out rate resulting in small samples for analysis and not reporting power. Despite these concerns, it is worth noting that these are two of a small number of studies that include an RD control group and feature a quasi-randomized design in which RD participants were randomly assigned to complete either a grapheme/phoneme correspondence or control (math) intervention. The Yamada (2011) study raised moderate concern due to high drop-out rate and no information on intervention duration, however it is one of the only studies to report that experimenters conducting the analysis were blind to group assignment.

--------------------*insert table 1 here*-----------------------

3.1.4 MEG & fMRI studies of reading intervention: Summary

In sum, evidence from functional neuroimaging studies reveals changes in both the typical reading network and regions outside this network over the course of intervention. Increased activation in hubs of the left hemisphere reading network were reported in several studies, indicating that function of this typical network can be recovered to some extent with training (Heim et al., 2015; Horowitz-Kraus et al. 2014; Richards, Aylward, Field, et al., 2006; Shaywitz et al., 2004). In some studies, activation changes occurred such that group differences that were evident prior to intervention were no longer present after intervention (Aylward et al., 2003; Meyler et al., 2008; Richards, Aylward, Berninger, et al., 2006). These changes sometimes involved increased activation in the right hemisphere to match levels of controls, which could reflect normal engagement of these regions rather than atypical recruitment of compensatory neural pathways to support reading. Evidence of increased activation in right hemisphere and sub-cortical regions has also been reported (Gebauer et al., 2012; Meyler et al., 2008; Partanen et al., 2019; Nugiel et al., 2019), with the most consistent effects in right IFG (Horowitz-Kraus et al., 2014; Meyler et al, 2008; Odegard et al., 2008; Partanen et al., 2019; Temple et al., 2003). Importantly, several of these studies linked activation increases in the right IFG to improvement in reading ability (Temple et al., 2003; Odegard et al., 2008). Responsiveness to intervention has also been associated with greater brain activation in reading network hubs and their right-hemisphere counterparts (Nugiel et al., 2019; Odegard et al., 2008; Simos et al., 2007b; Temple et al., 2003). Further research is needed to identify consistent effects related to improvement in reading ability.

3.1.5: Neural changes in response to intervention: Functional connectivity

Reading intervention has also been associated with changes in functional connectivity. Richards and colleagues have conducted several studies to examine such effects. In the first of these, children with RD showed greater connectivity of the left IFG with bilateral middle frontal gyrus, supplementary motor area, left precentral gyrus, and right superior frontal gyrus and IFG during a phoneme-grapheme mapping task than TD readers prior to intervention (Richards &

Berninger, 2008). Group differences in functional connectivity were no longer present after a 3-week intervention.

In a later study, Richards and colleagues (2016) examined changes in functional connectivity in children with RD following a computerized training program that targeted reading and writing skills. Functional connectivity between the right IPL and left anterior cingulate gyrus during a multiple-sentence comprehension task decreased from pre-intervention to post-intervention. In contrast, functional connectivity between the right IPL and right IFG during single sentence reading comprehension increased from pre-intervention to post-intervention.

Following this study, Richards and colleagues (2017) examined changes in both functional and structural connectivity in children with various types of language-based learning disabilities who completed a similar reading and writing intervention program. Four groups of children were tested: children with dysgraphia, children with RD, children with oral language difficulties, and a control group with typical reading and writing skills. Behaviorally, improvement in handwriting and oral sentence syntax construction was observed across groups, and some improvement was observed in spelling tasks, but spelling improvements were minimal in the group with RD. Graph theory analysis of functional connectivity during a spelling task revealed modest increases in connectivity with the right IFG and decreases in connectivity with the left cingulate gyrus. A group by time interaction in the left inferior cingulate gyrus showed an increase in connectivity specific to the group with oral language difficulties.

The authors published a second report on this study in which they examined functional connectivity during a set of reading tasks (Richards et al., 2018). They found that local functional connectivity in right cingulate gyrus during word-level reading increased in the two reading disabilities groups, but decreased in the TD and dysgraphia groups. Local functional connectivity during sentence reading comprehension increased in the left superior frontal gyrus and left IFG in the reading disabilities groups, but decreased in the other groups. Local functional connectivity during multi-sentence reading comprehension in the left insula decreased from pre-post intervention, driven by effects in the dyslexia and dysgraphia groups. Local functional connectivity in right middle frontal gyrus during multi-sentence reading comprehension decreased in the reading disability groups, and increased in the dysgraphia and TD groups.

Horowitz-Kraus and colleagues (2019) focused on changes in functional connectivity during a lexical decision task in groups with RD and/or attention-deficit and hyperactivity disorder (ADHD) as well as a TD control group following participation in a computer-based intervention program that targeted reading and executive functioning skills. They conducted an independent component analysis of the functional data to identify networks for connectivity analysis, and our summary of the results refers to these components. Functional connectivity between the low-level visual component (bilateral fusiform gyrus) and the dorsal attention component (bilateral precuneus/posterior cingulate) and between the attention component (bilateral anterior cingulate) and semantic/articulation component (bilateral insula) and between the attention component and higher-level visual component (bilateral lingual gyri) increased over the intervention period in the RD group. In the group with ADHD+RD, functional connectivity increased between the low-level visual component and the executive function component (bilateral superior frontal gyri) and between the attention component and dorsal attention component. In contrast, connectivity increased between the low-level visual component and the phonological component (bilateral IPL) and between the attention component and the memory component (bilateral parahippocampal gyri) in the TD group. Gain in reading speed positively correlated with increased functional connectivity between the low-level visual component and the executive function component and with increased functional connectivity between the low-level visual component and the dorsal attention component across groups.

Together, findings related to functional connectivity changes provide evidence that the integration of distributed functional networks can support improvement in reading ability. Several studies reported effects involving connectivity in fronto-parietal networks (Richards et al., 2016) and anterior cingulate cortex (Horowitz-Kraus et al., 2019; Richards et al., 2016, 2017), which could reflect modulation of attention-related networks during reading. Optimal levels of connectivity among various networks are likely needed to support reading as intervention-related changes involved both increases and decreases in connectivity.

----------------------insert table 2 here------------------------------

3.1.6: Neural changes in response to intervention: White matter structure

Research on structural brain changes associated with intervention is quite limited. One study of white matter microstructure showed increased fractional anisotropy (FA; a measure of white matter integrity) in a left frontal tract in children with RD following a six-month intervention focused on word-level decoding (Keller & Just, 2009). These children showed significant gains in word and pseudoword reading scores, contrary to a group of children with RD who continued usual classroom instruction over the same period and showed no white matter changes. Regression analysis revealed a positive relationship between change in FA in the left frontal tract and pseudoword reading fluency, but a negative relationship was found between change in FA and sight word reading fluency.

Davis and colleagues (2010) examined relationships between improvement in reading and white matter structural connectivity among regions of the bilateral reading network measured after intervention. Increased reading scores on timed and untimed word and pseudoword reading measures were positively correlated with connection strength (symmetrized connection ratio based on number of streamlines connecting pairs of seed regions) between the left IPL and left insula. Additionally, improvement in timed word reading was positively associated with connection strength between left IFG and left inferior frontal sulcus. Negative correlations were found between change in untimed pseudoword reading and connection strength between left thalamus and left STG/STS, between right insula and right STG/STS, and between right IFG and right thalamus.

Richards and colleagues (2017; 2018) examined changes in white matter structure in children with dysgraphia, children with RD, children with oral language difficulties, and a control group with typical reading and writing skills associated with a reading and writing intervention). Decreases in diffusivity (including radial, axial, and mean diffusivity) from pre- to post-intervention were reported across groups in corona radiata, superior longitudinal fasciculus, and additional superior and middle frontal regions (Richards et al., 2017). In addition, associations between functional connectivity and measures of white matter structure were shown at post-intervention, but not pre-intervention scanning (Richards et al., 2018). These included a significant positive correlation between axial diffusivity in left superior frontal white matter and local functional connectivity in right IFG during word reading, a significant positive correlation between mean diffusivity in left superior corona radiata and local functional connectivity in left middle frontal gyrus during sentence reading, and a significant positive correlation between mean diffusivity in left anterior corona radiata and right middle frontal gyrus.

In a recent study, Huber and colleagues (Huber et al., 2018) examined changes in white matter microstructure over the course of an eight-week reading intervention. Participants were scanned prior to intervention, at two times during intervention (approx. 2-3 weeks apart), and after intervention. Findings in tracts of interest defined *a priori* showed that mean diffusivity in the left arcuate fasciculus and inferior longitudinal fasciculus decreased with increasing hours of intervention. Changes in FA showed the opposite effect, with increases in the same regions with increasing hours of intervention. Further analysis showed a significant quadratic effect of FA in the left arcuate fasciculus. Changes in mean diffusivity were specific to the intervention group. Over the course of intervention, decreasing mean diffusivity in the left inferior longitudinal fasciculus was associated with increasing reading scores. Notably, the trajectories of tract development in the intervention group did not match the trajectories that would be expected if the tract structure became more similar to TD controls. Contrary to the effects in the arcuate fasciculus and inferior longitudinal fasciculus, mean diffusivity in the posterior corpus callosum showed a stable positive association with reading ability over time. Exploratory analyses of additional tracts revealed significant correlations between change in mean diffusivity in the left thalamic radiation, right thalamic radiation, left corticospinal, right corticospinal, left cingulum, left inferior fronto-occipital, and right arcuate tracts and change in reading scores. No significant correlations were found between FA development and improvement in reading.

The above studies provide evidence to link white matter plasticity to reading intervention. Increases in FA and decreases in mean diffusivity and radial diffusivity could index development of more efficient white matter pathways to support communication among distant cortical and sub-cortical structures involved in reading. Associations between structural connectivity changes and reading performance were consistent with networks that showed functional connectivity changes with reading intervention (Davis et al., 2010). Furthermore, direct associations between white matter structure and functional connectivity (Richards et al., 2018), and incremental change in white matter microstructure over the course of reading intervention (Huber et al., 2018) give insight to plausible mechanisms of change in the brain networks that support reading.

3.1.7: Neural changes in response to intervention: Gray matter structure

Only a few studies have reported intervention-related changes in gray matter (GM) structure to date. In one study, increases in GM volume were observed after an eight-week multisensory remediation program in a small sample of children with RD (Krafnick et al., 2011). Specifically, GM volume increases relative to initial testing in left anterior OT extending into hippocampus, bilateral precuneus, right hippocampus, and right cerebellum were present immediately following intervention, and remained after an 8-week null period, at which time an additional cluster of increased GM volume was observed in the right caudate. Nominal positive correlations between change in GM volume and change in behavioral scores were identified for phonological awareness in the left precuneus and for pseudoword reading in the right cerebellum, providing preliminary evidence to link morphometric changes in the brain to intervention response.

More recently, Romeo and colleagues examined changes in cortical thickness in children who participated in a summer reading intervention program (Romeo et al., 2017). On average, the intervention group did not show improvement in reading scores across the intervention period, but rather maintained reading performance, while the waiting control group showed decreases in reading scores. When considering individual participants, 20 children in the intervention group showed improvement in reading scores across the intervention (responders) and 19 declined in reading scores (non-responders). Interestingly, intervention responders had lower socioeconomic backgrounds than non-responders. Change in cortical thickness from pre- to post-intervention differed between responders and non-responders, with greater thickening in responders in bilateral middle-inferior temporal cortex, IPL, precentral cortex, and paracentral/posterior cingulate cortex, right STG-insula, and left middle temporal regions. In addition, lower socioeconomic status and greater RD severity were correlated with cortical thickening.

Together, these studies link changes in GM structure to reading intervention, with effects in hubs of the reading network as well as sub-cortical and right hemisphere regions (Krafnick et al., 2011; Romeo et al., 2017). Importantly, both of these studies reported effects related to response to intervention, though in different regions.

-----------------------------insert table 3 here----------------------------------

3.2 Neural predictors of response to intervention

Initial patterns of brain activation and functional connectivity during reading-related tasks have been examined in relation to reading outcomes in an effort to identify neural predictors of children's responsiveness to intervention. With regard to regional activation, several studies show associations between baseline measures of brain activation and improvement in reading skills over the course of intervention in children and adolescents with RD. Simos and colleagues (2005) used MEG to examine brain activation in a bilateral set of reading-related regions of interest (ROIs) during letter-sound naming and pseudoword reading in beginning readers. Children received intervention if they were classified as high-risk for RD based on a screener of pre-literacy skills in kindergarten. Those who subsequently improved in reading skills showed greater activation prior to intervention in bilateral IFG relative to the low-risk group that was not assigned to intervention, and more symmetric bilateral activation across time points relative to the left-lateralized pattern in the low-risk group.

In another set of MEG studies, Rezaie and colleagues (2011b) found that greater baseline activation in left STG and IPL regions during pseudoword reading was associated with greater improvement in word and pseudoword reading fluency after a nine-month intervention in adolescents with RD. In a second article, the authors reported on pre-intervention activation during printed word recognition in an overlapping sample (Rezaie et al., 2011a). Activity in left middle temporal, left ventral OT, and right parahippocampal regions during the printed word task was associated with improvement in word reading efficiency. This pair of studies shows that greater initial activation in hubs of the reading network is associated with greater subsequent improvement in reading and underscores the specialization of pathways supporting pseudoword reading (dorsal route via STG/IPL) versus word reading (ventral route via OT) (Coltheart et al., 2001).

Turning to fMRI findings, Farris and colleagues found that improvement in word and pseudoword reading over a 2-year intervention period positively correlated with pre-intervention activation during phonological processing in bilateral IFG, left middle frontal gyrus, right medial frontal lobe, and left insula (Farris et al., 2016). In another study, participants who responded to an intensive 15-hour intervention showed greater pre-intervention activation during a lexical decision task relative to non-responders in a broad set of regions including right temporal pole/STG, hippocampus, middle temporal gyrus, left precentral gyrus, postcentral gyrus, superior frontal gyrus, IPL, inferior occipital gyrus, and bilateral OT, cerebellum and parahippocampal gyrus. Regression analysis revealed that activation in left IPL, right cerebellum, and right OT was positively associated with improvement in reading (Barquero, 2015). In a recent publication from an overlapping sample, Aboud and colleagues (2018) reported that pre-intervention activation in the left IPL during a lexical decision task also positively correlated with improvement in reading over the intervention period.

