ACADEMIC AND NEUROIMAGING OUTCOMES OF SCHOOL-BASED READING INTERVENTIONS

by

MARITA HELMI PARTANEN

B.A., The University of Waterloo, 2005
M.Sc., The University of British Columbia, 2010
M.Ed., The University of British Columbia, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(School Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2017

ABSTRACT

Early intervention is important for decreasing the prevalence of reading disabilities. However, despite receiving treatment, some children continue to struggle with reading and therefore they require ongoing supports. Intensive and individualized programs may be beneficial for the lowest-performing readers, however, empirical review of intensive and individualized programs has not been widely conducted. Furthermore, there is a neurobiological basis to reading impairments. Children with poor reading skills have differences in brain function and structure when compared to typically-developing readers, and there may be changes in the brain after intervention. However, the combination of multiple reading tasks in functional brain imaging along with measures of grey and white matter structure has not been conducted previously. Therefore, the purpose of my dissertation was to evaluate the academic and neurobiological outcomes of an intensive reading program as well as to determine the predictors of reading success. In Chapter 2, poor readers receiving intensive instruction were compared to other poor readers receiving small group supports as well as to good readers not receiving additional supports. Performance on academic and cognitive measures were evaluated before and after 3 months of instruction and one year later. In Chapters 3 and 4, poor readers and good readers completed functional imaging tasks (Chapter 3) and scans of grey and white matter (Chapter 4) before and after 3 months of instruction. The results showed that students in the intensive program had improved word recognition and decoding fluency immediately after intervention and one year later. Changes after intervention were also shown in functional brain activity during a rhyming task, but not during a spelling task or in grey and white matter structures. However, baseline reading and spelling skills, brain activity in the left hemisphere, and white matter organization in the right hemisphere were associated with gains in reading skills over time. Although improvements in reading were shown, a significant gap between poor and good readers persisted in the third and fourth grades. Overall, this dissertation illustrates the importance of an intensive reading program and the need for continuing supports, and that both academic and neuroimaging measures are associated with reading outcome.

LAY ABSTRACT

When children have difficulty learning to read, tutoring programs help many of them to develop good reading skills. Some of these children, however, continue to have poor reading skills and therefore they need ongoing tutoring. Also, the underlying cause of a reading disability may be related to differences in genetics or brain development. The main purpose of this dissertation was to examine the outcomes of reading tutoring programs in school, which were measured by academic tests and brain imaging. The results showed that students who completed an intensive reading program had increased reading skills over time, and there were changes in brain function but not structure. Before the intervention started, it was also shown that literacy skills, brain function, and brain structure were related to improved reading over time. Overall, these studies show the importance of early reading programs and what factors predict improvement in reading skills.

PREFACE

The research presented in this dissertation was conducted at The University of British Columbia. Ethics approval was obtained from The University of British Columbia Children's and Women's Health Centre of BC Research Ethics Board [certificate # CW13-0113 / H13-01050]. I was involved in the design, participant recruitment, data collection, analysis, and original writing of the studies in this dissertation, under the supervisory direction of Dr. Deborah Giaschi and Dr. Linda Siegel. There were also research assistants from Dr. Giaschi's lab and staff at the Child and Family Research Imaging Facility who assisted with data collection and analysis. None of the text of the dissertation is taken directly from previously published articles. A version of Chapter 2 has been submitted for publication, along with my supervisors as co-authors. Chapters 3 and 4 are under preparation for publication.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would first and foremost like to thank my supervisors, Dr. Debbie Giaschi and Dr. Linda Siegel, for the support you have provided me throughout my graduate training. The work in this dissertation would not have been possible without your guidance in all areas of school, career, and personal life. I am glad to have had such supportive supervisors, and it is hard to believe that 10 years have passed since I started at the Giaschi Lab. I look forward to continuing to work and collaborate with you as I move into my career.

I would also like to acknowledge my committee members and reviewers, Dr. Sterett Mercer, Dr. Bruce Bjornson, Dr. Rachel Weber, and Dr. Laurie Ford. I am thankful to have learned from a diverse range of disciplines and training backgrounds, which started in the Brain Mapping Lab and continued through my coursework and clinical training in School Psychology. Your ongoing input and support have helped me to get where I am today.

The research in this dissertation would not have been possible without the involvement of the schools, teachers, students, and families from the North Vancouver School District. I would particularly like to acknowledge Susan McLean, Vanessa Barattini, Jeff Ballou, and the school psychologists in the district. You supported me throughout this process and I know that you are invested in the outcomes of all of your students. I am very thankful to have mentors and colleagues like you.

Last but not least, there are many lab mates, friends, and family members who have been with me since the beginning. Kim Meier, thank you for all of the statistical advice and for getting excited about stats, just like me. Also, to Kevin Fitzpatrick, Danny Kim, and Lynne Williams, I am glad to have colleagues who I can have an entire conversation with in acronym form, and I am sure this MRI language will only improve. To the friends I met in Neuroscience and School Psychology: you were always the cheerleaders and supportive force when we were trying to balance coursework, clinical training, and research, and this definitely helped me to finish this dissertation in a timely manner. Finally, to my parents Äiti ja Isä, brothers Marko and Simo, sister-in-law Gill, nephews Hendrik and Sawyer, and grandmother Mummu. Kiitos paljon for everything that you do for me. I promise that although this is the last of my student years, I will always continue to learn.

CHAPTER 1. GENERAL INTRODUCTION

Reading Development

Skilled and efficient reading involves various linguistic and cognitive processes, including an understanding of the auditory and visual components of words as well as meaning of words. For instance, models of reading and reading disability propose that word recognition occurs in pathways that involve phonological, orthographic, and semantic components (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 1999). Some researchers have modelled reading through direct or indirect routes, which allows a person to read regular (e.g., 'save') and irregular (e.g., 'have') words as well as to decode pronounceable non-words correctly (e.g., 'mave'; as discussed in Seidenberg, 2005). Other researchers, however, have suggested that reading is acquired through statistical learning and that multiple connections between the visual components, phonemic structure, and meaning of words leads to word recognition (Harm & Seidenberg, 1999). Finally, an understanding of proper sentence structure as well as adequate working memory and cognitive processing speed are essential when reading extended text, so that the reader can understand and follow the story line. All of these processes have been implicated in typical reading development as well as in reading disabilities.

Children who are learning to read first learn about phonological awareness (Stanovich, 1988), which is the ability to detect, manipulate, and think about the sound structure in spoken language. However, orthographic skills which involve the acquisition of a mental 'dictionary' of words, are also important for reading development. Children first rely on phonology and, with experience, switch to orthography; this effect has been shown in reading aloud (Backman, Bruck, Hebert, & Seidenberg, 1984; Waters, Seidenberg, & Bruck, 1984) and in silent reading (Coltheart, Laxon, Rickard, & Elton, 1988; Johnston, Thompson, Fletcher-Flinn, & Holligan, 1995). Non-word reading tasks are typically used to assess phonological decoding ability, whereas reading or spelling of irregular words are used to assess orthographic skills. Research has shown that both phonological and orthographic skills contribute to later reading success or failure (MacDonald & Cornwall, 1995; Sprenger-Charolles, Siegel, Béchennec, & Serniclaes, 2003; Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997).

Beyond phonological and orthographic awareness, higher-level language skills are also associated with reading development. Semantics is the knowledge of word meaning, syntactic awareness relates to proper sentence structure, and morphological awareness is the knowledge of how words consist of smaller parts, such as prefixes, suffixes, and roots. Children with better semantic, syntactic, and morphological awareness have better reading skills (Biemiller & Boote, 2006; Jongejan, Verhoeven, & Siegel, 2007; Muter, Hulme, Snowling, & Stevenson, 2004; Nagy, Berninger, & Abbott, 2006; Nation & Snowling, 1998). Furthermore, both expressive and receptive language domains are related to reading, which includes the ability to put thoughts into words and sentences (expressive) and the ability to understand what has been said (receptive). For example, Nation & Snowling (2004) demonstrated that performance on expressive vocabulary and receptive listening tasks predicted unique variance in reading comprehension, even after accounting for variance in phonological awareness and non-word reading. Overall, these studies illustrate that there are multiple linguistic factors involved with reading, and these are typically measured by expressive or receptive language tasks.

Several cognitive processing skills have also been related to the reading process. Short-term memory includes components of memory span and working memory. Memory span refers to the amount of information that can be held in the short-term, whereas working memory requires the simultaneous retention and manipulation of information (Baddeley, 2003). There are greater working memory demands in complex tasks, such as reading comprehension. Research has shown that working memory is related to reading ability (Daneman & Carpenter, 1980; Georgiou, Das, & Hayward, 2008; Gottardo, Stanovich, & Siegel, 1996; Swanson & Howell, 2001), and it may also contribute unique variance to reading above phonological awareness (McCallum, 2006). In processing speed tasks, such as rapid naming, children are typically asked to name letters, colours, numbers, or simple objects as quickly as possible (Denckla & Rudel, 1976). It has been suggested that rapid naming is related to reading because both require the efficient retrieval of linguistic and perceptual information (Norton & Wolf, 2012). To this end, many studies have shown that performance on rapid naming tasks is a predictor of later reading achievement (Compton, 2003; de Jong & van der Leij, 1999; Kirby, Parrila, & Pfeiffer, 2003; Landerl & Wimmer, 2008; Wagner et al., 1997). Other measures of processing speed, including speeded motor and non-naming tasks, have also been related to reading performance in children (Kail, Hall, & Caskey, 1999). Overall, these studies indicate that cognitive factors, such as working memory and processing speed, contribute to reading accuracy and efficiency.

In summary, the aforementioned studies demonstrate that there are many linguistic and cognitive processes associated with skilled reading. A child who has acquired phonological and orthographic awareness and has adequate language and reading-related cognitive skills has the foundation for becoming a skilled reader. When a child has persistent difficulties with reading and they do not attain reading proficiency, this is considered a reading disability.

Reading Disabilities

It is estimated that 5 to 12% of the population has a reading disability (Katusic, Colligan, Barbaresi, Schaid, & Jacobsen, 2001), which is currently clinically defined as reading performance that is substantially below age-expected levels despite interventions in these areas (American Psychiatric Association, 2013). Also, it cannot be attributed to intellectual disabilities, sensory deficits, lack of educational opportunities, cultural or linguistic differences, or social, emotional, or physical health. As part of the DSM-5 diagnosis of a Specific Learning Disorder in reading, the person has difficulty with word reading accuracy, reading fluency, and/or reading comprehension. However, they also often struggle with academic skills such as spelling and math (Curtin, Manis, & Seidenberg, 2001; Vukovic, Lesaux, & Siegel, 2010). It is well established that children with reading impairments struggle with phonological encoding and decoding. Many research groups have shown that, in comparison to average readers, poor readers have increased errors on phoneme-level and word-level tasks, such as identifying sounds in words and reading pronounceable non-words (Bruck, 1992; Snowling, 1981; Stanovich & Siegel, 1994). Additionally, some poor readers have deficits in orthographic processing, which includes a lack of awareness for irregular words (e.g., ‘yacht’) that do not have phonological representations (Castles & Coltheart, 1993). Most children with reading disabilities have problems with phonological processing, and many have orthographic deficits as well (Castles & Coltheart, 1993; Edwards & Hogben, 1999). Pure orthographic and phonological subtypes of reading disabilities are rare, and one review suggested that phonological deficits are more common than orthographic deficits (Sprenger-Charolles, Siegel, Jiménez, & Ziegler, 2011).

In addition to orthographic and phonological impairments, people with reading disabilities often struggle on a variety of linguistic and cognitive processing tasks. For example, in comparison to typically-developing readers, children and adults with reading disabilities have difficulty on tasks of syntactic awareness (Siegel & Ryan, 1988; Willows & Ryan, 1986), morphological awareness (Siegel, 2008), and vocabulary knowledge (Ricketts, Nation, & Bishop, 2007). These types of challenges have been shown in both expressive and receptive language domains (Catts, Fey, Zhang, & Tomblin, 1999; Wise, Sevcik, Morris, Lovett, & Wolf, 2007). Furthermore, it has been consistently shown that poor readers have difficulty with working memory (Swanson, 1993; 1994), rapid naming (Badian, 1993; Wolf, 1986), and speeded motor (Shanahan et al., 2006) tasks. Longitudinal studies have illustrated that poor performance on language and cognitive tasks in young children is predictive of later reading impairments (Badian, 1995; de Jong & van der Leij, 1999; Muter et al., 2004) and thus, screening of linguistic and reading-related cognitive skills is often included in determining children who are at-risk of developing reading problems (Scanlon, Vellutino, Small, Fanuele, & Sweeney, 2005; Vellutino, Scanlon, Small, & Fanuele, 2006).

In addition to poor academic performance and processing impairments, children with reading difficulties may have negative mental health and social outcomes. Attention deficit/hyperactivity disorder (ADHD) is the most common comorbid diagnosis; it has been estimated that between 18% to 45% of children with reading disabilities also meet the criteria for ADHD (Germanò, Gagliano, & Curatolo, 2010). Furthermore, internalizing disorders such as depression and anxiety have been documented in children with reading disabilities at all ages (Mugnaini, Lassi, La Malfa, & Albertini, 2009). Children who struggle with school, including those with reading impairments, may have decreased feelings of self-efficacy and self-worth, which may lead to internalizing symptoms. There are also increased rates of social problems such as anti-social behaviour and homelessness (Barwick & Siegel, 1996). These negative effects of reading impairments may persist into adulthood (Undheim, 2003), thus it is important to improve reading skills at an early age.

Reading Interventions

There are many reading intervention programs available to educators, and therefore educators face the difficult decision of which ones to include in their lesson plans. It has been suggested that comprehensive reading programs should include an emphasis on phonological awareness, phonics (letter-to-sound correspondence), vocabulary, reading fluency, and reading comprehension (National Institute of Child Health and Human Development, 2000). There have been numerous efforts to develop programs that include one or more of these elements. Given the vast number of programs available, the following review will include those that have been widely used or evaluated with elementary school students, as this is the age range for the studies in this dissertation.

Standardized Reading Programs

Reading Recovery (Clay, 1993) is one program that has been extensively studied. It is intended for first grade students who have previously shown limited gains following high-quality classroom instruction and it is delivered by trained teachers for 20 weeks in a one-to-one format. Lessons are focused on phonological awareness, phonics, fluency, comprehension, and spelling, depending on the student's needs. If the student is not reading at their grade-level by the end of the program, they are referred for special education services. Several meta-analyses of Reading Recovery have found moderate effect sizes after intervention, with the greatest gains shown for print knowledge and text reading (D'Agostino & Harmey, 2016; D'Agostino & Murphy, 2004; Slavin, Lake, Davis, & Madden, 2011). However, the utility of this program has been questioned. For example, it may not be effective for students with poor phonological awareness (Chapman, Tunmer, & Prochnow, 2001) or those who are most at-risk of reading failure (Reynolds & Wheldall, 2007), which indicates that students who need intervention may not learn the strategies they need to develop reading proficiency. Furthermore, improvement after Reading Recovery has been primarily on measures that were designed for the program (D'Agostino & Murphy, 2004), called Observation Survey of Early Literacy Achievement (Clay, 2005), and thus positive results should be interpreted with caution. Finally, Reading Recovery may not be more beneficial than other one-to-one interventions (Elbaum, Vaughn, Tejero Hughes, & Watson Moody, 2000) and it has been suggested that changes to the program along with further methodological rigor are needed to provide support for this program (Center, Wheldall, & Freeman, 1992; Reynolds & Wheldall, 2007; Tunmer & Chapman, 2003).

Programs from Lindamood-Bell have also garnered some research. There are several reading-focused programs at this center, including the Lindamood Phoneme Sequencing Program (LiPS; Lindamood & Lindamood, 1998), which was previously called Auditory Discrimination in Depth (ADD; Lindamood & Lindamood, 1975), as well as the Seeing Stars Symbol Imagery Program (Bell, 1997). The LiPS/ADD program places greater emphasis on developing phonological skills whereas the Seeing Stars program focuses more on building orthographic skills. Studies from several research groups have shown gains following intervention, particularly in phoneme awareness for the LiPS program (Pokorni, Worthington, & Jamison, 2004) and "symbol imagery" (a measure of orthographic awareness) for the Seeing Stars program (Christodoulou et al., 2017). While some of these studies showed improvements that extended to word reading and comprehension skills (Torgesen et al., 2001), others did not (Christodoulou et al., 2017; Pokorni et al., 2004). Porkoni and colleagues also indicated that many students continued to have significant reading deficits after the intervention.

The purpose of the Phono-Graphix program (McGuinness, McGuinness, & McGuinness, 1996) is to build phonological awareness and phonics skills in an intensive format, however, it has received less empirical review. Several small studies indicated gains in phonological (Dias & Juniper, 2002; Wright & Mullan, 2006) and basic reading (Endress, Weston, Marchand-Martella, Martella, & Simmons, 2007) skills after treatment, the latter of which is a composite score of word identification and decoding skills. A more rigorous research design was conducted by Denton and colleagues who randomly assigned students to interventions, one of which included the Phono-Graphix program (Denton, Fletcher, Anthony, & Francis, 2006). There were moderate to large gains in word recognition, decoding, phonemic decoding fluency, and comprehension, however, the authors noted that many students continued to read in the below average range after intervention.

Read Naturally (Hasbrouck, Ihnot, & Rogers, 1999) is designed to promote fluent reading skills for students between the second and sixth grades, using strategies such as repeated readings and teacher modelling of reading. A computerized version of Read Naturally similarly emphasizes repeated readings and computer modelling of reading. Most published studies of Read Naturally have been case reports, which have shown improved oral reading fluency with tutor-led (Erickson, Derby, McLaughlin, & Fuehrer, 2015) or computer-led (Gibson, Cartledge, Keyes, & Yawn, 2014; Keyes, Cartledge, Gibson, & Robinson-Ervin, 2016) instruction. As described above for the Phono-Graphix program, Denton et al. (2006) also evaluated Read Naturally in a more rigorous design. There were large gains in sight word fluency and paragraph reading fluency after this program, but many students continued to have significant reading impairments.

The Orton-Gillingham multisensory method has been widely used and focuses on increasing skills in phonological awareness, phonics, morphology, syntax, and vocabulary

(Academy of Orton-Gillingham Practitioners and Educators, Inc.). The main feature of Orton-Gillingham instruction is that it involves visual, auditory, and somatosensory (i.e., movement or touch) modes of presenting information. However, a review of this approach indicated that while some studies had improved reading outcomes following intervention, other studies did not (Ritchey & Goeke, 2006). Furthermore, a recent case study of 5 students with reading impairments showed no added benefit of a multisensory program to structured language instruction (Schlesinger & Gray, 2017). The support for the Orton-Gillingham multisensory approach is largely anecdotal and empirical evidence is minimal at this time, however, it remains an area of active research.

There are also several computerized programs that have been developed to remediate reading deficits, such as GraphoGame (Lyytinen, Erskine, Kujala, Ojanen, & Richardson, 2009), Reading Acceleration Program (Breznitz et al., 2013), and Fast ForWord Language (Scientific Learning Corporation). In GraphoGame, the child listens to phonemes and then chooses the letter on the screen that matches the phoneme. The game progresses to syllables, words, and non-words after the child has developed letter-to-sound correspondence skills. Several studies have shown that poor readers in elementary school had improved literacy skills immediately after this program and up to one year later (Heikkilä, Aro, Närhi, Westerholm, & Ahonen, 2013; Jere-Folotiya et al., 2014; Kyle, Kujala, Richardson, Lyytinen, & Goswami, 2013; Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2010). In the Reading Acceleration Program, the goal is to develop better reading fluency and comprehension by introducing time constraints when reading. Several studies have demonstrated gains in reading fluency and comprehension following this program, which have been shown in children and adults with and without reading difficulties (Breznitz et al., 2013; Horowitz-Kraus, Vannest, & Holland, 2013; Horowitz-Kraus et al., 2014; Niedo, Lee, Breznitz, & Berninger, 2014; Snellings, van der Leij, de Jong, & Blok, 2009). Finally, Fast ForWord is designed to improve reading and language skills by using acoustically-modified speech patterns, which is based on the theory that poor readers have difficulty perceiving rapid speech transitions (Tallal, 1980). Initial evidence showed that language-delayed children improved on language tasks following intensive intervention (Merzenich et al., 1996; Tallal et al., 1996). However, other studies have shown no gains in reading after this program (Agnew, 2004; Pokorni et al., 2004) and it may not be more effective than standard practice or other computer-based language interventions (Gillam et al., 2008; Strong, Torgerson, Torgerson, & Hulme, 2011).

There are numerous programs that have been developed to ameliorate reading deficits and further reviews are provided by What Works Clearinghouse, a national center in the United States (https://ies.ed.gov/ncee/wwc/). Consistent findings, however, are that gains may not extend to reading fluency or reading comprehension and that longitudinal evaluations are lacking. Furthermore, although early intervention is successful for many students, there is large variability in intervention response such that 8 to 80% show minimal or no improvement, depending on the study (reviewed in Al Otaiba & Fuchs, 2002). Therefore, these students will require ongoing and intensive programming to support their academic progress.

Evaluation of Intensive Reading Interventions

Intensive reading programs are typically defined as small group or individual supports that occur on a daily basis for 100 or more sessions of instruction (Denton, 2012; Vaughn & Wanzek, 2014). Using standardized reading programs or protocols, studies have illustrated the importance of increased intervention intensity for students in the primary grades. Interventions that started in Kindergarten or Grade 1 primarily focused on increasing phonemic awareness through programs such as the Orton-Gillingham approach (Foorman et al., 1997), Open Court reading series (Simmons et al., 2008), Reading Mastery program (Simmons et al., 2008), or SpellRead program (Metsala & David, 2017; Rashotte, MacPhee, & Torgesen, 2001), and activities such as segmenting and blending games, sight word instruction with flash cards and spelling, and repeated readings (Blachman, Tangel, Ball, Black, & McGraw, 1999; Scanlon et al., 2005; Simmons et al., 2008; Torgesen et al., 1999). Improved early literacy and reading skills were shown immediately after program completion, and up to 10 years later (Blachman et al., 2014). Across studies, programs were delivered for 30 to 60 minutes per day for 2 to 5 days per week for a total of 25 to 173 hours of instruction (Scammacca, Vaughn, Roberts, Wanzek, & Torgesen, 2007).

In a more extensive daily intervention, Torgesen and colleagues examined reading programs that occurred for 100 minutes per day over 8 to 9 weeks, for a minimum of 67.5 hours of instruction (Torgesen et al., 2001). They compared the Lindamood Auditory Discrimination in Depth (ADD) and Embedded Phonics (EP) programs, which both emphasized phonemic awareness but the EP program included more practice with word reading and comprehending text. Students between 8 to 10 years old who were previously identified with a learning disability showed significant growth in all reading measures with mean scores within the standardized average range. Although the ADD program had slightly larger gains immediately after intervention, both programs had similar performance at 1- and 2-year follow-up sessions. A recent randomized control study examined the Lindamood-Bell Seeing Stars program (Bell, 1997), which focuses on developing phonological and orthographic awareness, sight word recognition, and reading comprehension for children 6 to 9 years old (Christodoulou et al., 2017). This program was implemented for 4 hours per day over 6 weeks in the summer for a total of 100 to 120 hours of instruction, and this was compared to a “non-intervention” group (i.e., children did not receive instruction from Lindamood-Bell, however, some had private instruction from other tutors). Children showed improvements in oral reading fluency and orthographic awareness after the intensive program, whereas children in the non-intervention group showed declines in word and non-word reading.

Researchers have also examined multi-tier or response to intervention (RTI) models. This model starts with high-quality instruction for all students in the classroom using research-supported strategies (i.e., tier 1). If a student is not showing adequate progress from universal classroom instruction, they are provided with small group supports (i.e., tier 2) and then more intensive and individualized programming if they do not improve from small group supports (i.e., tier 3). Although there are many studies that have examined tier 1 and 2 programs (see Denton, 2012 for a review), fewer have specifically evaluated tier 3 programs. For example, several research groups have shown significant gains in reading when standardized and intensive protocols were used in the primary grades, which is promising given that these students showed a previous low response to universal or small group instruction. These studies included programs such as Early Interventions in Reading (Al Otaiba et al., 2014), Phono-Graphix (Denton et al., 2006), and Read Naturally (Denton et al., 2006), or activities that emphasized phonemic awareness, sight-word recognition, decoding, and reading fluency (Gilbert et al., 2013), as well as vocabulary, passage reading, and comprehension (Vaughn, Wanzek, & Murray, 2009). Instruction was provided between 30 and 100 minutes per day for 4 or 5 days per week for a total of 35 to 120 hours of instruction.

These studies have provided some support for standardized and intensive programs to improve reading skills, including an emphasis on phonemic awareness, decoding, word study, guided reading, writing exercises, and comprehension strategies (Scammacca et al., 2007) as well as using a multi-tiered approach (Gersten et al., 2009). However, standardized protocols on their own may not be beneficial for students who have the poorest reading skills. For example, Gilbert and others (2013) used a 3-tiered program and classified students as being “responsive” or “unresponsive” to tier 1 and tier 2 instruction based on performance on word identification measures. Students who were “unresponsive” were then randomly assigned to continue receiving tier 2 instruction to serve as a control group or to receive daily one-to-one tier 3 instruction. The same standardized program was used, where the difference between the two groups included increased frequency of support and decreased student-to-teacher ratios. These groups had similar reading skills prior to intervention and showed improvements after further instruction, however, there were no differences between the two groups in reading outcomes despite the increased intensity. Therefore, the authors suggested that standardized programs without individualization may not be beneficial for the lowest-performing students. For example, an important aspect of individualized programs is that research-supported strategies continue to be implemented (e.g., exercises with phoneme awareness), however, the specific content, activities, and materials may vary for each student.