Another recent study showed differences in activation during a sentence comprehension task prior to intervention, such that subsequent responders had higher activation in right OT regions relative to non-responders and reduced activation in precuneus relative to non-struggling readers (Nugiel et al., 2019). On the other hand, non-responders had reduced activation in right posterior middle temporal gyrus and left postcentral white matter relative to non-struggling readers.

Relationships between activation prior to training and later reading outcomes have also been reported in pre-readers. Karipidis and colleagues (2018) found that pre-reading children at familial risk for RD who initially had greater activation in left OT cortex during an implicit grapheme-phoneme audiovisual target detection task showed greater improvement in pseudoword reading fluency after a single session of artificial letter training (GraphoGame; Lyytinen et al., 2009) and a 5-7-month reading instruction period. In addition, children who had reading fluency scores in the average range after intervention showed increased activation in the left planum temporale, a structure involved in phonological processing, during the congruent condition (artificial grapheme matched to trained phoneme) relative to the incongruent condition (artificial grapheme mis-matched with trained phoneme). In contrast, children with poor reading fluency outcomes showed a trending effect in the opposite direction in the left planum temporale. These findings may reflect early specialization in the left OT cortex for visual processing of orthographic information, and engagement of the left planum temporale for audio-visual integration in subsequent typical readers, alongside impaired audio-visual integration in subsequent poor readers.

Along with standard functional activation approaches, functional connectivity prior to intervention has been examined as a predictor of reading outcomes. In one study, pseudoword reading improvement positively correlated with initial connectivity between the left IFG and right medial frontal lobe (Farris et al., 2016). In addition, Aboud and colleagues (2018) investigated associations between improvement in reading and functional connectivity among reading and executive functioning networks prior to intervention. Examination of connectivity patterns revealed that children who responded to intervention had greater connectivity between left middle temporal and right inferior frontal regions relative to non-responders and typical readers, and connectivity positively correlated with improvement in reading. Additionally, activation in the left dorsolateral prefrontal cortex, a region known to be involved in executive functioning and higher order cognition, positively correlated with connectivity among left temporal, left parietal, and bilateral inferior frontal regions, and these activation-connectivity associations were stronger in children whose reading performance improved.

Together, these studies provide evidence that greater activation in regions of the posterior left hemisphere reading network (temporal, IPL, and OT regions) prior to intervention is related to greater improvement in reading after intervention. In addition, there is some support for links between reading gains and pre-intervention activation in the right parahippocampal gyrus (Rezaie, et al., 2011a, Barquero, 2015) along with cerebellum, hippocampus, and right hemisphere regions homotopic to those of the left-hemisphere reading network (Barquero, 2015; Nugiel, et al., 2019; Simos, et al. 2005). Further, emerging research using functional connectivity approaches provides evidence that greater initial connectivity among frontal regions that support executive function and left hemisphere regions associated with language and reading may also promote response to intervention (Aboud, et al. 2018; Farris, et al. 2016).

------------------------------------insert table 4 here-------------------------------------

3.3. META-ANALYSIS RESULTS

Our quantitative meta-analysis aimed to characterize the relationship between reading intervention and brain activation changes in 8 studies that met inclusion criteria (Table 5). The analysis included data from a total of 151 participants. The median age at start of intervention across studies was 9.95 years (range: 5.6-44 years). The median duration of intervention across studies was 8 weeks. The meta-analysis did not yield any significant effects (threshold: p_{FWE} = .05)

Table 5.

Studies included in the quantitative meta-analysis.

Study	N	Task contrast	Analysis Contrast	Voxel-wise Threshold	# of foci	Mean age at start of study (years)	Weeks of Intervention
Eden, 2004	19	sound deletion > word repetition	Post vs. Pre for intervention group > non-intervention RD group	$p < .001$, unc.	15	44	8
Gebauer, 2012	10	pseudoword lexical decision > fixation	Post vs. Pre in training group	z > 2.0	7	11.5	5
Heim, 2015	33	word/pseudoword reading > baseline	Post vs. Pre in RD intervention group	$p < .05$, FWE-corrected	2	10	4
Partanen, 2019	29	words > symbols	Poor readers > Good readers at Post vs. Pre	z > 2.3	1	8.555	12
Richards, 2006, *J.NL*	8	word pair spelling decision > letter string matching	Post vs. Pre in orthographic treatment group	z > 2.4	5	10.808	3
Shaywitz, 2004	25	audio-visual letter identification > baseline	Follow-up > Pre in RD experimental intervention group	$p < .05$	7	7.9	32
Temple, 2003	20	letter rhyming > letter matching	Post vs. Pre in RD group	$p < .005$, unc.	14	9.9	8
Yamada, 2011	7	one-back letters > false fonts	Post vs. Pre in at-risk group	z > 2.33	41	5.6	NR

N indicates number of participants included in the contrast of interest.
Mean age reported for participants included in the contrast of interest.
NR=Not Reported

3.4 Assessment of Risk to Internal Validity

We assessed risk to internal validity of primary studies included in our systematic review, and the results of our assessment are summarized in Figure 2 using a “traffic light” plot that indicates the rating for each individual study in each quality appraisal domain. A summary of quality ratings across studies on each domain is presented in Figure 3.

Figure 2. Summary of the assessment of risk to internal validity by study and quality domain. Rows indicate each primary study by first author and year. Columns indicate each quality appraisal domain upon which studies were evaluated. Colors indicate raters’ judgements: Red = No/Poor, Yellow = Partial/Fair, Green = Yes/Good, Blue = No information reported, Gray = Not applicable. Quality domains are listed below table.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13
Aboud, 2018	X	+	?	?	+	X	X	+	?	+	-	?	-
Aylward, 2003	X	X		?	?	X	+	+	?	+	+	?	+
Bach, 2013	+	+	+	?	+	X	X	+	+	+	-	?	+
Barquero, 2015	+	+	+	?	+	?	?	-	+	+	-	?	+
Brem, 2010	+	+	+	?	+	?	?	+	+	+	-	?	+
Davis, 2010	-	X		?	+	X	?	+	+	+	-	?	+
Davis, 2011	-	X		?	+	X	?	+	?	+	-	?	+
Eden, 2004	+	X		?	+	+	+	+	+	+	+	?	+
Farris, 2016	X	X		?	-	?	?	+	?	+	-	?	+
Gebauer, 2012	+	X		?	-	X	?	+	?	-	+	?	+
Heim, 2015	+	-	+	?	-	?	?	+	?	-	+	?	-
Horowits-Kraus, 2014	X	X		?	+	?	?	+	+	+	+	?	+
Horowitz-Kraus, 2019	X	X		?	+	-	?	+	+	+	X	?	-
Huber, 2018	X	X		?	-	X	X	+	-	+	+	?	+
Karipidis, 2017	X	X		?	+	+	?	+	?	+	-	?	+
Keller and Just, 2009	+	+	+	?	+	?	?	+	?	+	+	?	+
Koen, 2018	+	+	+	?	?	X	?	+	+	+	-	?	+
Krafnick, 2011	X	X		?	+	+	?	+	?	+	+	?	+
Meyler, 2008	X	X		?	+	X	+	+	?	+	+	?	+
Nuigel, 2019	X	+	+	?	+	+	?	-	?	+	+	?	-
Odegard, 2008	X	X		?	?	?	?	+	+	+	-	?	+
Partanen, 2019	X	X		?	+	X	X	+	-	+	+	?	-
Rezaie, 2011 (Dev N)	X	X		?	+	+	+	+	?	+	+	?	+
Rezaie, 2011 (J INS)	X	X		?	+	+	+	+	?	+	+	?	+
Richards, 2006 (J NL)	X	+	+	?	+	?	?	+	?	+	+	?	+
Richards, 2006 (Dev N)	X	X		?	-	X	-	+	+	+	-	?	+
Richards, 2007	+	+	+	?	-	+	?	-	?	+	+	?	+
Richards, 2008	X	X		?	-	?	?	?	?	+	-	?	+
Richards, 2016	X	X		?	+	+	?	+	?	+	+	?	+
Richards, 2017	X	X		?	+	+	+	+	+	+	-	?	+
Richards, 2018	X	X		?	?	+	X	+	-	+	-	?	-
Romeo, 2017	+	-	+	+	+	+	+	+	?	+	+	?	+
Shaywitz, 2004	+	X		?	+	-	?	+	+	+	-	?	+
Simos, 2002	X	X		?	+	?	?	+	+	+	-	?	+
Simos, 2005	X	X		?	+	-	?	?	?	+	+	?	+
Simos, 2007 (Neu)	X	X		?	+	X	?	+	?	+	+	?	+
Simos, 2007 (J LD)	X	X		?	+	-	?	+	?	+	+	?	-
Temple, 2003	X	X		?	+	?	?	+	+	+	+	?	+
Yamada, 2011	X	X		-	+	X	-	+	?	+	+	?	+

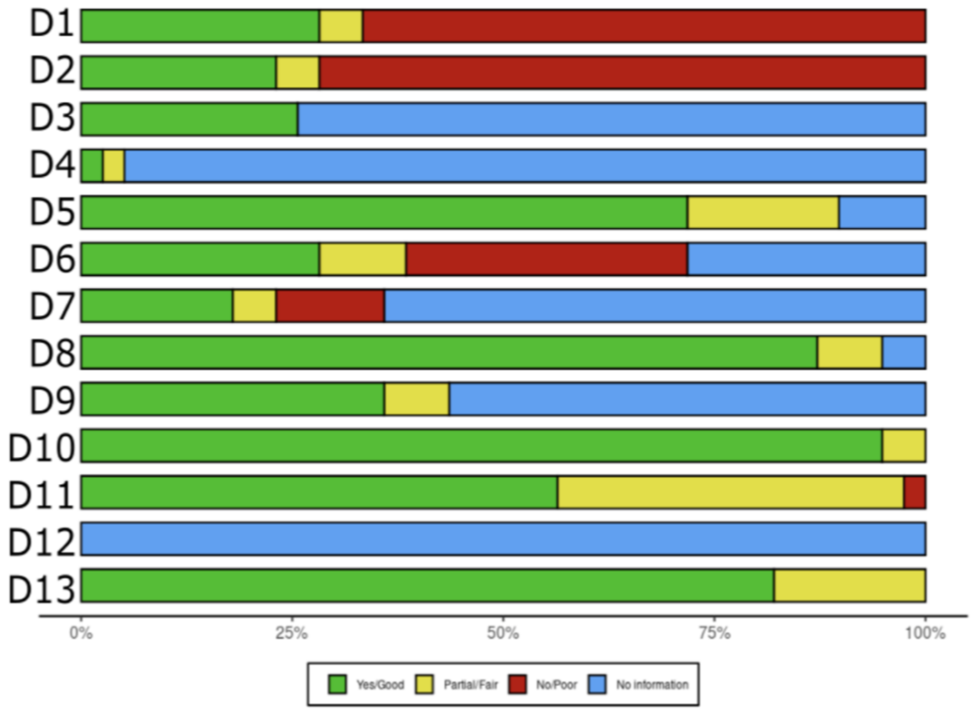
Judgement

- X No/Poor
- \- Partial/Fair
- \+ Yes/Good
- ? No information
- Not applicable

Quality Domains

D1: Was the treatment group compared to a control group of individuals with RD who did not receive reading intervention?
D2: Was the study described as randomized?
D3: Was the method of randomization adequate?
D4: Were the people assessing the outcomes blinded to the participants' group assignments?
D5: Were the groups similar at baseline on important characteristics that could affect outcomes?
D6: Was the overall drop-out rate from the study at endpoint 20% or lower?
D7: Was the differential drop-out rate (between treatment groups) at endpoint 15% or lower?
D8: Was there high adherence to the intervention procedure for each treatment group?
D9: Was there high adherence to the intervention duration for each treatment group?
D10: Were behavioral outcomes assessed using valid and reliable measures, implemented consistently across all study participants?
D11: Were imaging outcomes assessed using valid and reliable measures, implemented consistently across all study participants?
D12: Did the authors report that the sample size was sufficiently large to be able to detect a difference in the main outcome between groups with at least 80% power?
D13: Were all participants analyzed in the group to which they were originally assigned?

Figure 3. Summary of quality ratings by domain. Color bars indicate percent of studies with each rating for each domain. Domains are listed below Figure 3.



Quality Domains

D1: Was the treatment group compared to a control group of individuals with RD who did not receive reading intervention?

D2: Was the study described as randomized?

D3: Was the method of randomization adequate?

D4: Were the people assessing the outcomes blinded to the participants' group assignments?

D5: Were the groups similar at baseline on important characteristics that could affect outcomes?

D6: Was the overall drop-out rate from the study at endpoint 20% or lower?

D7: Was the differential drop-out rate (between treatment groups) at endpoint 15% or lower?

D8: Was there high adherence to the intervention procedure for each treatment group?

D9: Was there high adherence to the intervention duration for each treatment group?

D10: Were behavioral outcomes assessed using valid and reliable measures, implemented consistently across all study participants?

D11: Were imaging outcomes assessed using valid and reliable measures, implemented consistently across all study participants?

D12: Did the authors report that the sample size was sufficiently large to be able to detect a difference in the main outcome between groups with at least 80% power?

D13: Were all participants analyzed in the group to which they were originally assigned?

4. DISCUSSION

Our literature review illustrates that there is no single route to reading remediation; there is evidence for changes in both left hemisphere hubs of the reading network and homotopic right hemisphere regions, as well as in frontal and subcortical structures. In fact, some studies show a mix of effects within the same participants (Eden et al., 2004; Gebauer et al., 2012; Horowitz-Kraus et al., 2014; Meyler et al., 2008;

Temple et al., 2003), suggesting that intervention may enhance activation of the typical reading network *and* engage alternate cognitive systems to support reading. This is consistent with our hypothesis that reading remediation would involve both changes in activation in the left hemisphere reading network and homotopic right hemisphere regions.