Evaluations of both individualized and intensive programs have been scarce. In one study, multiple tiers of support were provided to students between Kindergarten and the third grade (O'Connor, Harty, & Fulmer, 2005). Ten students completed an intensive and individualized program, which occurred for 30 minutes per day on a daily basis, and the total amount of instructional time per student varied according to when they started the program and their individual progress on phonemic awareness, letter knowledge, and reading tasks. The results showed that 40% of these students had improved reading skills after intervention and they maintained average reading performance until the third grade. However, the remaining 60% continued to have poor reading skills and they required ongoing intensive supports. A randomized control trial evaluated a tier 3 individualized program in comparison to typical school instruction (Denton et al., 2013). Grade 2 students who had inadequate response to tier 1 and tier 2 instruction were randomly assigned to either the tier 3 program or typical school instruction. The tier 3 program occurred daily for 45 minutes for approximately 76.5 hours total. All teachers emphasized word reading, fluency, comprehension, and writing in each session, however, they differentiated instruction and objectives for each student. The tier 3 program showed greater gains than regular school instruction in word- and sentence-level reading, but not in extended text reading. Finally, studies with older children have similarly shown that intensive supports lead to higher reading scores when compared to a control group, however, no differences between individualized and standardized programs were found (Vaughn et al., 2009; 2011).

Overall, these studies suggest that intensive programming may be most effective for students with the poorest reading skills. However, it has been consistently shown that a proportion of students continue to read below grade-level after intervention. Also, gains in reading fluency and comprehension have been smaller than other reading skills such as word recognition and non-word decoding (Simmons et al., 2008; Torgesen et al., 2001; Vaughn et al., 2009), and each student's response to intervention may be quite variable (Denton et al., 2006). Furthermore, the evidence for using individualized instruction rather than standardized programs is mixed. It is possible that individualization of instruction leads to greater gains than standard (packaged) programs because the educator specifically targets the areas of weakness for that student, but further evidence is needed. *Overall these results suggest that further evidence is needed for intensive and individualized programs, and thus the purpose of my first study (Chapter 2) was to examine reading programs that used a combination of intensity and individualization of instruction.*

Neurobiological Bases of Reading

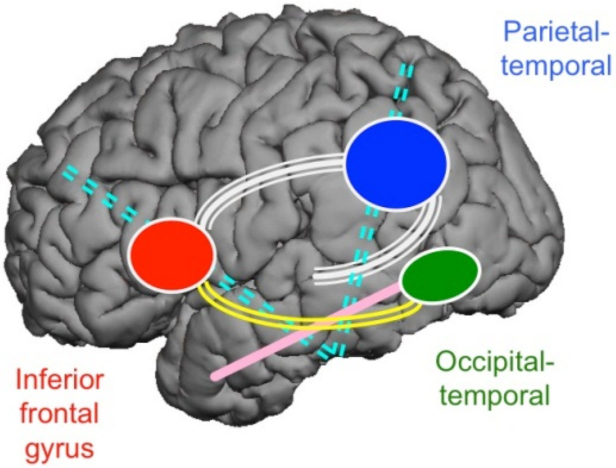
Historically, it has been thought that a reading disability is characterized by letter reversals, but research has shown that it is a more complex disorder with a neurobiological basis (Lyon, Shaywitz, & Shaywitz, 2003). Neuroimaging studies have shown that both brain function and brain structure change with development, and there are differences between typical readers and those with reading impairments on these measures.

Brain Function

There are multiple non-invasive neuroimaging methods that are used to measure brain function, such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), or magnetoencephalography (MEG). fMRI measures change in blood oxygenation level dependent (BOLD) signal (Ashby, 2011), EEG measures change in neuronal electrical fields (Brandeis, Michel, & Amzica, 2009), and MEG measures change in neuronal electromagnetic fields (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). Behavioural tasks can be presented to the participant in the scanner to evaluate fluctuations in BOLD, electrical fields, or electromagnetic fields, which are then used to infer brain activity during a specific task such as reading. On the other hand, an assessment of the brain in "resting state" occurs when the participant is not conducting a specific task and they are instructed to lay still while images are being obtained (Cole, Smith, & Beckmann, 2010). These images are then typically used to infer functional connectivity between brain regions based on signal fluctuations. The advantage of MRI is the high spatial resolution (1-10 mm) in comparison to EEG and MEG which have lower spatial resolution (1-10 cm; Sakkalis, 2011). Conversely, the advantage of EEG and MEG is the high temporal precision (<1 ms) in comparison to fMRI which has longer temporal resolution (1-3 sec; Sakkalis, 2011). All of these methods have been used to measure the development of reading and reading disability.

A well-supported neurobiological network of reading involves three primary regions in the left hemisphere, including occipital-temporal, parietal-temporal, and inferior frontal regions (Shaywitz, Mody, & Shaywitz, 2006) (see Figure 1.1). Young children who are learning to read activate the left parietal-temporal and occipital-temporal cortex during word recognition and phonological reading tasks (Simos et al., 2001), and the left inferior frontal gyrus is engaged with increasing age and reading ability (Turkeltaub, Gareau, Flowers, Zeffiro, & Eden, 2003). Some studies have shown greater activation for adults in left frontal, temporal, and parietal cortices (Booth et al., 2004; Simos et al., 2001; Turkeltaub et al., 2003), whereas others have found greater activity for children in primarily posterior regions of the brain (Church, Coalson, Lugar, Petersen, & Schlaggar, 2008). A recent meta-analysis indicated that activation patterns for children and adults are similar during reading tasks, including regions of occipital-temporal, posterior parietal, and inferior frontal cortex (Martin, Schurz, Kronbichler, & Richlan, 2015). However, developmental differences were shown in left superior temporal and bilateral supplementary motor regions for children and in bilateral posterior occipital-temporal, bilateral cerebellum, and left precentral cortex for adults. These results indicate that there are consistent cortical regions involved with reading, and there are some regions that change with development and experience.

Figure 1.1. Brain regions implicated in reading development and impairment



Note: Grey matter regions are shown as circles with corresponding text. White matter tracts are depicted as: 3 white lines (arcuate fasciculus), 2 yellow lines (inferior fronto-occipital fasciculus), solid pink line (inferior longitudinal fasciculus), and 2 teal dashed lines (corona radiata). Adapted from (Vandermosten, Boets, Wouters, & Ghesquière, 2012b).

The models of reading, discussed in the Reading Development section, are also supported by functional neuroimaging studies. A dorsal phonological system may include parietal-temporal, superior temporal, and inferior frontal cortices, whereas a ventral orthographic system may include middle temporal and occipital-temporal regions (Jobard, Crivello, & Tzourio-Mazoyer, 2003; Schlaggar & McCandliss, 2007). Furthermore, semantic knowledge may be processed in similar areas including angular, middle temporal, and fusiform gyri (Taylor, Rastle, & Davis, 2013). Many studies have shown that part of the fusiform gyrus is responsive to visual letter and word stimuli, suggesting it is part of the orthographic system (McCandliss, Cohen, & Dehaene, 2003), and that inferior frontal regions are activated on rhyming tasks, suggesting they are part of the phonological system (Bitan et al., 2007; Pugh et al., 1996). The mapping between orthography and phonology (e.g., spelling-sound conversion) may occur in parietal-temporal cortex, which includes angular and supramarginal gyri (Booth et al., 2004). Furthermore, differential processing of phonological and orthographic codes has also been shown in developmental studies. For example, adults had higher levels of activity than children within inferior frontal and angular gyri on tasks of spelling and rhyming (Booth et al., 2004). Overall, these results indicate that separate, as well as some overlapping, cortical regions are used for reading, which is consistent with behavioural reading models (Taylor et al., 2013).

As summarized in Shaywitz et al. (2002), children with reading disabilities show different brain activation patterns than control groups during reading. Specifically, they have decreased activity in left posterior regions, including left parietal-temporal and occipital-temporal cortex (Horwitz, Rumsey, & Donohue, 1998; Temple et al., 2001), and some may also have increased activity in right inferior frontal and middle temporal lobes during reading. This pattern of activation is not observed in average readers (Hoeft et al., 2007a; Shaywitz et al., 2003). Poor readers also have altered activation in left parietal-temporal regions on phonological tasks and in left occipital-parietal regions on orthographic tasks (Temple et al., 2001). It has been suggested that differences in temporal regions reflect deficits in orthographic processing and semantics, whereas differences in frontal and parietal regions reflect deficits in the mapping between orthography and phonology (Cao, Bitan, Chou, Burman, & Booth, 2006). Furthermore, differences between typically-developing and poor readers have been shown on a variety of reading and language tasks in the scanner, including phonological awareness (Kovelman et al., 2012), word and non-word reading (Brambati et al., 2006), reading fluency (Langer, Benjamin, Minas, & Gaab, 2015), sentence comprehension (Meyler et al., 2007), and passage reading (Simos, Rezaie, Fletcher, Juranek, & Papanicolaou, 2011). However, functional differences exist even when participants are not explicitly told to read (Brunswick, McCrory, Price, Frith, & Frith, 1999), are viewing letter strings (Reilhac, Peyrin, Démonet, & Valdois, 2013), or during rest (Finn et al., 2014). In summary, these results indicate that poor readers have altered brain activation patterns in comparison to typical readers, and this effect has been shown on various reading and non-reading tasks.

Brain Structure

Diffusion tensor imaging (DTI) is commonly used to model white matter structures (myelin) by measuring the amount of diffusion or movement of water molecules in the brain (Alexander, Lee, Lazar, & Field, 2007). In less restricted areas such as cerebrospinal fluid, there is a lot of movement of water molecules and therefore diffusion is high (i.e., more isotropic). In white matter, movement is more restricted and water molecules are more likely to follow the path of the axons, and thus diffusion is low (i.e., more anisotropic). These properties of water diffusion are used to provide an indirect assessment of white matter organization. Fractional anisotropy is one metric that is commonly used in DTI studies, which provides a measure of the amount of diffusion that is anisotropic within a specified brain region. It is expected that areas of high white matter organization will have high fractional anisotropy. In contrast, areas that are less developed or damaged by disease will have lower white matter organization and thus lower fractional anisotropy. Another DTI method is tractography, which models white matter tracts using 3D connectivity patterns (Alexander et al., 2007). Tracts are estimated by starting at a pre-determined location, called the seed point (e.g., inferior frontal gyrus), and a probabilistic algorithm estimates the direction of the tract until it reaches a termination point (e.g., angular gyrus). Mean fractional anisotropy can then be calculated within the tract to provide an assessment of white matter organization.

Myelination of the brain increases with age (Qiu, Tan, Zhou, & Khong, 2008), and increased fractional anisotropy in various brain regions or tracts has been associated with better phonological awareness (Yeatman et al., 2011), word reading accuracy (Beaulieu et al., 2005; Klingberg et al., 2000), non-word reading accuracy (Klingberg et al., 2000), word reading fluency (Nagy, Westerberg, & Klingberg, 2004), and paragraph reading fluency (Lebel et al., 2013). There may be at least 4 white matter tracts involved with the reading process (see Figure 1.1) including: arcuate fasciculus, which connects frontal, parietal, and temporal regions; inferior fronto-occipital fasciculus, which connects occipital and frontal regions; inferior longitudinal fasciculus, which connects occipital and temporal regions; and corona radiata, which connects subcortical structures with cortex (Catani & Thiebaut de Schotten, 2008; Vandermosten et al., 2012b). DTI tractography studies have linked the dual route model of reading to specific white matter tracts. For example, the arcuate fasciculus may be part of the dorsal phonological route (Vandermosten, Boets, Poelmans, Sunaert, Wouters, & Ghesquiere, 2012a; Yeatman et al., 2011), whereas the inferior fronto-occipital fasciculus may be part of the ventral orthographic route (Vandermosten et al., 2012a). These results indicate that white matter microstructure and white matter tracts are important for reading development.

Children and adults with reading impairments have decreased fractional anisotropy in the left parietal-temporal region (Deutsch et al., 2005; Klingberg et al., 2000), implicating deficits in the arcuate fasciculus and corona radiata (Vandermosten et al., 2012a). There have also been reports of increased fractional anisotropy in children with reading difficulties compared to typical readers, including areas of the left inferior frontal gyrus, superior temporal gyrus, fusiform gyrus, middle occipital gyrus, putamen, and hypothalamus (Rimrodt, Peterson, Denckla, Kaufmann, & Cutting, 2010). Tractography studies have shown that differences between good and poor readers also exist within specific white matter tracts. These have included tracts that connected: putative "visual word form area" with areas of temporal, occipital, parietal, parahippocampal, and entorhinal cortex (Fan, Anderson, Davis, & Cutting, 2014a); thalamus with left prefrontal cortex and sensorimotor cortex (Fan, Davis, Anderson, & Cutting, 2014b); and the cerebellum with parietal-temporal, occipital-temporal, and inferior frontal regions (Fernandez et al., 2016). Overall, these studies suggest that poor readers have alterations in white matter microstructure and white matter tracts.

In addition to white matter, an assessment of grey matter is commonly conducted using T1-weighted MRI scans, which are static images that provide contrast between different structures in the brain. For example, cerebrospinal fluid appears dark and fat appears bright. Grey matter indices, such as volume, area, and thickness (Fischl & Dale, 2000), are identified based on the contrast differences between grey and white matter boundaries. Earlier studies of grey matter development typically focused on grey matter volume, which showed variations in volume as a child develops (Gogtay et al., 2004). However other studies have also assessed cortical surface area and cortical thickness, which are components of grey matter volume. Over time, children have increased thickness in regions such as inferior frontal and superior temporal gyri, but also decreased thickness in the rest of the brain (Sowell, 2004). Decreases in volume, area, and thickness continue into adolescence and adulthood (Lemaitre et al., 2012; Tamnes et al., 2010), and there are linear or curvilinear changes over time depending on the brain region (Gennatas et al., 2017; Tamnes et al., 2010). Furthermore, grey matter indices have been related to reading ability, including grey matter volume in left fusiform (Simon et al., 2013) and cortical thickness in left fusiform, left intraparietal sulcus, bilateral angular gyrus, and bilateral superior temporal gyrus (Blackmon et al., 2010).

Convergent findings have indicated that poor readers have reduced grey matter volume in left and right superior temporal regions (Richlan, Kronbichler, & Wimmer, 2012), and in left superior/middle temporal and orbitofrontal gyri (Eckert et al., 2016). Furthermore, differences between typically-developing and poor readers have been shown in cortical thickness and surface area. Increased surface area and grey matter volume was found in bilateral fusiform and inferior frontal gyri for typically-developing readers compared to poor readers, whereas right supramarginal gyrus was thicker for the poor reader group (Frye et al., 2010). These results, however, may be impacted by gender differences (Altarelli et al., 2013; Evans, Flowers, Napoliello, & Eden, 2013) or by low total grey matter volume (Eckert et al., 2016). Another area of debate has been whether poor readers have leftward asymmetry of the planum temporale in the superior temporal gyrus, which is typically shown in good readers (i.e., left larger than right). Inconsistencies between studies may be due to methodological issues, however, more rigorous studies have shown that poor readers have a lack of asymmetry between left and right planum temporale (Altarelli et al., 2014; Bloom, Garcia-Barrera, Miller, Miller, & Hynd, 2013). Furthermore, this effect may only be shown in boys who are poor readers (Altarelli et al., 2014). In summary, these results show that children with reading impairments may have neuroanatomical, as well as functional differences, that contribute to their poor reading skills.

Neuroimaging Studies of Reading Interventions

Many reading interventions have been developed to ameliorate reading difficulties, and some of these programs have also been evaluated with neuroimaging methods. These studies have typically focused on standardized reading protocols or programs, and primarily included measures of brain function rather than brain structure. Across studies, it has been shown that following treatment, poor readers show changes in brain activation patterns during reading and at rest, in white matter microstructure and connectivity, and in grey matter volume and cortical thickness.

Standardized Programs

Using computerized reading programs, several research groups have illustrated positive effects following treatment for reading impairments in children. These studies have focused on improving phonics skills with the GraphoGame program (Lyytinen et al., 2009), orthographic awareness through the Lindamood-Bell Seeing Stars program (Bell, 1997), and reading fluency with the Reading Acceleration Program (Breznitz et al., 2013). These programs were delivered for as short as 3 hours with 6-year-old children who were non-readers (Lovio, Halttunen, Lyytinen, Näätänen, & Kujala, 2012), and up to 4 to 8 weeks with children with reading impairments (Horowitz-Kraus et al., 2014; Horowitz-Kraus & Holland, 2015; Horowitz-Kraus, DiFrancesco, Kay, Wang, & Holland, 2015a; Horowitz-Kraus, Toro-Serey, & DiFrancesco, 2015b; Krafnick, Flowers, Napoliello, & Eden, 2011). These studies illustrated that improvements in letter knowledge, basic reading, fluency, and comprehension were associated with changes in brain activity during a lexical decision task (Horowitz-Kraus et al., 2014), event-related potentials (Lovio et al., 2012), functional connectivity patterns during a resting state paradigm (Horowitz-Kraus & Holland, 2015; Horowitz-Kraus et al., 2015a; 2015b), and grey matter volume in the fusiform gyrus, hippocampus, precuneus, and cerebellum (Krafnick et al., 2011). Overall, these results indicate that after relatively short computerized programs, there are gains in reading skills, functional brain activity, and grey matter volume.

Some neuroimaging studies have included various standardized reading programs within one study and/or programs were not described in detail. However, similarly to other studies, improved reading skills were associated with increased functional activity in right inferior frontal gyrus during a rhyming task (Hoeft et al., 2011) and with increased event-related potentials during a phonological lexical decision task (Hasko, Groth, Bruder, Bartling, & Schulte-Korne, 2014). One research team examined the effects of intervention according to a child's history of reading supports and current reading skills. Children who had remediated spelling and reading skills had stronger functional connectivity between left fusiform and right middle occipital gyrus and between left fusiform and right medial prefrontal cortex (Koyama et al., 2013). In contrast, poor readers continued to have thicker cortex in left fusiform and right temporal lobe compared to typically-developing readers, and this effect was found whether the child received remediation or not (Ma et al., 2015). These authors argued that having a history of reading disabilities is associated with cortical thickness differences, and these differences persist even following treatment. However, these studies did not directly compare the effects of remediation in a prospective design and thus subtle changes in grey matter may not have been detected.

Intensive Programs

Intensive interventions, which includes 100 or more sessions of instruction or extensive daily intervention (i.e., 1 hour or more per day), have also been a focus of neuroimaging studies. Several programs have been evaluated, including Phono-Graphix (McGuiness et al., 1996), Lindamood Phoneme Sequencing (Lindamood & Lindamood, 1998), and Fast ForWord Language (Scientific Learning Corporation). Instruction was provided for 20 minutes to 3 hours per day and between 46.5 to 112.5 hours total. After intervention, children with reading impairments showed large gains in phonological awareness, word reading, and reading comprehension as well as increased activation in left superior temporal gyrus on a non-word rhyming task (Simos et al., 2002), bilateral areas on a letter rhyming task (Temple et al., 2003), and bilateral frontal, parietal, and thalamic regions on a rapid temporal processing task (Gaab, Gabrieli, Deutsch, Tallal, & Temple, 2007). In adults, gains in phonological skills and word reading accuracy were shown after intervention but not in reading rate or comprehension (Eden et al., 2004). Similarly to children, there were regions of increased brain activity including parietal and fusiform gyri as well as in right inferior frontal, superior parietal, and superior temporal cortex after intervention.

An examination of less specific, but intensive, programs has also shown changes in brain function and structure following treatment. In a series of studies, Richards and colleagues examined 3-week programs that emphasized phonological, orthographic, or morphological awareness. In addition to reading gains, there were changes in brain activity and functional connectivity after treatment that were alike to typical readers (Aylward et al., 2003; Richards & Berninger, 2008; Richards et al., 2000; 2002; 2006). Similar results were shown with 6- to 8-month programs, whereby activation patterns for poor readers resembled those of good readers immediately after intervention and one year later (Meyler, Keller, Cherkassky, Gabrieli, & Just, 2008; Shaywitz et al., 2004). Changes in white matter microstructure were also found after a 6-month intensive program (Keller & Just, 2009). Overall, these studies illustrated that brain activation patterns appeared to "normalize" after intensive instruction.

Response to Intervention

There have been few neuroimaging studies that have specifically examined the effects of RTI models or multiple tiers of instruction. In one randomized control trial, students who showed a low response to tier 1 instruction completed a tier 2 program for 17 weeks and then were identified as "responders" or "non-responders" (Davis et al., 2010; 2011). The non-responders had less activation in left angular gyrus compared to responders during a letter-sound matching task (Davis et al., 2011), and greater white matter connectivity between left angular gyrus and left insula was correlated with increased reading skills (Davis et al., 2010). These studies were relatively small (*n* = 5 for each group) and the programs were assessed at the end of intervention; thus, changes that occurred over time could not be ascertained. In a magnetoencephalography (MEG) study, students in Grades 6 to 8 with poor reading skills completed tier 2 instruction throughout one school year (Rezaie et al., 2011). MEG scans were obtained prior to intervention, and were retrospectively analyzed to determine differences between "adequate" and "inadequate" responders to treatment in comparison to typically-developing readers. The results showed that inadequate responders had reduced functional activity in middle and superior temporal gyri as well as occipital-temporal cortex in comparison to adequate responders and typical readers. Similarly to Davis and colleagues, only one time point of data was collected and thus analyses could not evaluate changes over time. Finally, one study has examined the effects of tier 3 instruction with students between 7 to 9 years old (Simos et al., 2007). These students completed Phono-Graphix and Read Naturally programs for 1 to 2 hours per day over 16 weeks. Using MEG, students completed a timed reading task in the scanner before and after each phase of intervention. Over time, there was increased brain activity in bilateral regions of middle temporal gyri, which was associated with improved in-scanner reading accuracy.

Several studies have evaluated the neurobiological basis of children's individual response to intervention. These studies were not RTI models per se, but they emphasized the importance of treatment success. In one study, students in Kindergarten and Grade 1 who had poor pre-reading skills received enhanced classroom instruction or pull-out group support for 40 minutes per day throughout the school year (Simos et al., 2005). Using MEG, results showed that there were differences in brain activity between those who had average reading skills and those who showed an adequate response to intervention. Similarly, several studies from one research group examined the effects of intensive intervention, which occurred for 4 days per week for a total of 230 to 333 hours of instruction (Farris et al., 2011; Farris, Ring, Black, Lyon, & Odegard, 2016; Odegard, Ring, Smith, Biggan, & Black, 2008). Children who were treatment "responders" had greater activity and functional connectivity in right frontal regions when compared to "non-responders" or typical readers (Farris et al., 2011; Odegard et al., 2008). Furthermore, growth in non-word and real word reading was associated with increased activity in bilateral frontal and insular cortex during an object rhyming task (Farris et al., 2016), and increased white matter fractional anisotropy in the corpus callosum was related to increased functional connectivity (Farris et al., 2011). These studies were conducted either before (Farris et al., 2016) or after (Farris et al., 2011; Odegard et al., 2008) intervention, however, a prospective analysis from pre-treatment to post-treatment has not been conducted.

In summary, these studies have shown that poor readers who have undergone reading remediation exhibit brain activation patterns closely resembling those of average readers, and there are also changes in white and grey matter indices. The location of these changes, however, has varied widely between studies. For example, changes have been found in bilateral frontal, parietal, temporal, and fusiform gyri as well as in corpus callosum, hippocampus, thalamus, and cerebellum. This variability in results indicates that response to intervention may vary according to a particular program or group of students. Furthermore, a prospective examination of intervention effects on both brain function and brain structure has not been conducted with magnetic resonance imaging (MRI), specifically with individualized programs. Finally, most of these studies included interventions that were implemented by the research team, which does not reflect the “real world” experience of how most students with reading impairments receive remediation. *Therefore, the purpose of my second and third studies was to examine the prospective changes in brain function (Chapter 3) and brain structure (Chapter 4) after school-based interventions, of which were provided in multiple tiers of instruction. These changes were assessed in students with poor reading skills as well as in typically-developing readers using several MRI techniques.*

Current Research Goals

The main purpose of this dissertation was to examine the academic and neurobiological outcomes following school-based reading interventions. First, I examined the immediate and one-year outcomes of an individualized, intensive reading program using a prospective longitudinal design (Chapter 2). Students were invited from one school district in Canada. Teachers identified students as belonging to one of 3 groups: good readers who were receiving universal instruction only (tier 1); poor readers who were receiving small group supports (tier 2); or poor readers who were starting the district’s intensive reading program (modified tier 3). This program was considered a modified tier 3 program because it was provided in a separate school from the student’s home school. Also, it was delivered by specialized teachers for most of the school day, or an average of 188 hours of instruction across 3 months. All students’ performance on tests of literacy, language, and cognition was measured before intervention, immediately after the 3-month intensive intervention ended, and one year after the intervention ended. My specific research questions in Chapter 2 were: 1) does individualized and intensive intervention lead to gains in literacy skills in comparison to small group instruction and typical development?; 2) are these gains in literacy maintained one year after the intensive intervention has ended?; and 3) which measures are predictive of persistently low reading skills? I hypothesized that changes in reading, spelling, and phonological awareness would be shown for both groups of poor readers after intervention, however, gains would be larger for students in the intensive program and this would be maintained one year later. I further hypothesized that baseline scores of reading, spelling, phonological awareness, rapid naming, and working memory would be predictive of persistent reading deficits immediately after the program or one year later.

Next, I evaluated whether there were changes in brain function after reading intervention, and if these results were modulated by type of reading task or relative difficulty (Chapter 3). The same group of students from the Chapter 2 study were invited to the studies in Chapters 3 and 4. In the MRI scanner, participants completed functional MRI tasks of orthography (spelling) and phonology (rhyming) before and after 3 months of instruction. These fMRI tasks were designed to have trials that were relatively easy or difficult based on printed word frequency. My research questions in Chapter 3 were: 1) are there changes in functional brain activity after school-based intervention on tasks of orthography and phonology?; 2) what are the functional neuroimaging predictors of improved reading skills?; and 3) are the effects of intervention shown on relatively easy or difficult reading tasks? I hypothesized that there would be differences in brain activity between good and poor readers on the spelling and rhyming tasks, and there would be changes following intervention for the poor readers. Furthermore, I hypothesized that baseline functional neuroimaging scores would be predictive of gains in reading skills over time. I did not have a specific hypothesis regarding the effect of difficulty on intervention outcome, as there have been previous conflicting results.