Our quantitative meta-analysis failed to show any significant converging effects across 8 studies. Our null effects contradict a previous meta-analysis of reading intervention studies that showed significant effects in the left thalamus, right insula/IFG, left IFG, right posterior cingulate, and left middle occipital gyrus (Barquero et al, 2014). Though both the Barquero et al. meta-analysis and our meta-analysis included 8 studies, they did not entirely overlap in the studies included. We used slightly stricter inclusion criteria to achieve a more homogeneous set of studies (though there were still differences, as described below). The Odegard et al, 2008 and Meyler et al, 2008 studies included in the Barquero et al, 2014 meta-analysis were omitted from our analysis as they only reported post-intervention group contrasts, but not pre-post-intervention change. In studies using only a post-intervention design, intervention effects cannot be validly discerned and may be obscured due to unknown confounds. We also included two studies that were published after the Barquero et al meta-analysis: Heim et al. (2015) and Partanen et al. (2019). Factors that contribute to the heterogeneity observed within the studies considered in our meta-analysis are discussed below.

4.1 BRAIN CHANGES ASSOCIATED WITH INTERVENTION

4.1.1 Changes in the left hemisphere reading network hubs

Many studies show "normalization" of brain activation in response to intervention, evidenced by increased engagement of left hemisphere reading network regions (especially STG, IPL and OT), and/or post-intervention activation that no longer differs from TD peers (Aylward et al., 2003; Davis et al., 2011; Karipidis et al., 2018; Rezaie et al., 2011; Richards et al., 2006; Richards & Berninger, 2008; Shaywitz et al., 2004; Simos et al., 2002; Simos, Fletcher, Sarkari, Billingsley-Marshall, et al., 2007; Simos, Fletcher, Sarkari, Billingsley, et al., 2007). These findings indicate that aberrant function in the typical reading network can be recovered through intervention to a certain extent, and normalization of brain activation has been observed over as little as 2-3 weeks of intervention (Aylward et al., 2003; Richards, Aylward, Berninger, et al., 2006; Richards & Berninger, 2008). However, it is important to note that the elimination of brain differences between RD and TD groups does not necessarily indicate complete behavioral remediation, and in most cases, RD groups continued to perform more poorly on behavioral measures of reading after intervention.

With respect to specific regions, hypoactivation in left TP regions including STG and IPL has been consistently linked to RD (Maisog et al., 2008; Richlan et al., 2009), and there is evidence of increased activation in these regions following intervention (Meyler et al., 2008; Simos et al., 2002; Simos, Fletcher, Sarkari, Billingsley-Marshall, et al., 2007; Simos, Fletcher, Sarkari, Billingsley, et al., 2007; Temple et al., 2003). Explicit training may facilitate the development of STG and IPL systems involved in phonological processing and integrating phonological and orthographic information. Evidence from functionally illiterate adults further supports this possibility, as increases in gray matter volume in left parietal regions and STG were positively correlated with improvement in reading following intensive intervention (Boltzmann et al., 2017).

The left ventral OT cortex (including the fusiform gyrus, the location of the putative visual word form area [VWFA]) is another key hub of the typical reading network that shows reduced activation in RD (Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan, Kronbichler, & Wimmer, 2009). Several studies showed increased activation in left OT following intervention (Eden, Jones, et al., 2004; Heim et al., 2015; Shaywitz et al., 2004). This could reflect specialization of that region to respond selectively to print. Alternatively, the functional role of activation increase in the left OT could be interpreted in the context of the recently proposed “multiplex model of VWFA function” (Chen et al., 2019). According to this model, the VWFA has discrete structural and functional connections to language and attention networks, and connectivity with each of these networks is uniquely associated with behavioral performance on related tasks. We cannot draw conclusions about whether increased left OT activation is associated with increased connectivity with language and/or reading networks based only on localized functional activation. Recent evidence of changes in functional connectivity with reading intervention revealed that increased reading speed was associated with increased connectivity between bilateral OT regions and the dorsal attention network (Horowitz-Kraus et al., 2019). These findings indicate that integration of the left OT with attentional networks may facilitate reading, and could provide an alternate route to reading efficiency that does not rely on strong connectivity with the temporal language network. Importantly, the intervention program used in this study focused on reading fluency, so examinations of left OT connectivity with language and attention networks in the context of more common phonics-based reading intervention are needed.

Intervention-related increases in activation and connectivity have also been observed in the left IFG (Aylward et al., 2003; Davis et al., 2010; Richards et al., 2018; Shaywitz et al., 2004; Temple et al., 2003; Yamada et al., 2011). The left IFG is consistently involved in reading and phonological analysis in typically developing readers (Martin, Schurz, Kronbichler, & Richlan, 2015; Taylor, Rastle, & Davis, 2013). Several meta-analyses show that RD groups exhibit reduced activation in the left IFG along with hyperactivation in proximal regions such as the left anterior insula, pre/post-central gyrus, and subcortical regions including thalamus and basal ganglia relative to TD groups (Hancock et al., 2017; Maisog et al., 2008; Richlan et al., 2009). This illustrates a complex pattern of activation in anterior left hemisphere regions in RD which must be considered carefully with regard to response to reading intervention. The role of left IFG in both phonological analysis and articulatory recoding complicates the interpretation of left IFG activation changes with reading intervention (Hancock et al., 2017). Increased activation in the left IFG could be associated with increased engagement of this region for phonological analysis and indicate normalization of function. Alternatively, hyperactivation of the bilateral IFG/insula in response to intervention could reflect increased reliance on compensatory strategies involving attention and working memory (Shaywitz et al., 2002) or articulatory coding (Hancock et al., 2017). Carefully designed fMRI tasks are needed to dissociate frontal activation associated with phonological analysis from that associated with articulatory coding and general cognitive processes in future intervention studies.

Numerous studies show intervention-related effects in the main hubs of the reading network, and increased activation in these regions has been linked to individual differences in reading improvement (Bach et al., 2013; Davis et al., 2011; Horowitz-Kraus et al., 2014; Simos et al., 2007 Neurospychology). Nonetheless, it remains unclear whether the post-intervention patterns of activation in individuals with RD are truly "normalized" as these regions may be engaged differently in individuals with RD. Moreover, our quantitative meta-analysis and many empirical studies show effects both within *and* outside of the reading network, indicating that effects in left hemisphere regions are likely integrated with compensatory mechanisms encompassing hubs across the brain.

4.1.2 Intervention-related changes in the right hemisphere: IFG

Increased activation in the right IFG following intervention was among the most commonly reported functional effects (Meyler et al., 2008; Odegard, Ring, Smith, Biggan, & Black, 2008; Richards, Aylward, Berninger, et al., 2006; Richards et al., 2017; Shaywitz et al., 2004; Temple et al., 2003; Yamada et al., 2011), and elevated activation in the right IFG was positively associated with improvement in reading and phonological processing (Odegard et al., 2008; Temple et al., 2003). Although our meta-analysis did not yield significant effects, an earlier meta-analysis that used slightly different inclusion criteria identified a significant intervention-related effect in the right insula/IFG. Furthermore, several studies indicate that activation and connectivity of the right IFG prior to intervention predicts later gains in reading ability (Aboud et al., 2018; Farris, Ring, Black, Lyon, & Odegard, 2016; Hoeft et al., 2011). Together, these pre- and post-intervention findings support the hypothesis that intervention enhances activation in regions that are responsive to reading tasks prior to intervention, and children with greater initial activation of these regions are predisposed to greater engagement of such regions with intervention. Regarding cognitive mechanisms, activation of the right IFG is thought to support reading through its involvement in articulatory recoding (Hancock et al., 2017; Pugh et al., 2001), working memory and attention (Shaywitz et al., 2002).

As noted in the previous section, different sub-divisions of the IFG and anterior insula may be involved in distinct cognitive mechanisms, and individuals may show distinct patterns of response within these regions.

Importantly, the right IFG is engaged to some extent in typical readers, and it is unclear whether compensatory engagement of this region in RD involves hyperactivation relative to typical readers (Horowitz-Kraus et al., 2014), or increased engagement to match the level of typical readers (e.g., Richards, Aylward, Berninger, et al., 2006). For example, one set of intervention studies showed changes in activation of right hemisphere regions to levels that no longer differed from TD groups, which may indicate optimization of reading-related activation in both hemispheres (Aylward et al., 2003; Richards, Aylward, Berninger, et al., 2006; Richards & Berninger, 2008). One possibility is that there is an optimal level of right IFG activation for efficient reading, and over- or under-activation of this region may be detrimental. Furthermore, RD is known to be heterogeneous in nature (Pennington, 2006), and the ability to engage the right IFG as a compensatory mechanism may be available to children who exhibit specific sub-types of RD. Some children with RD may be able to incorporate working memory and attentional functions supported by the right IFG into a compensatory reading network (e.g., involving connectivity between right IFG and IPL, Richards, et al., 2016); others may have difficulties integrating these cognitive processes with the auditory-visual mapping processes that are important to reading, resulting in a persistently inefficient reading network. Further research is needed to characterize the role of right IFG function in children who do and do not respond to intervention, and who represent various behavioral profiles within RD.

4.1.3 Intervention-related changes in the right hemisphere: STG & IPL

In addition to effects in the right IFG, reading intervention has been associated with changes in posterior right hemisphere homologues of the reading network. Activation in the right STG and IPL has been observed following phonological intervention in adults with RD (Eden et al., 2004), and activation of the right TP cortex has been positively associated with reading ability in adults with RD (Rumsey et al., 1999). A recent study in children showed that increasing activation in the right IPL during printed word rhyming was associated with improvement in decoding skills in children with RD (Partanen et al., 2019).

With regard to changes in brain structure, Romeo and colleagues (2017) found that intervention responders showed greater cortical thickening in bilateral IPL and right STG, among other regions.

On the other hand, some studies showed changes that shifted away from these posterior right hemisphere regions, or showed that activation in these regions was associated with persistent reading difficulties. Simos and colleagues (2002) observed a reversal from right-lateralized posterior STG activation prior to intervention to left-lateralized posterior STG activation after intervention. Gebauer and colleagues (2012) found that activation increased in right temporal and lateral occipital regions in the RD waiting control group, but not the training group, during a lexical decision task, and suggested that this may reflect compensation via serial grapheme-phoneme decoding. They also reported a negative correlation between improvement in spelling and activation increase in right posterior regions and the cerebellum, indicating that engaging these regions may be inefficient. These findings are in line with several studies that have shown increases in right temporal and IPL regions in children who do *not* respond well to treatment relative to those who do (Odegard et al., 2008; Simos, Fletcher, Sarkari, Billingsley, et al., 2007).

To further complicate this picture, an earlier study showed that children with RD had reduced activation in the right IPL during sentence comprehension before intervention and showed no activation difference in this region from their TD peers at post-intervention imaging; this effect was driven by reduced activation in bilateral IPL regions in the TD group, rather than increased activation in the RD group (Meyler et al., 2008). This finding suggests that TD children exhibit a developmental reduction in IPL activation during sentence reading while children with RD continue to show a similar level of activation in this region across the intervention period that could reflect less efficient processing and/or developmental delay. In contrast, Horowitz-Kraus and colleagues (2014) reported that the degree of activation increase from pre- to post-intervention during lexical decision (words > pseudowords) in right hemisphere regions including IPL was positively correlated with reading gains in the TD group, but not the RD group.

The contradictory findings from these studies could arise in part from fMRI task differences (word-level decoding vs. sentence comprehension). This is an important consideration given that the contrast of sentence reading versus fixation baseline (Meyler et al., 2008) is likely to show significant activation in a much broader set of regions than a specified contrast of words versus pseudowords (Horowitz-Kraus, et al., 2014), which is likely to isolate activation related to word recognition and semantic retrieval, and largely remove effects of phonological decoding. The focus of intervention (phonics vs. fluency) and study design could also explain the discrepant findings. Notably, the TD group in the Meyler study received only standard classroom instruction, while the TD group in the Horowitz-Kraus study completed the same 4-week Reading Acceleration Program as the RD group. Thus, the pattern of change in Meyler and colleagues' TD group shows a general effect of reading development, while the pattern of change in the Horowitz-Kraus study likely reflects training-induced changes. Moreover, Horowitz-Kraus and colleagues suggested that activation increases in the right IPL could underlie phonological processing and/or working memory, and the RD group could differ from the TD group in the recruitment of these processes with the reading acceleration training. Together these studies highlight the complex set of factors that contribute to patterns of typical reading development and response to intervention. Carefully designed studies including control groups of children with and without RD who do and do not receive intervention are needed to distinguish training-specific effects from trajectories of typical and atypical reading development.

4.1.4 Intervention-related changes in the right hemisphere: OT

Our review reveals mixed evidence with regard to the relationship between right OT activation and improvement in reading ability. Several early studies that compared RD and TD groups linked persistent reliance on the right OT cortex with poorer reading ability (Shaywitz et al., 2002; Turkeltaub et al., 2003), and decreases in right OT activation have been shown following intervention (Shaywitz et al., 2004). In another study, intervention-related increases in activation in right OT regions were associated with less improvement in spelling (Gebauer, et al., 2012). Hyperactivation of the right OT cortex in individuals with RD could reflect a developmental failure of the left OT cortex (putative visual word form area) to specialize function for print processing, leading to use of a diffuse bilateral visual processing regions to read.

However, contradicting evidence from longitudinal research points to a beneficial role of right OT activation. Aylward and colleagues (2003) found that activation in the right OT during morpheme mapping increased with intervention in the RD group to a level that no longer differed from controls. Greater initial activation in the right OT has been related to better reading outcomes following general reading instruction (Hoeft et al., 2007) and intensive intervention (Nugiel et al., 2019), and post-intervention activation in the right OT was positively correlated with improvement in reading (Nugiel et al., 2019). The involvement of right OT activation may therefore be more nuanced than previously thought.

Temporal aspects of right OT activation that are not detectable through fMRI could better distinguish beneficial vs. detrimental activity in this region. Using MEG, which affords greater temporal precision, Simos and colleagues (2005) observed distinct effects related to the timing and the degree of neural activity in kindergarten children at differing behavioral risk of RD. The timing of activation in low-risk children and high-risk responders differed at initial testing in kindergarten, such that the high-risk group showed temporal overlap in OT and TP activity, while the low-risk group showed sequential activation onset in OT followed by TP regions. Both groups showed similar sequential patterns of activation after one year of systematic reading instruction. This change at the brain level tracks with the behavioral response to intervention in the high-risk group. With regard to degree of activation, children with low risk of RD had greater right OT activation following instruction than high-risk responders. This finding is contrary to what would be expected if right OT serves as a protective mechanism in high-risk children. Instead, it provides evidence for bilateral distribution of OT activation early in typical reading development. Note, however, that risk in this study was determined based on behavioral assessment at kindergarten, not family history, and protective factors may differ based on the source of risk. In a later study of children in grades two and three, Simos and colleagues (2007a) observed that better reading performance was associated with decreased onset latency of activity in a right OT region following intervention, which the authors interpreted as increased processing efficiency in this region.