Finally, I examined changes in brain structure following intervention, which included both white matter and grey matter indices (Chapter 4). In addition to the functional MRI tasks described in Chapter 3, students also completed 3D-T1 anatomical and diffusion tensor imaging (DTI) scans before and after 3 months of instruction. I chose brain regions known to be involved with language processing, and evaluated grey matter volume, cortical thickness, and surface area within these regions and white matter microstructure between these regions. My research questions in Chapter 4 were: 1) are there changes in grey matter indices and white matter microstructure after intervention?; and 2) what are the structural neuroimaging predictors of improved reading skills? I hypothesized that following treatment, changes in both grey matter and white matter would be shown. I also hypothesized that measures more sensitive to change, such as cortical surface area, would be related to changes in reading.

This dissertation will provide further evidence for the immediate and long-term effects of school-based reading interventions, particularly for intensive and individualized programs. These results will also provide evidence for the multiple neurobiological factors that contribute to reading impairment and reading intervention.

CHAPTER 2. LONGITUDINAL OUTCOMES OF AN INDIVIDUALIZED AND INTENSIVE READING INTERVENTION

Introduction

The General Introduction described how early intervention is important for improving reading skills, however, between 8 to 80% of students show minimal or low response to treatment (reviewed in Al Otaiba & Fuchs, 2002). Therefore, these students need intensive and ongoing programming to support their reading impairments. Studies have illustrated the benefit of intensive interventions, however, a consistent finding across studies is that students with significant reading difficulties continue to read below grade-level after intervention and improvements may not extend to reading fluency or to comprehension (Simmons et al., 2008; Torgesen et al., 2001; Vaughn et al., 2009). It has been suggested that intensity and individualization can improve reading skills for these children, and that programs should be implemented in the context of multiple tiers of instruction (e.g., Gersten et al., 2009). However, further evidence for the combination of intensive and individualized programs is needed, as these are costly to implement and may or may not lead to a boost in reading skills.

Therefore, the purpose of the current study was to evaluate the immediate and one-year outcomes of an intensive, individualized reading program for third grade students who had previously received universal instruction and small group supports. This program had the unique arrangement of students attending all-day intervention in a separate classroom for 3 months, for an average of 188 hours of instruction. To my knowledge, this is the first behavioural study to examine this level of intensity and extent of intervention with students with significant reading difficulties. Students were invited from 23 elementary schools in one school district in British Columbia, Canada. Previous longitudinal studies in this district have shown the effectiveness of early universal intervention, including a reduction of reading disabilities between Kindergarten and Grade 7 (Lipka, Lesaux, & Siegel, 2006; Partanen & Siegel, 2014). In the current study, the purpose was to examine the outcomes of an intensive and individualized reading program. Performance of students in the intensive program was compared to that of poor readers who received small group intervention and to good readers who received universal instruction only.

Method

Ethics approval for this study was obtained from Clinical Research Ethics Boards at The University of British Columbia. Informed consent from parents and verbal assent from children was obtained prior to participation.

Participants

Students in Grade 3 were recruited from 23 elementary schools in one school district in British Columbia, Canada. Students were eligible for this study based on the types of supports they were receiving in school. Classroom teachers identified students as belonging to one of 3 groups: 1) good readers who were receiving no additional reading supports at school, 2) poor readers who were receiving small group supports, or 3) poor readers who were starting the district's intensive reading program. Students were ineligible for this study if they were currently in a French immersion instruction program or if they had known neurological impairments, developmental disorders, or uncorrected vision or hearing difficulties.

Following recruitment, we confirmed that students passed criteria as a "good" or "poor" reader. Good readers were defined as having at least average performance on measures of word recognition, decoding, fluency, and reading comprehension (i.e., > 95 standard score or > 37th percentile on all subtests). Poor readers were defined as at-risk or below average performance on the same reading measures (i.e., < 90 standard score or < 25th percentile on at least one subtest). The reading tests are described further below. With these criteria, there were 158 students at pre-test including 76 good readers with no additional reading supports, 42 poor readers with small group supports, and 40 poor readers in the intensive program. There were 21 students who were identified by their teacher as a good or poor reader, however they did not meet our reading group criteria and thus were not included in the current analyses. This included 8 good readers who obtained at least one standard reading score of 95 or below as well as 12 poor readers in the small group program and 1 poor reader in the intensive program whose reading scores were 90 or above on all reading subtests. There were several students in each group who moved to a different school, were no longer interested in participating, or were unavailable for testing and thus sample sizes varied for each time point.

Reading Programs

Universal instruction.

All students in this study received universal instruction from their classroom teacher. When students were in Kindergarten and Grade 1, phonological awareness training was emphasized through the school district's *Firm Foundations* program (North Vancouver School District, 2001). Activities focused on phonological awareness, early literacy, letter-to-sound correspondence, and language skills. In Grades 2 and higher, common strategies across schools were implemented using the school district's *Reading 44* program (North Vancouver School District, 1999). This program included 12 instructional activities that emphasized reading comprehension skills. For instance, strategies included obtaining prior knowledge, making predictions on what will be learned, determining unknown words by context, self-correcting, creating mental pictures, determining important ideas, getting information from text and pictures, summarizing, drawing conclusions, and reflecting on what was read.

Small group supports.

If a student was not showing adequate progress with universal instruction, they were eligible for small group supports from a certified Learning Assistance teacher, Learning Support Worker, and/or Educational Assistant. Most students were identified in Kindergarten or Grade 1 as having weak pre-reading skills on measures such as the Test of Phonological Awareness, Second Edition (TOPA-2; Torgesen & Bryant, 2004) or Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002). Activities and types of support were individualized to the student's needs. Examples of programs included Soar to Success (Houghton Mifflin Harcourt), Lexia Reading Core5 (Lexia Learning), Explode the Code (EPS Literacy), and RazKids (Learning A-Z). Strategies included guided reading, levelled readers, shared writing, scribing, graphic organizers, and using dictation software.

Students in small group programs typically received support for 40 minutes per day, for an average of 3.29 days per week (*SD* = 1.35) according to parent report. These supports have been in place since Kindergarten or Grade 1 (*M* = 3.03 years; *SD* = 0.74). Furthermore, parents indicated that 47% of these students had a history of after-school tutoring, which occurred for an average of 1.59 hours per week (*SD* = 0.61) since Grade 1 or 2 (*M* = 1.94 years, *SD* = 0.83). Some students also completed a psychoeducational assessment and received a diagnosis of a learning disability (11%), ADHD (8%), or both a learning disability and ADHD (6%). The other 75% of students did not have a current diagnosis.

Intensive reading program.

Students who showed inadequate progress from small group supports, as indicated by below average performance on measures such as DIBELS, Woodcock-Johnson Tests of Academic Achievement, Third Edition (WJ III ACH NU; Woodcock, McGrew, & Mather, 2001a, 2007), or Basic Reading Inventory (Johns, 2008), were eligible for an intensive reading program in the school district. The intensive program was offered to third grade students. Assessments were conducted to confirm that students met school district criteria for average reasoning and language abilities, a specific processing deficit (e.g., working memory) relative to their reasoning skills, and reading performance that was below age expectations (i.e., > 1 *SD* below the standardized norm for their age). Students were ineligible for this program if they had below average intellectual capacity, significant language delays, or if their reading difficulties were primarily due to mental health, behaviour, or second language acquisition.

The intensive program occurred every day for approximately 3 months with instruction from a certified teacher and Educational Assistant trained in the Orton-Gillingham approach. Each classroom had 7 students with the one teacher and one Educational Assistant. The primary emphasis was on teaching literacy skills for about 3.75 hours per day or an average of 188 hours per term (*SD* = 19.27). As the school day included approximately 5 hours of potential instructional time, this meant that most of the day was focused on literacy activities.

Furthermore, one-to-one instruction occurred for a minimum of 40 minutes per day with the teacher and 40 minutes per day with the Educational Assistant. The program was located in a separate school from the student's home school; therefore, students were immersed into the new school for the duration of the intervention (i.e., students continued to have math, physical education, and extracurricular activities). After program completion, students returned to their home school to continue receiving small group support and universal instruction as described above.

Activities and programs were tailored to the student's specific needs (e.g., word decoding, fluency) based on assessments that were conducted by school district psychologists and teachers prior to starting the program. Furthermore, teachers assessed student's reading skills over the course of the program with PM Benchmark materials (Nelson Education, Ltd.). Instruction was provided in one-to-one format with the teacher or support worker as well as in small groups, depending on the activity. Practice with decoding, sight words, and fluency included programs such as Phono-Graphix (Phono-Graphix Reading Company), Explode the Code (EPS Literacy), SRA Decoding Strategies (McGraw-Hill Education), and Read Naturally (Read Naturally, Inc.). Comprehensive programs included Read Well (Voyager Sopris Learning), Essential Skills Reading (Essential Skills Software, Inc.), Soar to Success (Houghton Mifflin Harcourt), RazKids (Learning A-Z), and Lexia Reading Core5 (Lexia Learning). The teachers also emphasized metacognitive awareness and social-emotional learning through programs such as MindUP (The Hawn Foundation) and Zones of Regulation (Social Thinking, Inc.). Examples of individualized strategies included flashcards, guided reading with an adult, repeated readings, building sight words and word families, puzzles with Dolch-list words, cloze procedures, vocabulary drills, graphic organizers, and shared writing. Parents were also required to complete daily homework with their children, which included reading to an adult for 15 minutes, math questions for 10 minutes, and practice with flashcards, decoding drills, and spelling.

Prior to starting the intensive program, students received small group supports for 40 minutes per day for an average of 3.81 days per week (*SD* = 1.10) since Kindergarten or Grade 1 (*M* = 3.12 years, *SD* = 0.81). About half (49%) of the students also received after-school tutoring, and this occurred for an average of 2.44 hours per week (*SD* = 1.64) since Grade 1 or 2 (*M* = 2.06 years, *SD* = 0.75). Finally, through a psychoeducational assessment, many students were diagnosed with a learning disability (59%), ADHD (3%), or both a learning disability and ADHD (18%). The remaining 20% of students did not have a formal diagnosis.

Study Design

In this prospective, longitudinal study, participants completed testing at 3 time points: before the intensive reading program started (pre-test), immediately after the 3-month intensive program finished (post-test), and one year after the intensive program finished when students were in the fourth grade (delayed test). All participants completed testing at similar intervals. As described below, measures of literacy, language, and cognition were completed at all time points, while screening tests and parent questionnaires were completed at pre-test only. Trained graduate students and research assistants administered the measures.

Of interest for this study was to compare the performance of the 3 groups of students over time. This included good readers who received universal instruction, poor readers who received universal instruction and small group supports, and poor readers who received universal instruction and small group supports followed by the intensive program.

Measures

Screening tests.

All students were screened for cognitive and visual impairments on their first visit. An estimate of intellectual functioning was conducted using the Test of Nonverbal Intelligence, 4th Edition (TONI-4; Brown, Sherbenou, & Johnsen, 2010). This test measures inductive reasoning, which required students to identify the rule underlying a visual pattern. In the test manual, authors reported estimates of reliability and validity including internal consistency for 7-9 year olds (coefficient $\alpha = 0.94 – 0.96$), test-retest reliability between 1-2 weeks ($r = 0.88 – 0.93$), and concurrent validity with the Comprehensive Test of Nonverbal Intelligence, Second Edition ($r = 0.73 – 0.79$). For this study, we required that participants had at least adequate reasoning skills (i.e., standard score ≥ 70) in order to eliminate anyone who may have significant cognitive impairments. All participants met this criterion.

Near visual acuity was assessed with the Waterloo near vision test (Cho & Woo, 2004), where the student read letters of progressively smaller size with each eye at a distance of 40 cm. There were 7 participants (3 good readers, 1 poor reader with small group supports, 3 poor readers in intensive program) who had near visual acuity in one or both eyes that was outside the normal range for children aged 6 to 12 years old (i.e., LogMar > 0.20 or Snellen > 20/32; Dobson, Clifford-Donaldson, Green, Miller, & Harvey, 2009). These participants were excluded from the analyses. Therefore, the final sample size at the first time point was 73 good readers, 41 poor readers with small group supports, and 37 poor readers in the intensive program.

Parent questionnaires.

On a researcher-developed questionnaire, parents provided demographic information, parent education level, supports inside and outside of school, assessment history, and diagnoses of a learning disability or ADHD. Parents also completed the Conners 3 Parent Short Form (Conners, 2008), which provided ratings of the student's level of inattention, hyperactivity, learning problems, executive functioning, aggression, and peer relationships. Estimates of reliability and validity from the test manual included internal consistency for the total sample (coefficient $\alpha = 0.85 – 0.92$), test-retest reliability between 2-4 weeks ($r = 0.73 – 0.97$), and concurrent validity with similar subscales from the Behavior Assessment System for Children, Second Edition Parent Rating Scale and the Child Behavior Checklist ($r = 0.46 – 0.93$).

Literacy tests.

Reading and spelling skills were assessed using subtests from the Kaufman Test of Educational Achievement, Second Edition (KTEA-II; Kaufman & Kaufman, 2004). Students completed alternate forms of the KTEA-II for each visit to minimize practice effects. Authors reported estimates of reliability and validity including split-half reliability coefficients for 7-9 year olds, alternate form reliability between 11-60 days, and concurrent validity with similar subtests from the Wechsler Individual Achievement Test, Second Edition (WIAT-II).

Word recognition.

Word recognition skills were measured on the KTEA-II Letter and Word Recognition subtest. Students were required to read aloud single words of increasing difficulty at their own pace. The test manual reported split-half reliability ($r = 0.95 – 0.98$), alternate form reliability ($r = 0.85$), and concurrent validity with the WIAT-II Word Reading subtest ($r = 0.79$).

Decoding.

Word decoding skills were assessed on the KTEA-II Nonsense Word Decoding subtest. On this subtest, the student read aloud non-words using English pronunciation rules at their own pace. The test manual reported split-half reliability ($r = 0.92 – 0.96$), alternate form reliability ($r = 0.90$), and concurrent validity with the WIAT-II Pseudoword Decoding subtest ($r = 0.86$).

Word recognition fluency.

The KTEA-II Word Recognition Fluency subtest measured word reading efficiency. Students read aloud a list of words as quickly and accurately as they could within one minute. The test manual reported alternate form reliability ($r = 0.89$) and concurrent validity with the WIAT-II Word Reading and Pseudoword Decoding subtests ($r = 0.71 – 0.89$).

Decoding fluency.

The KTEA-II Decoding Fluency subtest measured non-word reading efficiency. On this task, the student was required to read aloud a list of pronounceable non-words within one minute. The test manual reported alternate form reliability ($r = 0.92$) and concurrent validity with WIAT-II Word Reading and Pseudoword Decoding subtests ($r = 0.85 – 0.88$).

Reading comprehension.

Reading comprehension, or the ability to extract meaning from text, was measured with the KTEA-II Reading Comprehension subtest. The student was asked to act out (e.g., “wave goodbye”) or answer questions about a passage they had read. The passage was in front of them when they were answering questions, and questions were designed to be literal or inferential. The test manual reported split-half reliability ($r = 0.94 – 0.97$), alternate form reliability ($r = 0.76$), and concurrent validity with the WIAT-II Reading Comprehension subtest ($r = 0.69$).

Spelling.

Spelling skills were assessed on the KTEA-II Spelling subtest, where the student wrote words that were orally presented to them. The test manual reported split-half reliability (*r* = 0.94 – 0.96), alternate form reliability (*r* = 0.91), and concurrent validity with the WIAT-II Spelling subtest (*r* = 0.91).

Cognitive and language tests.

Students also completed measures of cognitive and language skills, that are related to reading acquisition, on the KTEA-II and Woodcock-Johnson III Tests of Cognitive Abilities (WJ III COG NU; Woodcock, McGrew, & Mather, 2001b, 2007). Estimates of split-half reliability for 7-9 year olds and test-retest reliability with a time interval less than one year are reported from the WJ III COG test manual.

Phonological awareness.

Phonological awareness was assessed with the KTEA-II Phonological Awareness subtest, which requires the student to segment and delete phonemes in orally-presented words, as well as to verbally produce rhyming words. The test manual reported split-half reliability (*r* = 0.82 – 0.90) and alternate form reliability (*r* = 0.58). There was not a similar subtest on the WIAT-II to evaluate concurrent validity.

Receptive language.

The KTEA-II Listening Comprehension subtest was used to measure receptive language skills. On this subtest, students listened to short passages and then answered questions about them and therefore requires both language and memory skills. The test manual reported split-half reliability (*r* = 0.82 – 0.88), alternate form reliability (*r* = 0.56), and concurrent validity with the WIAT-II Listening Comprehension subtest (*r* = 0.52).

Memory span.

Memory span was assessed on the WJ III COG Memory for Words test. On this test, students heard a list of words and then were asked to verbally repeat the words verbatim. The test manual reported split-half reliability (*r* = 0.72 – 0.79) and test-retest reliability (*r* = 0.77).

Working memory.

On the WJ III COG Numbers Reversed test, the student heard a sequence of numbers and was asked to verbally repeat the numbers in reverse order. This test measures working memory capacity. The test manual reported split-half reliability (*r* = 0.84 – 0.86), but no estimate of test-retest reliability.

Rapid naming.

The KTEA-II Naming Facility subtest provides an assessment of rapid naming. The student was asked to name pictures of objects, colours, and letters as quickly as possible. The test manual reported split-half reliability (*r* = 0.87 – 0.90) and alternate form reliability (*r* = 0.76). There was not a similar subtest on the WIAT-II to evaluate concurrent validity.

Processing speed.

Processing speed for numbers was measured on the WJ III COG Visual Matching test. Participants were asked to identify matching numbers under a time limit (i.e., in a row of 6 numbers, the student circled the 2 numbers that were the same). The test manual reported split-half reliability ($r = 0.87 – 0.89$) and test-retest reliability ($r = 0.86$).

Results

The following analyses examined differences between groups on demographic variables, the effects of intervention, and predictors of group membership. The results are reported with a false discovery rate (FDR) to correct for multiple comparisons (Benjamini, Krieger, & Yekutieli, 2006); a FDR value of $q < .05$ was considered significant.

Demographics

Table 2.1 shows demographic information for the 3 groups. Group differences were analyzed using analysis of variance (ANOVA) for continuous variables and chi-square tests for categorical variables. Students were approximately 8.5 years old at the beginning of this study, with no differences between groups in gender or age (gender: $\chi^2(2) = 3.85$, $q = .324$; pre-test age: $F(2, 148) < 1$, $q = .864$; post-test age: $F(2, 140) < 1$, $q = .864$; delayed test age: $F(2, 103) < 1$, $q = .864$). The majority of students spoke English as their first language and similar distributions for language status were shown across groups, $\chi^2(2) = 3.25$, $q = .324$. Most parents had a college or university degree and parent education level was similar across groups (mother: $\chi^2(8) = 16.40$, $q = .136$, father: $\chi^2(8) = 10.68$, $q = .324$). Mean non-verbal intelligence was within the standardized average range for all groups, however, scores were higher for good readers compared to the poor readers, $F(2, 148) = 6.51$, $q = .014$. Furthermore, there were group differences on the Conners 3 Parent rating scale, $F(2, 141) = 24.75$, $q < .001$. The 2 poor reader groups had higher ratings of inattention, hyperactivity, learning problems, executive functioning, aggression, and peer relations when compared to the good reader group (all $q < .05$). Students in the intensive program also had higher ratings of learning problems than students who received small group supports ($q = .004$). There were no differences in age, gender, language status, parental education, or nonverbal IQ between those who were retained in the study versus those who dropped out (all $q > .05$).

Table 2.1. Demographic information for the 3 groups

Variable	Good reader	Poor reader with small group supports	Poor reader with intensive intervention
Sample size (N, % male)			
Pre-test	73 (49%)	41 (68%)	37 (57%)
Post-test	68 (49%)	40 (68%)	35 (54%)
Delayed test	65 (51%)	18 (78%)	23 (52%)
Age (M years, SD)			
Pre-test	8.54 (0.35)	8.51 (0.35)	8.53 (0.37)
Post-test	8.78 (0.33)	8.80 (0.33)	8.81 (0.36)
Delayed test	9.72 (0.35)	9.67 (0.38)	9.71 (0.34)
Language			
English as primary language (N, %)	58 (84%)	35 (92%)	35 (95%)
Mother's education (N, %)			
Some high school	0 (0%)	2 (5%)	1 (3%)
High school diploma	1 (3%)	4 (11%)	2 (5%)
At least one year college	4 (6%)	5 (13%)	8 (22%)
College or university degree	34 (49%)	13 (34%)	17 (46%)
Graduate or professional degree	30 (43%)	14 (37%)	9 (24%)
Father's education (N, %)			
Some high school	0 (0%)	3 (8%)	2 (6%)
High school diploma	3 (5%)	5 (13%)	6 (17%)
At least one year college	15 (23%)	7 (18%)	4 (11%)
College or university degree	25 (39%)	12 (32%)	13 (36%)
Graduate or professional degree	21 (33%)	11 (29%)	11 (31%)
Nonverbal IQ (M, SD) [a, *]	107.21 (10.21)	102.17 (9.44)	100.78 (9.49)
Parent ratings of ADHD symptoms (M, SD) [b]			
Inattention *	52.01 (8.77)	63.40 (13.81)	62.65 (13.33)
Hyperactivity *	53.87 (11.30)	60.84 (14.35)	61.59 (15.87)
Learning problems *	48.43 (8.75)	65.39 (10.47)	71.19 (7.75)
Executive functioning *	52.28 (8.38)	60.92 (13.08)	61.27 (13.61)
Aggression *	49.39 (10.03)	54.74 (13.14)	55.00 (14.97)
Peer relations *	52.00 (11.36)	59.84 (17.27)	58.95 (17.99)

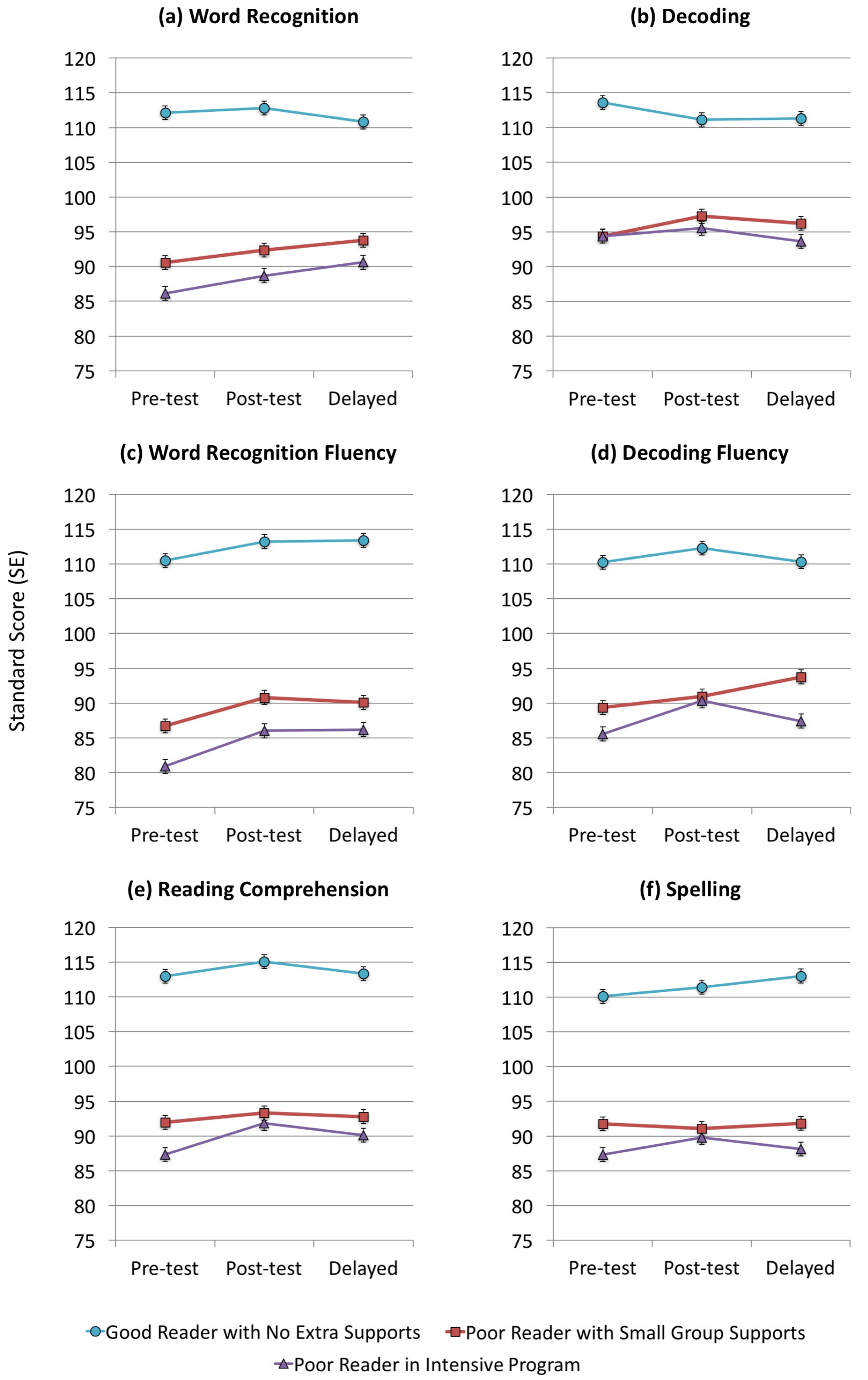
Note: [a] Standard score mean = 100, SD = 15; [b] T-score mean = 50, SD = 10; *Significant difference between groups, FDR corrected $q < .05$. Sample sizes and percentages vary as some questions were left blank on the parent questionnaire.