Altogether, the role of the right OT cortex in response to reading intervention remains uncertain, and developmental changes in lateralization of function may underlie effects in different age groups.

Individual differences in risk and protective factors may also influence the capacity to incorporate right OT activation into an efficient processing stream for reading, or to disengage noisy, inefficient activity patterns in this area. Importantly, no brain structure works in isolation, and in the following section we turn to functional integration of reading and other cognitive processing networks as a potential mechanism for improving reading skills.

4.1.5 The role of general cognitive processes and inter-network connectivity

Individuals with RD may rely on general cognitive processes (e.g., executive function, attention, working memory, general learning mechanisms) to overcome deficits that impair more direct routes to reading (e.g., rapid phonological decoding, automatized word recognition) (Haft et al., 2016; Xi Yu et al., 2018). Neuroimaging evidence supports this hypothesis by linking intervention to changes in activation, structural and functional connectivity, and gray matter volume in frontal regions associated with higher order cognition and subcortical regions associated with general learning processes (Aboud et al., 2018; Barquero et al., 2014; Farris et al., 2016a; Horowitz-Kraus et al., 2019, 2014; Horowitz-Kraus, Toro-Serey, et al., 2015; Keller & Just, 2009; Krafnick et al., 2011; Meyler et al., 2008; Richards et al., 2018; Temple et al., 2003).

Importantly, greater connectivity among hubs of the reading network and regions associated with higher order cognition before intervention and increasing connectivity in these circuits over the course of intervention have been linked to better reading outcomes following intervention (Aboud et al., 2018; Horowitz-Kraus et al., 2019; Horowitz-Kraus, Toro-Serey, et al., 2015). These connectivity effects were observed both during reading tasks (Aboud et al., 2018; Horowitz-Kraus et al., 2019) and during rest (Horowitz-Kraus et al., 2015). Thus, a coupling between reading network hubs in the posterior left hemisphere and higher order cognitive mechanisms in the prefrontal cortex could facilitate improvement in reading ability. Connectivity among these networks may reflect engagement of cognitive strategies to analyze visual (print) input to overcome difficulties with automatic word recognition and/or rapid orthographic-phonological mapping. For example, Horowitz-Kraus and colleagues (2019) have linked increased connectivity among lower-level visual networks and higher-level attention and executive functioning networks following reading intervention to improvement in reading. The anterior cingulate cortex is a key structure in the executive functioning network that may facilitate attention and error monitoring (Horowitz-Kraus et al., 2014). Several other studies have shown intervention-related increases in activation in the anterior cingulate cortex (Horowitz-Kraus et al., 2014; Meyler et al., 2008; Richards et al., 2017; Temple et al., 2003; but see Richards et al., 2016). These findings are consistent with observations that classroom behavior ratings, including attention and cognitive control, predict response to intervention (Torgesen et al., 1999).

Connectivity among the reading network and executive functioning networks is commonly observed across good and poor readers, however these networks likely play different roles in TD and RD readers (Nicolson and Fawcett 2019). Early in typical reading acquisition, children rely on working memory, inhibitory control and other executive functioning mechanisms to support phonological decoding and word-reading skills (Welsh et al., 2010; Blair and Razza, 2007). Once a child "cracks the code", more efficient reading circuitry is established to support rapid decoding and automatic word recognition (Chyl et al., 2018, 2019). This allows skilled readers to utilize higher order cognitive mechanisms, like executive functions, to attend to the text, monitor comprehension, and integrate new information into their existing knowledge (Arrington et al., 2014; Sesma et al., 2009). In RD, impairments in the typical reading network impede rapid decoding and word recognition and force readers to persistently rely on alternative cognitive processes to support the decoding level of reading (Langer et al., 2019; Koyama et al., 2013). The result is diffuse, inefficient, and effortful processing of text that often blocks access to meaning and understanding. In order for these compensatory cognitive mechanisms to benefit individuals with RD, they must be integrated into efficient networks that can be engaged with minimal effort (Bailey et al., 2018). Systematic instruction and/or intervention may help people with RD to develop compensatory strategies and promote plasticity to build more efficient neural pathways for reading.

It is important to note that greater functional connectivity does not always indicate better functioning. Several studies have reported hyper-connectivity in RD relative to TD groups prior to intervention that is reduced following intervention, and in some cases decrease in functional connectivity was associated with improvement in reading (Richards, et al., 2016; Richards & Berninger, 2008). Similarly, Mohammadi and colleagues (2020) found hyper-connectivity in functionally illiterate adults prior to an intensive reading intervention, along with intervention-related increases *and* decreases in connectivity thought to reflect rewiring for a more efficient reading network. These findings highlight the complexity of retuning functional networks to optimize efficiency of processing. With regard to some functional connectivity reductions, de-coupling between reading regions and regions of the default mode network may be associated with improved task-oriented functioning (Koyama et al., 2013; Richards et al., 2016).

Complementary findings from diffusion-weighted imaging studies show that changes in structural connectivity and white matter integrity are also associated with improvement in reading performance (Davis et al., 2010; Huber et al., 2018; Keller & Just, 2009; Richards et al., 2017). Changes occurred in both positive and negative directions, indicating that rewiring of connections may be associated with building stronger connections in some pathways while reducing the connectivity in other, potentially inefficient, pathways. Notably, Huber and colleagues (2018) scanned children several times over the course of an 8-week intervention and reported significant changes in white matter microstructure after only 2-3 weeks of intensive intervention. This finding suggests that the brain can readily respond to intensive training to improve the efficiency of connections in the brain for successful reading. Associations between white matter microstructure and reading skill in three reading-related tracts showed deviation from the typical trajectory over the course of intervention, rather than normalization. This study design provides important insight to the trajectories of brain changes that occur over the course of intervention, but this challenging, costly design is seldom applied to study reading intervention. Future studies using frequent neuroimaging in various modalities over the course of instruction could be immensely beneficial for understanding the neural mechanisms of reading intervention.

4.1.6 Compensation via subcortical learning and memory mechanisms

Neuroimaging studies of reading intervention provide some support for compensatory mechanisms involving declarative memory that align with the procedural deficit hypothesis. Increased activation and gray matter volume have been observed in the hippocampus and adjacent medial temporal structures following intervention (Eden et al., 2004; Gebauer et al., 2012; Krafnick et al., 2011; Temple et al., 2003), which supports a compensatory role of these structures in supporting declarative memory strategies in place of procedural memory. For example, Gebauer and colleagues (2012) identified intervention-specific activation increases in bilateral parahippocampal gyri during processing of misspelled words. The parahippocampal cortex is a medial-temporal structure associated with spatial learning, episodic memory and processing contextual associations (Aminoff et al., 2013), and engagement of this region likely reflects declarative memory strategies.

Several studies reported changes in the basal ganglia that may be related to procedural learning deficits in RD. One intervention study showed decreased activation in the caudate nucleus, a structure of the basal ganglia associated with procedural learning, immediately following intervention and at follow-up (Shaywitz, et al., 2004). In this case, intervention did not recover typical activation for procedural learning, but further disengaged this region. Conversely, Meyler and colleagues (2008) found that activation in another basal ganglia region, the left putamen, was greater in the RD group relative to the TD group following intervention. This finding indicates that procedural learning mechanisms may be recovered with reading intervention.

The cerebellum is another key structure associated with procedural learning due to its involvement in the development of automatization (Nicolson et al., 2010). Gebauer and colleagues (2012) noted intervention-specific increases in cerebellum activation in addition to the increased activation observed in the parahippocampus. Thus, reading intervention may engage multiple learning systems within the context of a single intervention and task. Recently, Partanen and colleagues (2019) reported increased activation in the cerebellum during a spelling task following intervention in an RD group and general reading instruction in a TD group. The authors noted that the region of the cerebellum that showed this effect is associated with motor function and was likely activated due to button presses. Interestingly, both studies that reported intervention-related effects in the cerebellum found the effects during spelling-related tasks, which may indicate that the cerebellum is particularly involved in orthographic processing. It is important to note most studies of reading are designed to capture cortical activation and are not optimized to characterize cerebellar activation with great specificity. The cerebellar clusters reported in the Partanen (2019) and Gebauer (2012) studies are fairly large and include white matter and gray matter spanning several sub-regions of the cerebellum, making it difficult to interpret specific functional effects. Research targeting cerebellar structure and function is needed to clarify the role of this complex, multifaceted structure in reading and reading remediation.

4.2 NEURAL PREDICTORS OF RESPONSE TO INTERVENTION

Patterns of activation and features of brain structure that are uniquely associated with response to intervention may be useful for identifying how individuals are likely to respond to treatment. According to several studies, activation in posterior regions of the typical left-lateralized reading network prior to intervention was positively associated with improvement in reading (Aboud et al., 2018; Karipidis et al., 2018; Rezaie et al., 2011b, 2011a). Based on these findings, children who started out with more "typical" patterns of activation improved the most with intervention; though Karipidis and colleagues (2018) noted a dissociation such that greater left OT activation was present in subsequent good readers, but greater left superior temporal (planum temporale) activation was present in subsequent poor readers. This could reflect a developmental delay or failure to specialize left OT activation for automatic letter/word recognition, along with enduring reliance upon the temporal phonological processing system to decode simple grapheme-phoneme correspondences in those with poorer reading outcomes. In one study, lower activation in the left IFG prior to intervention was associated with greater improvement in reading skills, pointing to distinct effects in posterior and anterior left hemisphere regions (Partanen et al., 2019). Children who show initial engagement of the typical left hemisphere reading network prior to intervention may be able to strengthen this network with training, and thus improve in reading via a pathway that is similar to that of typical readers.

On the other hand, there is also evidence that greater initial engagement of right hemisphere and/or bilateral regions and interhemispheric connectivity are associated with greater improvement following reading intervention (Aboud et al., 2018; Barquero, 2015; Farris et al., 2016; Nugiel et al., 2019). These findings are consistent with earlier evidence that activation in right hemisphere homologues of the reading network was associated with better reading outcomes in struggling readers following subsequent classroom instruction and community intervention (Hoeft et al., 2011, 2007). Early bilateral activation and widespread bilateral connectivity prior to intervention may facilitate plasticity in the network to recruit a more bilateral compensatory network for reading in children who have impairments in the left posterior reading network.

Indeed, some neural characteristics present in children at risk of RD who go on to have typical reading skills have been proposed as protective factors that facilitate recruitment of alternative brain networks to support reading development (Yu et al., 2018). Children with a family risk of RD who went on to have typical reading skills have been shown to have greater activation in right IFG at the pre-reading stage relative to typical readers with no family risk (Yu et al., 2020), and more rapid development of white matter in the right superior longitudinal fasciculus (SLF) relative to at-risk children who had poor reading outcomes (Wang et al., 2017). Children who had typical reading outcomes despite elevated risk for RD based on weak pre-literacy skills had greater white matter organization in the right SLF at kindergarten relative to at-risk children with poor reading outcomes (Zuk et al., 2020). Together, these findings provide evidence that early recruitment of right hemisphere mechanisms may facilitate the development of alternative pathways to support reading, and with adequate instruction, protect against severe reading difficulties.

4.3 Assessment of internal validity of primary studies

Though our review illustrates the promise of neuroimaging studies to provide insight to the neurobiological changes associated with reading intervention, our assessment of the quality of the primary studies raised some concerns that limit our ability to draw conclusions about causal effects. There were several domains in which most studies we reviewed were rated strongly: similarity of groups at baseline (D4), adherence to intervention procedures (D7), valid and reliable assessment of behavioral outcomes (D9), and analysis of participants according to the groups to which they were initially assigned (D13). However, there were also several domains for which most studies did not report any information: blinding of the researchers who assessed outcomes (D3), differential drop-out rates between groups (D6), whether sample sizes were sufficient to detect effects with at least 80% power (D11). Future studies should take care to report on these factors which may influence the validity of the research and impact of the findings.

The chief concern raised by our quality assessment was the lack of adequate control groups and failure to provide information on randomization (i.e., domains D1, D2 and D3 of quality appraisal). Only 30% of studies in our review reported findings comparing an RD intervention group to a control group that did not receive reading intervention. This represents a major problem in this literature as a whole because without adequate control groups it is impossible to determine whether the reported brain changes were *caused* by the intervention or by maturation or the experience of persistent reading difficulties. The use of waiting control groups provides one approach to designing studies with adequate comparison groups, but still provides the opportunity for all participants to receive the potentially beneficial intervention at the conclusion of the study (e.g. Gebauer et al., 2012; Romeo et al., 2017). The possibility that the additional attention received during intervention is driving post-intervention results presents another challenge for the interpretation of results in brain and behavior. To help disambiguate such effects, studies can be designed to compare experimental reading intervention to a control (non-reading) intervention so that attention children receive is matched, and only the instructional component is manipulated.

Many studies compared RD intervention groups to TD controls, which is problematic because this results in groups that are not matched prior to treatment (D5), making group comparisons difficult to interpret. In addition, lack of randomization may result in biased treatment assignments. As such, we cannot be sure to what extent any confounding variables may have affected the treatment outcomes and related brain changes. We also noted inconsistency in how comparison groups were defined across studies. In some studies, reading intervention groups were compared that groups that received "business as usual" reading instruction or groups that received a non-reading intervention, while in other studies, RD intervention groups were compared to TD groups that did not receive intervention. Results can be driven by effects within the comparison groups, so this heterogeneity may partly explain the lack of convergence across studies. Interpretation and synthesis of effects across studies is complicated by these methodological choices, especially since there are cases in which reading intervention groups were collapsed with business-as-usual reading instruction groups for neuroimaging analysis (e.g., Nugiel et al., 2019), and others in which intervention groups were compared to business-as-usual instruction control groups (e.g. Shaywitz et al, 2004). Thus, it is crucial for researchers to clearly lay out the aims and rationale of their studies and to distinguish studies that aim to evaluate the neural changes resulting from a specific intervention from those that aim to identify neural correlates of reading improvement in general.

Attrition and data loss constituted another area of concern overall. Only 64% of primary studies reported drop-out rates (i.e., domains D6 and D7) between the treatment and control groups. Hence, results should be interpreted with caution because differential drop-out rates across treatment and control groups might have introduced attrition bias, such that there were systematic differences between participants who completed a study and those who dropped out. In the case of the studies in our review, data loss could occur due to failure to complete intervention and/or exclusion of data due to poor imaging quality. This raises concern as poor imaging quality may cooccur with underlying comorbidities such as attention deficits or more severe cases of RD, which may lead to final analysis groups that differ from the originally recruited groups. While some attrition is expected, researchers must take care to carefully report information on groups used for imaging to ensure that they are representative of the full sample.

Though randomized studies are needed to provide important data about brain changes caused by reading intervention, there are some cases in which comparison of groups determined *post-hoc* are valid and informative. Specifically, examination of effects related to responsiveness to intervention require classification of groups or characterization of individuals based on improvement in reading skills, or lack thereof. These comparisons provide insight to whether effects at the neurobiological level are associated with improvement in reading or persistent reading difficulties. Naturally, these comparisons are most informative in the context of randomized intervention studies.

In addition to the study quality appraisal, we examined factors related to imaging methods in the studies included in our meta-analysis that may influence the convergence of results across studies. Two factors of interest were the thresholds used to report fMRI results as significant and the brain coverage (field of view; FOV) of fMRI acquisition. With the exception of one study (Heim et al., 2015), the studies included in our meta-analysis reported effects at uncorrected thresholds which raises concern for a high rate of false positives reported in the primary studies (Eklund et al., 2016). Regarding FOV, a minimum inferior-superior FOV of 142 mm has been reported to achieve complete brain coverage in adults (Mennes et al., 2014). One study (Shaywitz et al., 2004) reported an inferior-superior FOV value less than 142 mm, and the authors noted that brain coverage ranged from the inferior aspect of the temporal lobes to the parietal convexity, indicating that superior and inferior brain regions (including cerebellum) were excluded from the imaging space. As a result, significant effects outside the bounds of the FOV may have been missed. In addition, three studies did not report sufficient information to determine the inferior-superior FOV (Gebauer et al., 2012; Temple et al., 2003; Yamada et al., 2011), leading to concern that fMRI brain coverage in these studies may have been incomplete. Brain regions most likely affected by this limited coverage include cerebellum, inferior temporal lobes, and/or superior parietal cortex. These limitations could contribute to the nature of results in our meta-analysis.

4.4. Limitations

Our quantitative meta-analysis was limited in scope due to the exclusion of studies for methodological reasons (i.e., ROI analysis, non-fMRI imaging modality), and small sample sizes within many of the included studies. Because only 8 studies met our inclusion criteria, we were unable to examine meta-analytic effects related to various factors that could contribute to the patterns of brain plasticity associated with intervention response, such as imaging tasks, intervention approaches, and individual differences among participants. A major limiting factor of our meta-analysis was the inability to include studies that used ROI analysis rather than whole brain analysis. Although ROI approaches are often well-justified, selecting ROIs based on the "typical reading network" biases investigation of intervention-related effects by limiting the analysis to regions expected to be activated by typical readers. Recruitment of alternative compensatory mechanisms can easily be missed in intervention studies that focus on changes in the left hemisphere reading network hubs. Exploratory studies with sufficient sample sizes to detect reliable effects across the whole brain are needed to fully capture brain changes associated with reading intervention.

A variety of reading-related fMRI tasks were used in the studies included in our meta-analysis and review, which may partly explain the heterogeneity in the patterns of brain activation observed (Murphy et al., 2019). As more fMRI studies of reading intervention become available, new, more selective, meta-analyses will be needed to identify task-specific effects. Similarly, our review and meta-analysis covered a variety of reading intervention approaches, including those focusing on phonics, fluency, attention, and mixed reading skills. The interventions also varied in duration and intensity. We were unable to quantitatively examine brain-level effects associated with different intervention styles. Several studies have examined brain changes related to different intervention approaches (Heim et al., 2015; Richards, et al., 2002; Richards et al., 2006a,b), but further research is needed to identify reliable effects.

Regarding individual differences among participants, factors such as age, initial skills, home literacy environment, and socioeconomic status may play a role. Behavioral interventions have shown greater efficacy in younger groups (e.g., Lovett et al., 2017), but it is difficult to dissociate intervention-specific effects from developmental effects in the brains of very young children, as bilateral activation of the language network is typical early in development (Olulade et al., 2020). Moreover, the focus of reading intervention may differ by age, with basic phonics training more prevalent in interventions targeting younger children, and more broad training spanning phonics, fluency, and comprehension more prevalent in interventions targeting older children and adolescents: a meta-analysis of behavioral intervention studies identified an interaction between grade at intervention and focus of intervention, with greater effects of phonics training in the younger grades and greater effects of comprehension training in the older grades (Suggate, 2010). Thus, age may be confounded with other intervention-related factors. In addition, other environmental factors may play a role in brain responses to reading intervention. For example, one study showed that lower socioeconomic status and greater severity of RD were associated with a higher degree of cortical thickening over the course of intervention (Romeo, et al., 2017). Empirical investigations of such factors on brain changes related to intervention remain scarce, and further research is needed.

The extant literature is also limited in that the majority of reported findings reflect group averages, but do not link changes in the brain to improvement in reading ability *per se*. That is, many studies report increases in group means in reading ability along with group-level changes in brain activation patterns, but these group averages do not explain individual differences among children in their responsiveness to intervention and associated changes in the brain. Thus, the effects presented in our review should be interpreted as intervention-related changes in brain activation, but not necessarily *improvement*-related changes. Indeed, numerous studies reviewed here have accounted for individual differences in intervention response and show distinct effects as a function of improvement in reading (Aboud et al., 2018; Davis et al., 2011; Farris et al., 2016; Heim et al., 2015; Horowitz-Kraus et al., 2019, 2014; Horowitz-Kraus et al., 2015; Nugiel et al., 2019a; Odegard et al., 2008; Simos et al., 2007; Rezaie et al., 2011, 2011; Romeo et al., 2017). These distinctions are crucial for identifying neural mechanisms of successful intervention, and future research *must* account for behavioral response to intervention. Additional empirical research is needed to identify relationships among specific intervention programs and individual differences in participant characteristics, and intervention response in brain and behavior. Further research is also needed to characterize the structural neural correlates of reading intervention, especially with regard to gray matter, which has only been addressed in a few studies to date.

Notably, several studies did not meet criteria for our review due to the focus on non-reading fMRI tasks or resting state functional connectivity (Gaab et al., 2007; Horowitz-Kraus et al., 2015; Olulade et al., 2013). Though outside the scope of this review, the studies by Gaab et al. and Olulade et al. provide unique insight to changes in processing complex auditory and visual stimuli following reading intervention. Horowitz-Kraus and colleagues (2015) reported intervention-related changes in resting state functional connectivity within the cingulo-opercular network, along with links to reading improvement. This study supports the utility of resting state fMRI approaches to study reading intervention-related changes in functional connectivity.

It remains unclear whether any specific region or pattern of activation is necessary and/or sufficient to evoke gains in reading ability in those with RD, and many studies revealed effects both within and outside of the primary reading network hubs. Moreover, several brain regions associated with the typical reading network are uniquely associated with other cognitive networks (i.e., attention, executive functioning), so distinguishing "normalized" vs. "compensatory" effects is not as simple as observing activation in specific regions. Differentiating pathways to successful reading as such may be less informative than characterizing brain networks associated with reading remediation, regardless of the "typical" reading-related functions the given regions. Rather, carefully identifying networks of activation and underlying structural changes associated with improvement in reading ability while considering various factors related to interventions and individual traits should drive research on the neural mechanisms of reading remediation. Further investigation can build upon this work to identify links to specific cognitive strategies and networks that support remediation.

Finally, publication bias and changes in conventional thresholds for reporting neuroimaging results may contribute to a lack of convergence in this literature. First, the publication of significant results, but not null findings, may lead to a false sense of substantial and significant brain changes associated with reading intervention. Furthermore, sample sizes and significance thresholds for reporting fMRI results have changed over the years in light of concern for inflated false positive rates in the fMRI literature (Eklund et al., 2016). Thus, the earlier studies included in our review and meta-analysis are likely to be susceptible to these issues, and findings must be interpreted with caution.

4.5. Conclusions

Major progress has been made in identifying the neural mechanisms associated with reading intervention and applying neuroimaging methods has added substantial value to reading intervention research, but the findings do not yet converge upon a set of effects consistently associated with improvement in reading ability. Substantial evidence of increased left STG, IPL and OT activation in groups that improved with reading intervention indicates that engagement of the typical reading network is one mechanism underlying successful remediation (Davis et al., 2011; Horowitz-Kraus et al., 2014; Simos et al., 2002; Simos et al., 2007a,b). On the other hand, several studies have linked greater activation/connectivity in right hemisphere regions as well as frontal and subcortical structures associated with domain-general cognitive processes to improvement in reading, lending credence to compensatory hypotheses (Horowitz-Kraus, et al., 2015; Horowitz-Kraus, et al., 2019; Nugiel et al., 2019; Odegard et al., 2008; Partanen et al., 2018). Together, these findings show that alternate routes to fluent reading can be successfully engaged when the typical left-lateralized phonological processing networks are not reliable. It is important to consider that mechanisms of change cannot be inferred based on location of change in the brain alone, so it remains unclear whether the observed changes are driven by cognitive processes engaged by typical readers or by alternative strategies. Rigorous research designs are needed to probe these subtle processes and relate them to changes in the brain and reading. Moreover, meta-analyses point to the involvement of right hemisphere and sub-cortical structures in typical readers (Martin et al., 2015; Murphy et al., 2019; Taylor et al., 2013), so engaging regions outside the left hemisphere cortical hubs of the reading network does not necessarily indicate compensation. Thus, we suggest moving away from describing intervention-related brain changes as “normalized” or “compensatory”, and instead aiming to characterize the complex interaction of cognitive systems that support improvement in reading. One important aspect in adjusting the view on brain systems supporting reading is to use exploratory whole brain analyses that may reveal regions outside the cortical “reading network” in order to identify effects in right hemisphere and sub-cortical regions, as limiting analyses to left hemisphere “reading network” ROIs may bias results toward a view of normalization of the reading network.

Altogether, the brain changes associated with improvement in reading remain unclear, and further empirical research on a larger scale is needed. Reading intervention studies with neuroimaging must be conducted with the rigor of randomized controlled trials in order to produce findings that inform and understanding of the *causal* mechanisms of reading intervention. Careful examination of brain changes associated with response to intervention is needed to distinguish between brain characteristics associated with improvement in reading versus persistent reading difficulties. Replication and more consistent use of fMRI tasks will provide important insight to task-specific effects, and later synthesis of studies will enable identification of task independent effects that may represent reliable changes in processing text. Similarly, further empirical research is needed to investigate effects related to intervention approach, duration and intensity. Systematic research in which fMRI task is held constant and intervention methods are manipulated, and vice versa, will be key to untangling these complex interacting factors.

Ultimately, this line of work will contribute to the understanding of neural mechanisms underlying remediation of reading difficulties. Future research targeting individual differences (risk profiles, age, socioeconomic status, etc.) may aid efforts to more specifically match at-risk children to intervention programs from which they are most likely to benefit. Studies of intervention-related factors may also support the development of targeted interventions. Further, this literature will inform directions for interventions applied at the brain level (e.g., transcranial magnetic stimulation, transcranial direct current stimulation) to support the development of efficient neural pathways.

Funding Acknowledgements

This work was supported by the following grants: NIH 2P50HD052120-11, NSF IGERT DGE-1144399, NSF GRFP DGE 1747453.

References

References marked with an asterisk indicate studies included in the synthesis review; those marked with a double-asterisk were additionally included in the quantitative meta-analysis.

*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

Al Otaiba, S., Schatschneider, C., & Silverman, E. (2005). Tutor-assisted intensive learning strategies in kindergarten: How much is enough?. *Exceptionality*, *13*(4), 195-208.

Albajes-Eizagirre, A., Solanes, A., Fullana, M. A., Ioannidis, J. P. A., Fusar-Poli, P., Torrent, C., ... Radua, J. (2019). Meta-analysis of Voxel-Based Neuroimaging Studies using Seed-based d Mapping with Permutation of Subject Images (SDM-PSI). *Journal of Visualized Experiments: JoVE*, (153), 1–7. https://doi.org/10.3791/59841

Araújo, S. (2015). The relationship between rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology*, *107*(3), 868–883. https://doi.org/10.1037/edu0000006

Arrington, C. N., Kulesz, P. A., Francis, D. J., & Fletcher, J. M. (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, E. H. H., Richards, T. L. L., Berninger, V. W. W., Nagy, W. E. E., Field, K. M. M., Grimme, A. . C., ... Cramer, S. C. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology*, *61*(2), 212–219. Retrieved from https://ezproxy.lib.uconn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true &db=cmedm&AN=12874401&site=ehost-live

*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. https://doi.org/10.1016/j.neuroimage.2013.05.062

Bailet, L. L., Repper, K. K., Piasta, S. B., & Murphy, S. P. (2009). Emergent literacy intervention for prekindergarteners at risk for reading failure. *Journal of Learning Disabilities*, *42*(4), 336-355.