Intervention Outcomes

Linear mixed modelling was used to account for the hierarchical nature of the data and to evaluate outcomes of the reading interventions. These models included longitudinal measurements on a specific test for each student and school. Classroom was not included as another level of data, as most students changed classrooms between the third and fourth grade. Unconditional models included student as a random effect in the first model and student and school as random effects in the second model. The results showed large variability between students (*M* intraclass correlation = 0.74, *SD* = 0.14), with minimal variance attributable to school differences (*M* intraclass correlation = 0.07, *SD* = 0.08). Likelihood ratio tests between the 2 models indicated that schools did not contribute a significant amount of variance, $\chi^2(1) < 1$, $p > .05$; therefore, school was not included as a random effect in the subsequent analyses. The conditional models included random effects of students and fixed effects factors of the 3 groups and 3 time points. Standard scores were used as the dependent variable and a covariate of non-verbal intelligence was included in all analyses, as there were significant differences between groups on this measure. The analyses were also conducted without students who had a reading comprehension deficit only ($n = 5$ students with small group supports) and the results described below were the same. Therefore, the results are reported with all participants described in Table 2.1.

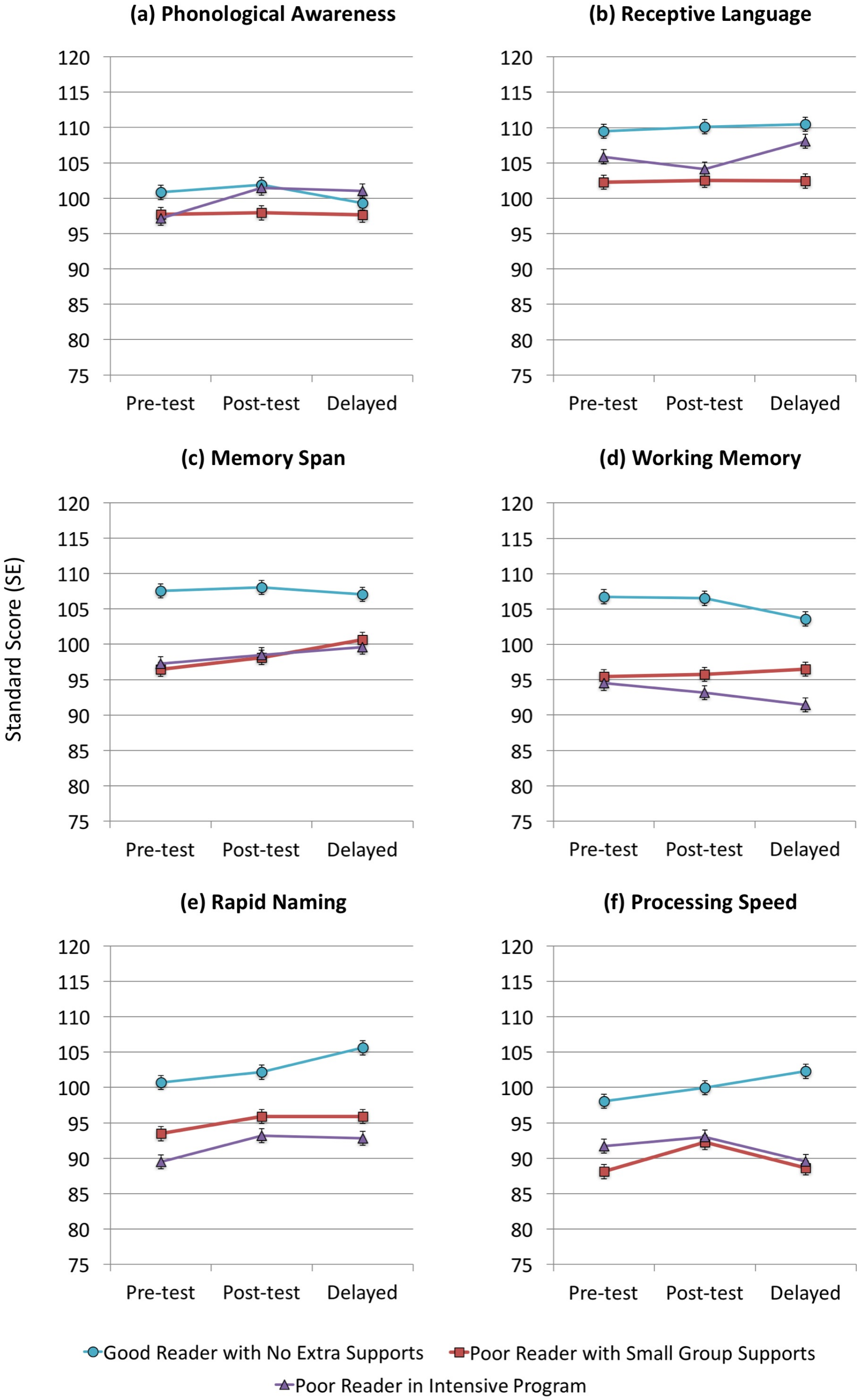
Figures 2.1 and 2.2 illustrate the results from the linear mixed modelling analyses. A significant main effect of group was shown on all measures ($q < .001$), except for phonological awareness, $F(2, 157.51) = 1.23$, $q = .207$. Pairwise comparisons showed that the intensive group had lower scores than other groups on word recognition, word recognition fluency, and decoding fluency ($q < .05$ on all tests), confirming that these students had the lowest word reading skills. The 2 poor reader groups, however, showed similar performance on tests of decoding, reading comprehension, spelling, receptive language, memory span, working memory, rapid naming, and processing speed ($q > .05$ on all tests). Furthermore, the good reader group had the highest scores, on average, for all measures ($q < .005$ on all tests) other than phonological awareness. Across groups, there were increased scores over time in word recognition, $F(2, 217.02) = 4.86$, $q = .009$, reading comprehension, $F(2, 209.35) = 4.97$, $q = .009$, word recognition fluency, $F(2, 196.79) = 30.99$, $q < .001$, decoding fluency, $F(2, 204.60) = 9.39$, $q < .001$, rapid naming, $F(2, 201.51) = 8.72$, $q < .001$, and processing speed, $F(2, 209.04) = 3.08$, $q = .043$.

Figure 2.1. Results for the literacy tests across time and by group



Note: Standardized mean = 100, SD = 15; Values represent estimated marginal means from the linear mixed models and bars denote 1 standard error (SE).

Figure 2.2. Results for the language and cognitive tests across time and by group



Note: Standardized mean = 100, SD = 15; Values represent estimated marginal means from the linear mixed models and bars denote 1 standard error (SE).

Significant group by time interactions, if present, would indicate that there was differential growth across groups, which may be related to the effects of intervention. There were significant interactions on word recognition, *F*(4, 217.36) = 3.50, *q* = .009, decoding, *F*(4, 214.42) = 3.48, *q* = .009, decoding fluency, *F*(4, 202.63) = 2.61, *q* = .035, spelling, *F*(4, 187.69) = 2.36, *q* = .047, and processing speed, *F*(4, 211.85) = 2.83, *q* = .026, tasks. Follow-up pairwise comparisons indicated that the intensive group improved in word recognition (*q* = .029) and decoding fluency (*q* = .010) after intervention, and this improvement was maintained one year later (*q* > .05 between post-test and delayed test). There were also improvements in spelling (*p* = .024) for the intensive program, and in word recognition (*p* = .048), decoding (*p* = .019), decoding fluency (*p* = .013), and processing speed (*p* = .020) for students who received small group supports; however, these findings did not survive FDR correction. Furthermore, the good reader group had small but significant improvements in spelling (*q* = .010) and processing speed (*q* = .010). Changes over time for the good readers were unexpected but likely reflects an effect of large sample size for this group. Therefore, effects sizes were calculated with the G*Power 3.1 program (Faul, Erdfelder, Lang, & Buchner, 2007), using the mean and standard deviation for pre-test and delayed test time points as well as the correlation between these 2 time points to determine the pooled standard deviation. Effect sizes were calculated for each group on the 6 literacy measures (see Table 2.2). These results showed that effect sizes for the good reader group were either minimal or small, whereas effect sizes for the small group and intensive programs were in the medium to large range (Cohen, 1992).

Table 2.2. Within-group effect sizes on the literacy tests

Measure	Good reader	Poor reader with small group supports	Poor reader with intensive intervention
Word Recognition	0.18	0.65	0.80
Decoding	0.25	0.30	0.04
Word Recognition Fluency	0.38	0.84	1.19
Decoding Fluency	0.02	0.61	0.62
Reading Comprehension	0.06	0.30	0.61
Spelling	0.44	0.13	0.33

Note: Cohen's *d* effect size was based on growth from pre-test to delayed test within each group. Light grey denotes medium effects (0.50 – 0.79) and dark grey denotes large effects (0.80 and greater).

Although analyses with all groups together indicated changes that are related to intervention beyond that of typical development, meaningful differences between the poor reader groups may be missed due to the variability and inclusion of the good reader group. Therefore, *a priori* contrasts between the 2 poor reader groups at post-test and delayed test were conducted. Group differences were examined for the word recognition, word recognition fluency, and decoding fluency tasks only, as the analyses above indicated that the intensive group had the lowest scores on these tasks. These contrasts showed that students in the intensive program had similar word recognition skills to small group programs immediately following intervention (q = .068) and one year later (q = .135). They also had similar decoding fluency after intervention (q = .497), but then the intensive group had worse performance one year (q = .039). Finally, although word recognition fluency scores for the intensive group were lower than the small group program after intervention (q = .039), they had similar performance one year later (q = .097).

I also examined the percentage of students who obtained scores 1 *SD* or below the standardized mean on reading tests (i.e., standard score < 85 or < 16th percentile) to determine students who were persistently poor readers. None of the students in the good reader group had scores below 1 *SD* at any time point. For students in the intensive program, 62% scored 1 *SD* or below on at least 2 reading tests prior to intervention (n = 23 of 37), which decreased to 40% following intervention (n = 14 of 35), and to 35% one year later (n = 8 of 23). For students in small group programs, 32% met this criterion on at least 2 reading tests at the first time point (n = 13 of 41), and this decreased to 23% after 3 months of instruction (n = 9 of 40), and to 17% one year later (n = 3 of 18).

Predictors of Reading Outcome

Finally, I examined whether pre-test measures could predict group membership of persistently low reading achievement at either post-test or delayed test using the criteria described above (i.e., < 85 standard score on at least 2 reading tests). Across all participants, this included a total 30 students who were persistently poor readers (n = 11 small group programs, n = 19 intensive program) and 48 students who were not (n = 30 small group programs, n = 18 intensive program). Discriminant analyses were conducted in 3 separate analyses due to small sample size; this included predictors of pre-test literacy measures, pre-test cognitive and language measures, and pre-test ratings of ADHD symptoms. The results indicated that pre-test literacy measures (canonical correlation = 0.72, Wilk's λ = 0.48, $\chi^2(6)$ = 53.60, q < .001) as well as cognitive and language measures (canonical correlation = 0.48, Wilk's λ = 0.77, $\chi^2(6)$ = 17.01, q = .003) were predictive of group membership, but ratings of ADHD symptoms were not (canonical correlation = 0.12, Wilk's λ = 0.99, $\chi^2(6)$ = 0.95, q = .346). Furthermore, 84.6% of students were correctly classified into groups using the pre-test literacy measures alone, compared to 70.7% using the cognitive and language measures and 61.3% using the ADHD ratings. An inspection of means and within-groups correlations indicated that the students with persistently low reading achievement had lower word recognition, decoding, reading comprehension, fluency, spelling, phonological awareness, and rapid naming compared to students who did not have persistent reading deficits.

Discussion

Improved Reading Skills After Intensive Intervention

Prior to intervention, third grade students who were enrolled in an individualized and intensive reading program had the lowest reading scores in the sample, specifically on measures of word recognition, word recognition fluency, and decoding fluency. After intervention, these students showed increased word recognition and decoding fluency when compared to the other groups. There were improvements in word recognition fluency and reading comprehension as well but these were similar across all groups, which implies that increases could have been due to normal development. However, the medium-to-large effect sizes on these tasks indicate that power to detect changes was limited by variability or sample size. These results are similar to prior studies of intensive and individualized programs, which showed improved word identification, decoding, fluency, and comprehension following intervention (Denton et al., 2013; O'Connor et al., 2005). The current study replicates and extends these findings to an all-day instructional program that was implemented by teachers. Therefore, this study demonstrates the importance of school-based reading interventions and provides further support for the use of intensive and individualized programming for students with the poorest reading skills.

One purpose of the intensive program was to accelerate reading growth to obtain similar performance to that of poor readers who were receiving small group supports. This effect was shown on word recognition and decoding fluency tasks immediately after the program ended in the third grade. After students returned to their home school and continued to receive small group supports, they had further gains in word recognition fluency, so that their scores were similar to other poor readers in the fourth grade. However, students in the intensive program showed declined performance in decoding fluency between the third and fourth grade. In summary, these results indicate that one year after intensive intervention, there is comparable performance between poor reader groups in word recognition and word recognition fluency. This result is similar to previous studies that showed gains in reading scores and then stability over time (Blachman et al., 2014), and provides evidence for the stability of word recognition skills after school-based reading programs.