Bailey, S. K., Aboud, K. S., Nguyen, T. Q., & Cutting, L. E. (2018). Applying a network framework to the neurobiology of reading and dyslexia. *Journal of Neurodevelopmental Disorders, 10*(1), 1–9. https://doi.org/10.1186/s11689-018-9251-z

Barquero, Laura A., Davis, N., & Cutting, L. E. (2014). Neuroimaging of reading intervention: A systematic review and activation likelihood estimate meta-analysis. *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0083668

*Barquero, Laura Alley. (2015). *Predicting Responsiveness to Reading Intervention with fMRI*. *ProQuest Dissertations and Theses*. Vanderbilt University. Retrieved from https://search.proquest.com/docview/2008722451?accountid=168248%0Ahttp://www.yi. edu.cn/educhina/educhina.do?artifact=&svalue=Predicting+Responsiveness+to+Readin+Interven tion+with+fMRI&stype=2&s=on%0Ahttp://pqdt.calis.edu.cn/Detail.aspx?pid=1075 3426%0Ah

Blair, C., & Razza, R. P. (2007). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Development*, 78(2), 647–663. http://eds.b.ebscohost.com.proxyiup.klnpa.org/ehost/pdfviewer/pdfviewer?vid=1&sid=be96 dddc-40b5-4e9a-916c4065f92670dc%40sessionmgr102

Boltzmann, M., Mohammadi, B., Samii, A., Münte, T. F., & Rüsseler, J. (2017). Structural changes in functionally illiterate adults after intensive training. Ne*uroscience,* 34*4*, 229– 242. H

*Brem, S., Bach, S., Kucian, K., Kujala, J. V, Guttorm, T. K., Martin, E., ... Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939–7944. https://doi.org/10.1073/pnas.0904402107

Chan, M. M., Yau, S. S., & Han, Y. M. (2021). The neurobiology of prefrontal transcranial direct current stimulation (tDCS) in promoting brain plasticity: A systematic review and meta-analyses of human and rodent studies. Neuroscience & Biobehavioral Reviews. https://doi.org/10.1016/j.neubiorev.2021.02.035

Chen, L., Wassermann, D., Abrams, D. A., Kochalka, J., Gallardo-diez, G., & Menon, V. (2019). The visual word form area (VWFA) is part of both language and attention circuitry. *Nature Communications*, *10*(5601), 1–12. https://doi.org/10.1038/s41467-019-13634-z

Chyl, K., Kossowski, B., Dębska, A., Łuniewska, M., Banaszkiewicz, A., Żelechowska, A., ... & Pugh, K. R. (2018). Prereader to beginning reader: changes induced by reading acquisition in print and speech brain networks. Jo*urnal of Child Psychology and Psychiatry*, 59*(1*), 76-87.

Chyl, K., Kossowski, B., Dębska, A., Łuniewska, M., Marchewka, A., Pugh, K. R., & Jednoróg, K. (2019). Reading acquisition in children: Developmental processes and Dyslexia-specific effects. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(10), 948–960. https://doi.org/10.1016/j.jaac.2018.11.007

Cutting, L. E., Clements-Stephens, A., Pugh, K. R., Burns, S., Cao, A., Pekar, J. J., ... Rimrodt, S. L. (2013). Not all reading disabilities are dyslexia: Distinct neurobiology of specific comprehension deficits. *Brain Connectivity*, *3*(2), 199–211. https://doi.org/10.1089/brain.2012.0116

D'Mello, A. M., & Gabrieli, J. D. E. (2018). Cognitive Neuroscience of Dyslexia. *Language Speech and Hearing Services in Schools*, *49*(4), 798. https://doi.org/10.1044/2018_LSHSS- DYSLC-18-0020

*Davis, N., Fan, Q., Compton, D. L., Fuchs, D., Fuchs, L. S., Cutting, L. E., ... Anderson, A. W. (2010). Influences of neural pathway integrity on children's response to reading instruction. *Frontiers in Systems Neuroscience*, *4*(150), 1–11. https://doi.org/10.3389/fnsys.2010.00150

*Davis, Nicole, Barquero, L., Compton, D. L., Fuchs, L. S., Fuchs, D., Gore, J. C., & Anderson, A. W. (2011). Functional correlates of children's responsiveness to intervention. *Developmental Neuropsychology*, *36*(3), 288–301. https://doi.org/10.1080/87565641.2010.549875

Denton, C. A. (2012). Response to intervention for reading difficulties in the primary grades: Some answers and lingering questions. *Journal of learning disabilities*, *45*(3), 232-243.

**Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., ... Cappell, K. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron*, *44*(3), 411–422. https://doi.org/10.1016/j.neuron.2004.10.019

Ehri, L., Nunes, S., Willows, D. M., Valeska Schuster, B., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250–287. https://doi.org/10.1598/RRQ.36.3.2

Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate*. Cambridge University Press. Erbeli, F., Hart, S. A., & Taylor, J. (2018). Examining the etiology of reading disability as conceptualized by the hybrid model. *Scientific Studies of Reading*, *22*(2), 167–180. https://doi.org/10.1080/10888438.2017.1407321

Erbeli, F., Hart, S. A., Wagner, R. K. & Taylor, J. (2018). Examining the etiology of reading disability as conceptualized by the hybrid model. *Scientific Studies of Reading*, *22*(2), 167–180. https://doi.org/10.1080/10888438.2017.1407321

*Farris, E. A., Ring, J., Black, J., Lyon, G. R., & Odegard, T. N. (2016). Predicting growth in word level reading skills in children with developmental dyslexia using an object rhyming functional neuroimaging task. *Developmental Neuropsychology*, *41*(3), 145–161. https://doi.org/10.1080/87565641.2016.1158264

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., ... Constable, R. T. (2014). Disruption of functional networks in dyslexia: A whole-brain, data-driven analysis of connectivity. *Biological Psychiatry*, *76*(5), 397–404. https://doi.org/10.1016/j.biopsych.2013.08.031

Frost, S. J., Landi, N., Mencl, W. E., Sandak, R., Fulbright, R. K., Tejada, E. T., ... & Pugh, K. R. (2009). Phonological awareness predicts activation patterns for print and speech. *Annals of dyslexia*, *59*(1), 78-97.

Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An fMRI study. *Restorative Neurology & Neuroscience*, *25*(3/4), 295–310. Retrieved from http://proxy.library.tamu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true &db=asn&AN=27060498&site=ehost-live

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. PloS one, 9(2), e89900.

**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). https://doi.org/10.1371/journal.pone.0038201

Glezer, L. S., Weisberg, J., O'Grady Farnady, C., McCullough, S., Midgley, K. J., Holcomb, P. J., & Emmorey, K. (2018). Orthographic and phonological selectivity across the reading system in deaf skilled readers. *Neuropsychologia*, *117*, 500–512. https://doi.org/10.1016/j.neuropsychologia.2018.07.010

Grigorenko, E. L. (2004). Genetic bases of developmental dyslexia: A capsule review of heritability estimates. *Enfance*, *56*(3), 273–288. https://doi.org/10.3917/enf.563.0273

Haft, S. L., Myers, C. A., & Hoeft, F. (2016). Socio-emotional and cognitive resilience in children with reading disabilities. *Current Opinion in Behavioral Sciences*, *10*, 133–141.https://doi.org/10.1016/j.cobeha.2016.06.005

Hancock, R., Richlan, F., & Hoeft, F. (2017). Possible roles for fronto-striatal circuits in reading disorder. *Neuroscience & Biobehavioral Reviews*, *72*, 243–260. https://doi.org/10.1016/J.NEUBIOREV.2016.10.025

Hart, H., Radua, J., Nakao, T., Mataix-Cols, D., & Rubia, K. (2013). Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention- deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. *JAMA psychiatry*, *70*(2), 185-198.

Hedenius, M., Ullman, M. T., Alm, P., Jennische, M., & Persson, J. (2013). Enhanced recognition memory after incidental encoding in children with developmental dyslexia. *PLoS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0063998

**Heim, S., Pape-Neumann, J., van Ermingen-Marbach, M., Brinkhaus, M., & Grande, M. (2015). Shared vs. specific brain activation changes in dyslexia after training of phonology, attention, or reading. *Brain Structure & Function*, *220*(4), 2191–2207. https://doi.org/10.1007/s00429-014-0784-y

Hickok, G., Poeppel, D. (2007) The cortical organization of speech processing. *Nat Rev Neuro*science, 8, 393–402. https://doi.org/10.1038/nrn2113

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., ... Gabrieli, J. D. E. (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences*, *108*(1), 361–366. https://doi.org/10.1073/pnas.1008950108

Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., ... Gabrieli, J. D. E. (2007). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience*, *121*(3), 602–613. https://doi.org/10.1037/0735-7044.121.3.602

Horowitz-Kraus, T., Grainger, M., Difrancesco, M., Vannest, J., Holland, S. K., & Consortium, C. A. (2015). Right is not always wrong: DTI and fMRI evidence for the reliance of reading comprehension on language-comprehension networks in the right hemisphere. *Brain Imaging and Behavior*, *9*, 19–31. https://doi.org/10.1007/s11682-014-9341-9

*Horowitz-Kraus, T., Hershey, A., Kay, B., & DiFrancesco, M. (2019). Differential effect of reading training on functional connectivity in children with reading difficulties with and without ADHD comorbidity. *Journal of Neurolinguistics*, *49*, 93–108. Retrieved from https://ezproxy.lib.uconn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cmedm&AN=31530970&site=ehost-live

Horowitz-Kraus, T., Toro-Serey, C., & DiFrancesco, M. (2015). Increased Resting-State Functional Connectivity in the Cingulo-Opercular Cognitive-Control Network after Intervention in Children with Reading Difficulties. *PloS One*, *10*(7), e0133762. https://doi.org/10.1371/journal.pone.0133762

*Horowitz-Kraus, T., Vannest, J. J., Kadis, D., Cicchino, N., Wang, Y. Y., & Holland, S. K. (2014). Reading acceleration training changes brain circuitry in children with reading difficulties. *Brain and Behavior*, *4*(6), 886–902. https://doi.org/10.1002/brb3.281

*Huber, E., Donnelly, P. M., Rokem, A., & Yeatman, J. D. (2018). Rapid and widespread white matter plasticity during an intensive reading intervention. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-04627-5

Katzir, T., Goldberg, A., Aryeh, T. J. B., Donnelley, K., & Wolf, M. (2013). Intensity vs. duration: Comparing the effects of a fluency-based reading intervention program, in after- school vs. summer school settings. *Journal of Education and Training Studies*, *1*(2), 61-73.

*Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., ... Brem, S. (2018). Simulating reading acquisition: The link between reading outcome and multimodal brain signatures of letter–speech sound learning in prereaders. S*cientific Reports*, 8. ttps://doi.org/10.1038/s41598-018-24909-8

*Keller, T. A., & Just, M. A. (2009). Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. *Neuron*, *64*, 624–631. https://doi.org/10.1016/j.neuron.2009.10.018

*Koen, B. J., Hawkins, J., Zhu, X., Jansen, B., Fan, W., & Johnson, S. (2018). The Location and Effects of Visual Hemisphere-Specific Stimulation on Reading Fluency in Children With the Characteristics of Dyslexia. *Journal of Learning Disabilities*, *51*(4), 399–415. https://doi.org/10.1177/0022219417711223

Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., ... Milham, M. P. (2013). Cortical signatures of dyslexia and remediation: an intrinsic functional connectivity approach. *PloS One*, *8*(2), e55454. Retrieved from https://ezproxy.lib.uconn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true &db=cmedm&AN=23408984&site=ehost-live

*Krafnick, A. J., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *NeuroImage*, *57*, 733–741. https://doi.org/10.1016/j.neuroimage.2010.10.062

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software, 28*(5), 1 - 26. doi:http://dx.doi.org/10.18637/jss.v028.i05/

Langer, N., Benjamin, C., Becker, B. L. C., & Gaab, N. (2019). Comorbidity of reading disabilities and ADHD: Structural and functional brain characteristics. *Human Brain Mapping*, *40(9)*, 2677–2698. https://doi.org/10.1002/hbm.24552

Lovett, M. W., Frijters, J. C., Wolf, M., Steinbach, K. A., Sevcik, R. A., Morris, R. D., ... Benson, N. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*, *109*(7), 889–914.

Lovett, M. W., Ransby, M. J., Hardwick, N., Johns, M. S., & Donaldson, S. A. (1989). Can dyslexia be treated? Treatment-specific and generalized treatment effects in dyslexic children's response to remediation. *Brain and Language*, *37*(1), 90–121. https://doi.org/10.1016/0093-934X(89)90103-X.

Lv, Q., Xu, G., Pan, Y., Liu, T., Liu, X., Miao, L., ... & Zou, Y. (2021). Effect of Acupuncture on Neuroplasticity of Stroke Patients with Motor Dysfunction: A Meta-Analysis of fMRI Studies. Neural Plasticity, 2021. https://doi.org/10.1155/2021/8841720

Lyon, G. Reid, Shaywitz Bennett A Sl, S. E., Brady, S., Catts, H., Dickman, E., Eden, G., ... Viall, T. (1995). PART I Defining Dyslexia, Comorbidity, Teachers' Knowledge of Language and Reading A Definition of Dyslexia. *Annals of Dyslexia in Annals of Dyslexia, 53*.

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E. & Richardson, U. (2009) In search of a science-based application: A learning tool for reading acquisition. *Scand. J. Psychol. 50*, 668–675.

Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A meta- analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences*, *1145*(1), 237–259. https://doi.org/10.1196/annals.1416.024

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. *Human Brain Mapping*, *36*(5), 1963–1981. https://doi.org/10.1002/hbm.22749

Mascheretti, S., Gori, S., Trezzi, V., Ruffino, M., Facoetti, A., & Marino, C. (2018). Visual motion and rapid auditory processing are solid endophenotypes of developmental dyslexia. *Genes, Brain and Behavior*, *17*(1), 70–81. https://doi.org/10.1111/gbb.12409

McGuinness, LA, Higgins, JPT (2020). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Syn Meth*. 1-7. https://doi.org/10.1002/jrsm.1411

Menghini, D., Carlesimo, G. A., Marotta, L., Finzi, A., & Vicari, S. (2010). Developmental dyslexia and explicit long-term memory. *Dyslexia*, *16*(3), 213-225. doi: 10.1002/dys.410

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. P*sychological Bulletin,* 1*38(*2), 322–352. ttps://doi.org/10.1037/a0026744

Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D. E., & Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia*, *46*(10), 2580–2592. https://doi.org/10.1016/j.neuropsychologia.2008.03.012

Miciak, J., Roberts, G., Taylor, W. P., Solis, M., Ahmed, Y., Vaughn, S., & Fletcher, J. M. (2018). The effects of one versus two years of intensive reading intervention implemented with late elementary struggling readers. *Learning Disabilities Research & Practice*, *33*(1), 24-36.

Mohammadi, B., Münte, T. F., Cole, D. M., Sami, A., Boltzmann, M., & Rüsseler, J. (2020). Changed functional connectivity at rest in functional illiterates after extensive literacy training. Ne*urological Research and Practice*, 2(*1*). Htps://doi.org/10.1186/s42466-020- 00058-0

Murphy, K. A., Jogia, J., & Talcott, J. B. (2019). On the neural basis of word reading: A meta-analysis of fMRI evidence using activation likelihood estimation. *Journal of Neurolinguistics*, *49*, 71-83.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (National Institute of Health Pub. No. 00-4769). Washington, DC: National Institute of Child Health and Human Development.

Newport, E. L., Landau, B., Seydell-Greenwald, A., Turkeltaub, P. E., Chambers, C. E., Dromerick, A. W., … Gaillard, W. D. (2017). Revisiting Lenneberg's hypotheses about early developmental plasticity: Language organization after left-hemisphere perinatal stroke. *Biolinguistics, 11*, 407–422.

Nicolson, R. I., & Fawcett, A. J. (2019). Development of dyslexia: The delayed neural commitment framework. *Frontiers in Behavioral Neuroscience, 13*, 1–16. https://doi.org/10.3389/fnbeh.2019.00112

Nicolson, R. I., Fawcett, A. J., Brookes, R. L., & Needle, J. (2010). Procedural learning and dyslexia. *Dyslexia*, *16*(3), 194-212. http://dx.doi.org/10.1002/dys.408

Nugiel, T., Roe, M. A., Taylor, W. P., Cirino, P. T., Vaughn, S. R., Fletcher, J. M., ... Church, J. A. (2019). Brain activity in struggling readers before intervention relates to future reading gains. *Cortex; a*

Journal Devoted to the Study of the Nervous System and Behavior, *111*, 286–302. Retrieved from https://ezproxy.lib.uconn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true &db=cmedm&AN=30557815&site=ehost-live

*Odegard, T. N., Ring, J., Smith, S., Biggan, J., & Black, J. (2008). Differentiating the neural response to intervention in children with developmental dyslexia. *Annals of Dyslexia*, *58*(1), 1–14. https://doi.org/10.1007/s11881-008-0014-5

Olulade, O. A., Napoliello, E. M., & Eden, G. F. (2013). Abnormal visual motion processing is not a cause of dyslexia. *Neuron*, *79(1)*, 180–190. https://doi.org/10.1016/j.neuron.2013.05.002

Olulade, O. A., Seydell-greenwald, A., Chambers, C. E., & Turkeltaub, P. E. (2020). The neural basis of language development : Changes in lateralization over age. Proceedings of the National Academy of Sciences, *117*(38). https://doi.org/10.1073/pnas.1905590117

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D. E., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: implications for early screening and theories of reading. *Developmental Science*, *20*(5), 1–18. https://doi.org/10.1111/desc.12471

Page, M. J., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71 doi: 10.1136/bmj.n71

**Partanen, M., Siegel, L. S., & Giaschi, D. E. (2019). Effect of reading intervention and task difficulty on orthographic and phonological reading systems in the brain. *Neuropsychologia*, *130*, 13–25. Retrieved from http://10.0.3.248/j.neuropsychologia.2018.07.018

Paulesu, E., Danelli, L., & Berlingeri, M. (2014). Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Frontiers in Human Neuroscience*, *8*, 1–20. https://doi.org/10.3389/fnhum.2014.00830

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*, 385–413. https://doi.org/10.1016/j.cognition.2006.04.008

Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Reviews Psychology*, *60*, 283–306. https://doi.org/10.1146/annurev.psych.60.110707.163548

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet*, *379*(9830), 1997–2007. https://doi.org/10.1016/S0140-6736(12)60198-6

Peterson, R. L., & Pennington, B. F. (2015). Developmental Dyslexia. *Annual Review of Clinical Psychology*, *11*(1), 283–307. https://doi.org/10.1146/annurev-clinpsy-032814-112842

Pugh, K. R., Mencl, W. E. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., ... Shaywitz, B. A. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders*, *34*(6), 479–492. https://doi.org/10.1016/S0021-9924(01)00060- 0

Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., ... Gore, J. C. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *Psychological Science*, *11*(1), 51–56. https://doi.org/10.1111/1467-9280.00214

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2*(3), 1–18. https://doi.org/10.1177/2332858416657343

*Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., & Papanicolaou, A. C. (2011a). Engagement of temporal lobe regions predicts response to educational interventions in adolescent struggling readers. *Developmental Neuropsychology*, *36*(7), 869–888. Retrieved from https://ezproxy.lib.uconn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true &db=cmedm&AN=21978010&site=ehost-live

*Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., & Papanicolaou, A. C. (2011b). 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. https://doi.org/10.1017/S1355617711000890

**Richards, T. L., Aylward, E. H., Berninger, V. W., Field, K. M., Grimme, A. C., Richards, A. L., & Nagy, W. (2006). Individual fMRI activation in orthographic mapping and morpheme mapping after orthographic or morphological spelling treatment in child dyslexics. *Journal of Neurolinguistics*, *19*(1), 56–86. https://doi.org/10.1016/j.jneuroling.2005.07.003

*Richards, T. L., Aylward, E. H., Field, K. M., Grimme, A. C., Raskind, W., Richards, A. L., ... Berninger, V. W. (2006). Converging Evidence for Triple Word Form Theory in Children With Dyslexia. *Developmental Neuropsychology*, *30*(1), 547–589. Retrieved from toddr@u.washington.edu

*Richards, T., Berninger, V., Winn, W., Stock, P., Wagner, R., Muse, A., & Maravilla, K. (2007). Functional MRI activation in children with and without dyslexia during pseudoword aural repeat and visual decode: Before and after treatment. *Neuropsychology*, *21*(6), 732.

*Richards, T. L., & Berninger, V. W. (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. *Journal of Neurolinguistics*, *21*(4), 294–304. https://doi.org/10.1016/j.jneuroling.2007.07.002

Richards, T. L., Corina, D., Serafini, S., Steury, K., Echelard, D. R., Dager, S. R., ... Berninger, V. W. (2000). Effects of a phonologically driven treatment for dyslexia on lactate levels measured by proton MR spectroscopic imaging. *AJNR. American Journal of Neuroradiology*, *21*(5), 916–922.

Richards, T., Nagy, W., Abbott, R., & Berninger, V. (2016). Brain connectivity associated with cascading levels of language. *Journal of Systems and Integrative Neuroscience*, *2*(3). https://doi.org/10.15761/JSIN.1000139

*Richards, T., Peverly, S., Wolf, A., Abbott, R., Tanimoto, S., Thompson, R., ... Berninger, V. (2016). Idea units in notes and summaries for read texts by keyboard and pencil in middle childhood students with specific learning disabilities: Cognitive and brain findings. *Trends in Neuroscience and Education*, *5*(3), 146–155. https://doi.org/10.1016/j.tine.2016.07.005

*Richards, T.L., Berninger, V. W., Yagle, K. J., Abbott, R. D., & Peterson, D. J. (2017). Changes in DTI diffusivity and fMRI connectivity cluster coefficients for students with and without specific learning disabilities in written language: Brain's response to writing instruction. *Journal of Nature and Science*, *3*(4), 1–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/28670621%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5488805

Richards, Todd L, Berninger, V. W., Aylward, E. H., Richards, A. L., Thomson, J. B., Nagy, W. E., ... Abbott, R. D. (2002). Reproducibility of proton MR spectroscopic imaging (PEPSI): comparison of dyslexic and normal-reading children and effects of treatment on brain lactate levels during language tasks. *AJNR. American Journal of Neuroradiology*, *23*(10), 1678–1685. Retrieved from https://ezproxy.lib.uconn.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cmedm&AN=12427623&site=ehost-live

*Richards, Todd L, Berninger, V. W., Yagle, K., Abbott, R. D., & Peterson, D. (2018). Brain's functional network clustering coefficient changes in response to instruction (RTI) in students with and without reading disabilities: Multi-leveled reading brain's RTI. *Cogent Psychology*, *5*. https://doi.org/10.1080/23311908.2018.1424680

Richardson, U., Thomson, J. M., Scott, S. K., & Goswami, U. (2004). Auditory processing skills and phonological representation in dyslexic children. *Dyslexia*, *10*(3), 215–233. https://doi.org/10.1002/dys.276

Richlan, F. (2012). Developmental dyslexia: Dysfunction of a left hemisphere reading network. *Frontiers in Human Neuroscience*, *6*(May), 1–5. https://doi.org/10.3389/fnhum.2012.00120

Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping*, *30*(10), 3299–3308. https://doi.org/10.1002/hbm.20752

*Romeo, R. R., Christodoulou, J. A., Halverson, K. K., Murtagh, J., Cyr, A. B., Schimmel, C., ... Gabrieli, J. D. E. (2017). Socioeconomic status and reading disability: Neuroanatomy and plasticity in response to intervention. *Cerebral Cortex*, *28*(7), 2297–2312. https://doi.org/10.1093/cercor/bhx131

Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K. L., Maisog, J. M., & Andreason, P. (1999). A functional lesion in developmental dyslexia: Left angular gyral blood flow predicts severity. *Brain and Language*, *70*(2), 187–204. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L29548843%5Cnhttp://dx.doi.org/10.1006/brln.1999.2158%5Cnhttp://rug.on.worldcat.org/atoztitles/link/?sid=EMBASE&issn=0093934X&id=doi:10.1006%2Fbrln.1999.2158&atitle=A+functional+lesi

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of learning disabilities*, *48*(4), 369-390.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, *96*(2), 265–282. https://doi.org/10.1037/0022-0663.96.2.265

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology, 15*(3), 232-246. doi: 10.1080/09297040802220029

**Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., ... Gore, J. C. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, *55*, 926–933. https://doi.org/10.1016/j.biopsych.2004.01.006

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E. E., Fulbright, R. K., Skudlarski, P., ... Gore, J. C. (2002). Disruption of posterior brain systems for reading in children with Developmental Dyslexia. *Biological Psychiatry*, *52*(2), 101–110. https://doi.org/10.1016/S0006-3223(02)01365-3

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., ... Gore, J. C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences*, *95*(5), 2636–2641. https://doi.org/10.1073/pnas.95.5.2636

Simos, P. G., Breier, J. I., Wheless, J. W., Maggio, W. W., Fletcher, J. M., Castillo, E. M., & Papanicolaou, A. C. (2000). Brain mechanisms for reading: the role of the superior temporal gyrus in word and pseudoword naming. *Neuroreport*, *11*(11), 2443-2446.

*Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., ... Papanicolaou, A. C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology*, *58*(8), 1203–1213. https://doi.org/10.1212/wnl.58.8.1203

*Simos, P. G., Fletcher, J. M., Sarkari, S., Billingsley-Marshall, R., Denton, C. A., & Papanicolaou, A. C. (2007). Intensive instruction affects brain magnetic activity associated with oral word reading in children with persistent reading disabilities. *J Learn Disabil*, *40*(1), 37–48. https://doi.org/10.1177/00222194070400010301

*Simos, P. G., Fletcher, J. M., Sarkari, S., Billingsley, R. L., Denton, C., & Papanicolaou, A. C. (2007). Altering the brain circuits for reading through intervention: A magnetic source imaging study. *Neuropsychology*, *21*(4), 485–496. https://doi.org/10.1037/0894- 4105.21.4.485

*Simos, Panagiotis G, Fletcher, J. M., Sarkari, S., Billingsley, R. L., Francis, D. J., Castillo, E. M., ... Papanicolaou, A. C. (2005). Early development of neurophysiological processes involved in normal reading and reading disability: a magnetic source imaging study. *Neuropsychology*, *19*(6), 787–798. https://doi.org/10.1037/0894-4105.19.6.787

Snowling, M. J., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology*, *81*(1), 1– 23. https://doi.org/10.1111/j.2044-8279.2010.02014.x

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia*, *7*(1), 12–36. https://doi.org/10.1002/dys.186

Study Quality Assessment Tools. (2021, July). Retrieved from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental psychology*, *46*(6), 1556-1579.

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of learning disabilities*, *49*(1), 77-96.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182–198. https://doi.org/10.1016/0093-934X(80)90139-X

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, *139*(4), 766–791. https://doi.org/10.1037/a0030266

**Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. E. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences*, *100*(5), 2860–2865. https://doi.org/10.1073/pnas.0030098100

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, *15*(1), 55–64.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology*, *91*(4), 579–593.

Torgesen, Joseph K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, *34*(1).

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nature Neuroscience*, *6*(6), 767–773. https://doi.org/10.1038/nn1065

Ullman, M. T., Earle, F. S., Walenski, M., & Janacsek, K. (2020). The neurocognition of developmental disorders of language. *Annual Review of Psychology*, *71*(1), 1–29. https://doi.org/10.1146/annurev-psych-122216-011555

Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews*, *51*, 205–222. https://doi.org/10.1016/j.neubiorev.2015.01.008

Vaughn, S., & Wanzek, J. (2014). Intensive interventions in reading for students with reading disabilities: Meaningful impacts. *Learning Disabilities Research & Practice*, *29*(2), 46-53. Wanzek, J., &

Vaughn, S. (2007). based implications from extensive early reading interventions. *School Psychology Review*, *36*(4), 541-561.

Verwimp, C., Bempt, F. Vanden, Kellens, S., & Economou, M. (2020). Pre-literacy heterogeneity in Dutch-speaking kindergartners: latent profile analysis. *Annals of Dyslexia*, *70*, 275–294.

Wang, Y., Mauer, M. V., Raney, T., Peysakhovich, B., Becker, B. L. C., Sliva, D. D., & Gaab, N. (2017). Development of tract-specific white matter pathways during early reading development in at-risk children and typical controls. *Cerebral Cortex*, *27*(4), 2469–2485. https://doi.org/10.1093/cercor/bhw095

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. *Journal of Learning Disabilities*, *51*(6), 612–624. https://doi.org/10.1177/0022219418775110

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta- Analyses of the Effects of Tier 2 Type Reading Interventions in Grades K-3. *Educational Psychology Review*, *28*(3), 551–576. https://doi.org/10.1007/s10648-015-9321-7

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). *Extensive Reading Interventions for Students With Reading Difficulties After Grade 3*. *Review of Educational Research* (Vol. 83). https://doi.org/10.3102/0034654313477212

Warmington, M., & Hulme, C. (2012). Phoneme awareness, visual-verbal paired-associate learning, and rapid automatized naming as predictors of individual differences in reading ability. *Scientific Studies of Reading*, *16*(1), 45–62. https://doi.org/10.1080/10888438.2010.534832

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The Development of Cognitive Skills and Gains in Academic School Readiness for Children From Low-Income Families. *Journal of Educational Psychology*, *102*(1), 43–53. https://doi.org/10.1037/a0016738

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*(3), 415–438. https://doi.org/10.1037/0022-0663.91.3.415

**Yamada, Y., Stevens, C., Dow, M., Harn, B. A., Chard, D. J., & Neville, H. J. (2011). Emergence of the neural network for reading in five-year-old beginning readers of different levels of pre-literacy abilities: an fMRI study. *NeuroImage*, *57*(3), 704–713. https://doi.org/10.1016/j.neuroimage.2010.10.057

Yeatman, J. D., & Huber, E. (2019). Sensitive periods for white matter plasticity and reading intervention. *bioRxiv*, 346759.

Yu, X., Zuk, J., Perdue, M. V., Ozernov-Palchik, O., Raney, T., Beach, S. D., ... Gaab, N. (2020). Putative protective neural mechanisms in pre-readers with a family history of dyslexia who subsequently develop typical reading skills. *Human Brain Mapping*, *41*(10), 1–19. https://doi.org/10.1002/hbm.24980

Yu, Xi, Zuk, J., & Gaab, N. (2018). What factors facilitate resilience in developmental dyslexia? Examining protective and compensatory mechanisms across the neurodevelopmental trajectory. *Child Development Perspectives*, *12*(4), 240–246. https://doi.org/10.1111/cdep.12293

Zijlstra, H., Van Bergen, E., Regtvoort, A., de Jong, P. F., & vam der Leij, A. (2020). Prevention of reading difficulties in children with and without familial risk: Short- and long-term effects of an early intervention. *Journal of Educational Psychology*, 1–20. https://doi.org/10.1037/edu0000489

Zuk, J., Dunstan, J., Norton, E., Yu, X., Ozernov-Palchik, O., Wang, Y., ... Gaab, N. (2020). Multifactorial pathways facilitate resilience among kindergarteners at risk for dyslexia: A longitudinal behavioral and neuroimaging study. *Developmental Science*, 1–18. https://doi.org/10.1111/desc.12983

Figure Captions

Figure 1.

PRISMA Flow Diagram for Systematic Review Literature Screening

Supplementary Materials

Literature Search Terms

PsychInfo, ERIC, Academic Search Ultimate, MedLine, and EBSCOhost eBook Collection:

(((DE "intervention” or TI reading interven* or AB reading interven*) OR (DE "remediation" or TI reading remedia* or AB reading remedia*) OR (DE "treatment" or TI reading treat* or AB reading treat*) OR (DE "reading instruction"))

AND

((DE "dyslexia" or TI dyslex* or AB dyslex*) OR (DE "reading disability" or TI reading disab* or AB reading disab*) OR (DE "reading difficulty" or TI reading difficult* or AB reading difficult*)) OR (DE "reading impairment" or TI reading impair* or AB reading impair*) OR (TI struggling reader* or AB struggling reader*) OR (TI poor reader* or AB poor reader*) OR (TI less-skilled read* or AB less-skilled read*))

AND

((TI neuroimag* or AB neuroimag*) OR (DE "neuro" or TI neur* or AB neur*) OR (DE "MRI") OR (DE "fMRI") OR (TI brain imag* or AB brain imag*) OR (TI brain activ* or AB brain activ*)))

PubMed:

((("reading"[MeSH Terms] OR "reading"[All Fields]) AND disability[All Fields]) OR ("reading"[MeSH Terms] OR "reading"[All Fields]) AND difficulty[All Fields]) OR ("dyslexia"[MeSH Terms] OR "dyslexia"[All Fields])

AND

(("neuroimaging"[MeSH Terms] OR "neuroimaging"[All Fields]) OR ("magnetic resonance imaging"[MeSH Terms] OR ("magnetic"[All Fields] AND "resonance"[All Fields] AND "imaging"[All Fields]) OR "magnetic resonance imaging"[All Fields] OR "fmri"[All Fields]) OR (("brain"[MeSH Terms] OR "brain"[All Fields]) AND activation[All Fields]))

AND

((("reading"[MeSH Terms] OR "reading"[All Fields]) AND ("Interv Sch Clin"[Journal] OR "intervention"[All Fields])) OR (("reading"[MeSH Terms] OR "reading"[All Fields]) AND ("teaching"[MeSH Terms] OR "teaching"[All Fields] OR "instruction"[All Fields])))

Google Scholar:

(~intervention OR ~remediation OR ~instruction OR ~treatment) AND (reading disability OR ~dyslexia OR reading impairment OR struggling reader OR poor reader) AND (neuroimaging OR MRI OR brain imaging OR brain activation)

Coding procedure:

Initial coding of data from full-text articles was completed by two coders. Initial coders read through each of the full-text articles, identified the relevant information, and entered the data into separate spreadsheets. Coding sheets from each coder were compared and discrepancies between coders were discussed and resolved by a third coder. Final coding was compiled in a separate spreadsheet which was used to determine study eligibility and extract information for the review and meta-analysis.

Supplementary Table 1: Coding variables for articles in systematic review

Variable	Definition and/or Coding Scheme
First Author	Record the name of first author
Full Author List	Record the names of all authors in the order they appear in the paper
Corresponding Author Name and Contact Information	Record the name and contact information, usually an email address, for the corresponding author
Publication Year	Record the year of publication
Journal	Record the name of the journal in which the article was published. If not a journal article, leave this field blank
Publication Type	Record if the publication is a journal article, dissertation, etc.
Sample Size (Full Sample)	Record the sample size including all participants across all analysis groups
Sample Size (Reading Disabled/Dyslexia)	Record the sample size for those in the reading disability, dyslexia, or other reading group of interest
Sample Size (Controls)	Record the sample size for control participants
Sample Size (Additional)	Record the sample size and notes about any additional samples such as a waiting control group with reading disability
Number of Males	Record number of males in the sample
Number of Females	Record number of females in the sample
Mean Age	Record mean age of the sample and standard deviation if available
Age Range	Record age range if available
Language of Testing (Native Language)	Record language that testing was conducted in and not if this was not the participants native language
Did the study control for age?	Record a 0 if the study did not control for age and a 1 if they did control for age
Did the study control for sex?	Record a 0 if the study did not control for sex and a 1 if they did control for sex
Was there a co-morbidity of interest?	Record a 1 if yes and a 2 if no.
Co-morbidity description	If the above is yes, explain the co-morbidity
Intervention Name	Record the name of the intervention program if available
Intervention Style	Record information on the teaching style used in the intervention program such as one-on-one teaching, computer-based intervention, small group intervention, etc.

Intervention Focus	Record information on the skill focus of the intervention such as phonics, grapheme-phoneme correspondence, mixed focus- phonics and orthography, etc.
Intervention Duration (Weeks)	Record information on the length of an intervention program in weeks
Intervention Session Frequency (Sessions/Week)	Record the number of intervention sessions per week for an intervention program. Also record information on the mean number of sessions attended by participants if provided
Intervention Session Length	Record the length in minutes of each session. Include information about the mean number of minutes of intervention each participant participated in if provided.
Socioeconomic Status as reported in the article	Record SES information as it is presented in the article.
Socioeconomic Status (Classification)	Based on the information reported above, classify if each article included primarily low-, middle-, high-, or mixed-income participants. *Note- because of the small number of studies that included SES information, this classification was not used in our final coding scheme
Socioeconomic Status (Parent Education)	Record information about parent education if it is reported in the article. Use a 1 to indicate completing high school or lower, a 2 to indicate completing college, and a 3 to indicate completing graduate school. Record if this is an average of both parents or a single parent's information. If a single parent, record if it is mother or father.
Imaging Modality	Record if the neuroimaging was functional MRI, structural MRI, MEG, DTI, etc.
Imaging Analysis Scheme	Indicate if the imaging analysis included a whole brain approach using a 1, or a regions of interest approach using a 2
Imaging Task as reported in article	Record the description of the imaging task as it is included in the article
Imaging Task Classification	Based on the reported information above, classify if each imaging task: 1. Phonological (including auditory stimulus only such as rhyming or sound matching) 2. Orthographic (letter matching) 3. Orthographic-phonological (letter-sound mapping) 4. Word/pseudoword reading (single words) 5. Sentence reading 6. other
Imaging Analysis Software	Record what software was used to analyze imaging data including information on specific packages/sequences used
Coordinate Space Reported	Record what space coordinates are reported in, such as MNI space, Talairach space, or other
Voxel Size	Record the voxel size
FWHM	Record the Full Width at Half Maximum-Smoothing Kernel in mm
Significance Threshold for Reporting (Uncorrected)	Record uncorrected thresholds for reporting significant results, typically as a p-value or z-score
Significance Threshold for Reporting (Corrected)	Record corrected thresholds for reporting significant results, typically as a p-value or z-score
Correction Type	Record correction type

Pre-intervention peak coordinates	Record all baseline pre-intervention peak coordinates from imaging data
Pre-intervention peak coordinates (Group differences and/or task-related peak activation)	Record all pre-intervention peak coordinates from specific groups or tasks that were significant. Also record information about the groups, tasks, etc.
Post-intervention peak coordinates	Record all baseline post-intervention peak coordinates from imaging data
Post-intervention peak coordinates (Group differences and/or task-related peak activation)	Record all post-intervention peak coordinates from specific groups or tasks that were significant. Also record information about the groups, tasks, etc.
Pre- to Post-intervention change in brain activation (by group)	Record all pre- to post-intervention brain activation changes for each group. Also record information on if these changes were seen on specific tasks, etc.
Peak coordinates of clusters showing a correlation with reading ability	Record coordinates of clusters showing a correlation with reading ability including information about in which groups these clusters were correlated with reading ability and which reading tests or subskills particular clusters were correlated with
T, Z, or F values of peak coordinates	Record the T, Z, or F values of peak coordinates and note which values were reported
P-values or Z-scores of peak coordinates	Record the p-values or z-scores reported for each peak coordinate
Is there additional data available vis OSF, preregistration, or supplemental materials?	Indicate if additional material is available by recording Yes or is not by recording No. If yes, indicate a link or way to access that information
Notes	Include any notes about the publication

Protocol for additional literature search of articles that have cited those included in our review:

Use Google Scholar to search for articles that have cited studies that met inclusion criteria for consideration

Search the study in Google Scholar, click "Cited by ###" link in google (make note of the number of articles that have cited this paper)

Select the "search within citing articles" option below the search bar

Enter search terms (based on our initial search):

(~intervention OR ~remediation OR ~instruction OR ~treatment) AND (reading disability OR ~dyslexia OR reading impairment OR struggling reader OR poor reader) AND (neuroimaging OR MRI OR brain imaging OR brain activation)

Make note of # of full-text studies considered, of these # excluded and why

Make note of # of studies included and study info

Add newly included studies to coding sheet

Supplementary Table 2: Questions and specific criteria used for quality analysis based on NIH recommendations for quality assessment of controlled intervention studies (https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools)

Domain	Criteria
1. Was the treatment group compared to a control group of individuals with RD who did not receive intervention?	Was there a waiting or control group of RD children used as a comparison for the intervention group?
2. Was the study described as randomized, a randomized trial, a randomized clinical trial, or an RCT?	Manuscript must explicitly include the word random. Those that simply imply randomization will be listed as No Report (NR).
3. Was the method of randomization adequate (i.e., use of randomly generated assignment)?	Was true or partial randomization used? Paying particular attention to how RD groups were randomized. Partial randomization is accepted if in RD groups.
4. Were the people assessing the outcomes blinded to the participants' group assignments?	Were those administering post-intervention behavioral or imaging or those analyzing data blind to participants group? Note in table if only one of the two aforementioned groups was blind.
5. Were the groups similar at baseline on important characteristics that could affect outcomes (e.g., demographics, risk factors, co-morbid conditions)?	Manuscript must include a report of baseline characteristics between groups. All baseline measures, with the exception of reading measures which are expected to be different at baseline, should be statistically similar.
6. Was the overall drop-out rate from the study at endpoint 20% or lower of the number allocated to treatment?	For the total number of participants, including both control participants and those in intervention, was the dropout rate less than 20%? Were all participants included in imaging both pre- and post-intervention? Was there imaging data loss due to movement that excluded more than 20% of participants at a given timepoint? Can be calculated by raters using reported sample sizes.
7. Was the differential drop-out rate (between treatment groups) at endpoint 15 percentage points or lower?	Was the drop-out rate of one group, or a few groups, particularly high? Can be calculated by raters using reported sample sizes if samples sizes are reported by group.

8. Was there high adherence to the intervention procedure for each treatment group?	Was the intervention program previously researched or commercially available and standard procedures were followed? Were those who administered the intervention highly trained?
9. Was there high adherence to the intervention duration for each treatment group?	Was the duration of intervention reported? Were participants who missed a high number of intervention sessions excluded? What was the mean amount of time spent in intervention? How much of the intervention program had participants completed upon conclusion of the project?
10. Were behavioral outcomes assessed using valid and reliable measures, implemented consistently across all study participants?	Were the measures used to assess behavior reliable? Were the same measures used across all groups including intervention and control groups?
11. Were imaging outcomes assessed using valid and reliable measures, implemented consistently across all study participants?	Were the measures used for neuroimaging reliable? Did all groups complete the same neuroimaging at the same timepoints including intervention and control groups? If MRI analysis used an ROI approach- How were ROIs selected? Were appropriate statistical thresholds used in imaging analysis? Raters instructed to leave detailed notes on imaging analysis including ROI selection and statistical thresholds used.
12. Did the authors report that the sample size was sufficiently large to be able to detect a difference in the main outcome between groups with at least 80% power?	Did authors explicitly report possible power given sample size?
13. Were all participants analyzed in the group to which they were originally assigned, i.e., did they use an intention-to-treat analysis?	Were any groups combined for analysis?