The number of students with persistent reading deficits also decreased over time, which was defined as below average performance on at least 2 reading tests (i.e., > 1 *SD* below the standardized mean). In the intensive program, 62% of students met this criterion before intervention which decreased to 35% one year later. Similarly, 32% of students in small group programs had below average performance on at least 2 reading tests at the beginning of the study, and this decreased to 17% in the fourth grade. This result implies a decline in significant reading impairments, however, a large gap remains between these students and their typical reader peers, which is a replication of previous research (Denton et al., 2006; 2013; Gilbert et al., 2013; O'Connor et al., 2005; Vaughn et al., 2009). Therefore, the current results provide further evidence that ongoing reading supports at school or home are needed for the students who have the poorest reading skills.

Predictors of Persistent Reading Deficits

I also examined whether baseline performance on measures of literacy, language, and cognition as well as ratings of ADHD symptoms were predictive of persistently low reading achievement after intervention. Pre-test measures of reading, spelling, phonological awareness, and rapid naming correctly classified students into 2 groups of persistent low achievers or not, whereas ratings of ADHD symptoms were not significant predictors of outcome. Furthermore, the reading and spelling measures had a better classification rate (84.6%) compared to language and cognitive measures (70.7%). This result is consistent with previous studies and replicates the idea that pre-test reading measures differentiate between “low” and “adequate” responders to intervention. For example, a recent meta-analysis suggested that baseline reading scores are a better predictor of outcome than cognitive measures (Stuebing et al., 2015). The results are consistent with this meta-analysis, and suggest that pre-test literacy skills are predictive of persistent reading deficits. Therefore, baseline literacy scores may be used as a screening tool for determining the students who will have ongoing reading deficits and require intensive and individualized supports.

Skills Not Improved After Intensive Intervention

Following intensive intervention, students did not show a statistically significant improvement on measures of spelling, decoding, language, short-term memory, or processing speed. Changes in listening comprehension, short-term memory, or processing speed were not hypothesized, however, areas of spelling and decoding were a direct focus of the intensive programs and thus were expected to improve. An inspection of raw scores indicated that most students increased in the number of words spelled correctly and words properly decoded following intervention. However, this improvement was similar to what would be expected by typical development and instruction, and suggests that additional practice with orthographic (i.e., spelling) and phonological (i.e., decoding, rhyming) skills may be required for this program. For example, some programs described in the General Introduction specifically focus on orthography (e.g., Lindamood-Bell Seeing Stars) or phonology (e.g., Phono-Graphix), and these may be beneficial for some students.

Although lower phonological awareness at the first time point was predictive of persistent reading deficits, there were non-significant group differences across time points on this task, as would be expected based on previous literature (e.g., Snowling, 1981). This result may be reflective of the early intervention programs that are in place in the school district and suggests that good and poor reader groups had similar rhyming and segmentation skills. However, despite adequate reliability from the test manual, there was a slight ceiling effect on this subtest that may have limited the variability in scores. Therefore, it is possible that group differences would have been shown on a different phonological awareness test but that cannot be ascertained in this study.

Considerations and Conclusions

There was a relatively small sample size at the third time point due to attrition and students being unavailable for testing. The current results may be biased towards those who were retained in the study, however, no differences in age, language, parental education, or nonverbal IQ were found between students who remained in the study compared to those who dropped out. Therefore, intervention effects cannot be attributable to these demographic variables.

Overall this study illustrated that students in the intensive program boosted their reading scores to levels comparable with students receiving small group supports. This result provides empirical support for an all-day intensive and individualized treatment program. However, ongoing interventions are needed and perhaps adaptations to the program may be warranted to illustrate even further gains, such as smaller class size, increased time in the program, or more tutors in the classroom (e.g., Fuch et al., 2014; Vaughn & Wanzek, 2014). This study also showed that students who received small group supports showed minimal gains over time, on average, and therefore a different model of service delivery may be needed. For example, high-quality instruction that starts in Kindergarten with an accelerated entry into intensive programs may be beneficial for students with the lowest reading scores (Compton et al., 2012). Finally, this study showed that baseline literacy scores were the best predictors of persistent reading deficits, which is consistent with results showing that low reading achievement is associated with reading outcome rather than results from cognitive measures (Stuebing et al., 2015). Therefore, baseline reading scores may be used as a screening tool to determine students who will require the most intensive supports.

In conclusion, this study replicates and extends previous studies of intensive reading programs. The results showed increased reading skills following an individualized, intensive program, which was implemented for most of the school day for 3 months. Gains in reading were considered to be in the medium to large effect size range, however, continued supports would be needed to achieve the reading level of good readers in the classroom. Earlier implementation and flexible RTI models may be one method to ameliorate persistent reading deficits.

CHAPTER 3. EFFECT OF INTERVENTION AND TASK DIFFICULTY ON ORTHOGRAPHIC AND PHONOLOGICAL READING SYSTEMS IN THE BRAIN

Introduction

The General Introduction described the cognitive and language processes involved with reading development. Specifically, reading models indicate that a combination of orthographic, phonological, and semantic knowledge is needed for word recognition (Coltheart et al., 2001; Harm & Seidenberg, 2005). In the dual route model, it has been suggested that indirect and direct routes are associated with learning to read (Coltheart et al., 2001), and children with reading impairments have deficits in orthographic and phonological reading skills (Castles & Coltheart, 1993; Snowling, 1981). These types of skills are typically measured by reading irregular words (e.g., 'yacht') to emphasize the orthographic route or by reading pronounceable non-words (e.g., 'brish') to emphasize the phonological route. However, other tasks such as spelling and rhyming elicit similar reading processes. The General Introduction also outlined the positive effects of reading intervention, including improved reading skills after intensive treatment (Denton, 2012; Vaughn & Wanzek, 2014). Following the dual route model, there have been commensurate gains in orthographic (Christodoulou et al., 2017) and phonological (Pokorni et al., 2004) awareness with programs that specifically target these skills.

Neuroimaging studies have illustrated that poor readers have altered patterns of activation compared to good readers (Shaywitz et al., 2002). Following treatment, there are associated changes in brain activity, such that poor readers closely resembled their typical reader peers immediately after and up to one year later after the intervention ended (Meyler et al., 2008; Shaywitz et al., 2004). The majority of these studies have included phonological measures of brain activity, including matching letters to phonemes (Davis et al., 2011; Farris et al., 2011; Odegard et al., 2008; Shaywitz et al., 2004), matching phonemes to morphemes (Aylward et al., 2003), deleting phonemes in words (Eden et al., 2004), rhyming pairs of letters (Temple et al., 2003), rhyming pairs of words (Hoeft et al., 2007; 2011; Richards et al., 2002), or rhyming pictures of objects (Farris et al., 2016). Others have used a lexical decision task, which is thought to activate the orthographic route because it emphasizes word meaning (Horowitz-Kraus et al., 2014; Richards et al., 2002). Few have specifically examined orthographic awareness, such as judging spelling accuracy (Richards et al., 2006). Evaluating both orthographic and phonological processes in the brain are important, as students may improve in one or both skills based on the skills that are emphasized in a treatment program.

Other functional MRI studies have shown that there is an effect of difficulty on brain activation patterns. For example, group differences between average and poor readers may only be shown on difficult reading tasks (Cao et al., 2006) or on both easy and difficult reading tasks (Pugh et al., 2008). In the latter study, the important aspect was the direction of activation. In comparison to typically-developing readers, poor readers had less activation on difficult tasks and increased activation on easy tasks, which indicates that brain activity was modulated by task difficulty. Finally, a recent meta-analysis indicated that the effect of difficulty in reading tasks was localized to phonological and attentional regions, including left inferior frontal gyrus, left superior parietal lobe, medial cingulate cortex, and the pons/cerebellum (Cattinelli, Borghese, Gallucci, & Paulesu, 2013). These studies illustrate that task difficulty is important in studies of reading, however, none have examined how reading intervention is modulated by task difficulty. This type of evaluation is important as intervention effects may only be shown on relatively easy or difficult tasks.

Therefore, the purpose of the current study was to evaluate the effect of reading intervention on orthographic and phonological systems in the brain, and whether these results were modulated by task difficulty. Children in this study were part of a larger intervention study where students received either small group or intensive supports in their school district (Chapter 2). There was also a group of good readers who did not receive additional reading supports.

Methods

Ethics approval for this study was obtained from Clinical Research Ethics Boards at The University of British Columbia and the Child and Family Research Institute. Informed consent from parents and assent from children was obtained prior to participation.

Participants

There were 37 students between 8 and 9 years old who were recruited for the MRI study. Two students declined to complete the MRI scan after practicing in a mock scanner, which meant there were 22 good readers (11 female), 9 poor readers receiving small group supports (2 female), and 4 poor readers receiving intensive supports (3 female) at the first time point ("pre-test"). Due to small sample size, the 13 poor readers were combined into one group for analyses. Five good readers and one poor reader did not return for a second scan; thus, there were 17 good readers and 12 poor readers at "post-test".

As part of the larger intervention study (Chapter 2), good readers did not receive additional reading supports in their school district while poor readers received either small group or intensive programming. All participants spoke English as their primary language, were predominately right-handed, had at least average non-verbal reasoning ability on the TONI-4 (i.e., standard score ≥ 70), and had near visual acuity that was within the normal or corrected-to-normal range (i.e., LogMar < 0.20 or Snellen < 20/32). None of the participants were diagnosed with ADHD. Most parents completed a university or college degree, according to parent report. Children were ineligible for this study if they were currently in a French immersion instruction program or if they had known neurological impairments, developmental disorders, or uncorrected sensory difficulties. In comparison to the Chapter 2 study, MRI data were not collected at the third time point (“delayed test”).

Screening, Literacy, and Cognitive Measures

All participants completed the following measures at the first time point. The Test of Nonverbal Intelligence, 4th Edition (TONI-4; Brown et al., 2010) was used to estimate overall intellectual functioning, standardized to the child’s age. Handedness was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971), which included 10 questions about which hand the child uses for certain activities (e.g., writing, using scissors). This measure provided a handedness quotient (HQ), where positive values indicated predominant right-handedness and negative values indicated predominant left-handedness. The Waterloo near vision test (Cho & Woo, 2004) was used to assess near visual acuity. Children were asked to read letters of progressively smaller size with each eye. Finally, parents completed the Conners 3 Parent Short Form (Conners, 2008) to provide ratings of a child’s inattention, hyperactivity, learning problems, executive functioning, aggression, and peer relationships.

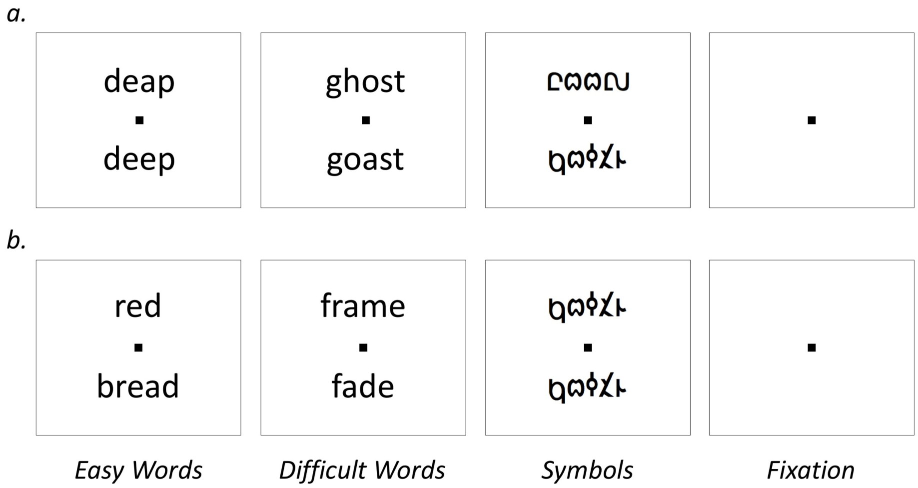
At both time points, students completed literacy, language, and cognitive subtests from the Kaufman Test of Educational Achievement, Second Edition (KTEA-II; Kaufman & Kaufman, 2004) and Woodcock-Johnson III Tests of Cognitive Abilities (WJ III COG NU; Woodcock, McGrew, & Mather, 2001b, 2007). This included measures of word recognition, decoding, fluency, reading comprehension, spelling, phonological awareness, listening comprehension, memory span, working memory, rapid naming, and processing speed. All subtests demonstrated adequate internal consistency, alternate form or test-retest reliability, and concurrent validity, as described in the test manuals.

MRI Measures

Each MRI session started with a 30-minute practice in a mock scanner at the Child and Family Research Imaging Facility. This allowed the children to practice the reading tasks and to get comfortable with the MRI environment. A 3.0 Tesla GE Discovery MR 750 scanner with a 32-channel head coil was used to collect images. A high-resolution T1-weighted 3D anatomic scan was obtained first and was used to align images from the functional scans. Images were acquired in 170 sagittal slices, with the following parameters: field-of-view (FOV): 256 x 256 mm; matrix: 256 x 256; TR: 7.6 ms; TE: 3.5 ms; flip angle: 8 deg; slice thickness: 1 mm; voxel size: 1 mm^3; scan length: 5 minutes. T2*-weighted scans using an echo-planar imaging (EPI) sequence were acquired for whole-brain functional data (TR: 2000 ms; TE: 30 ms). All images were acquired in 39 interleaved axial slices, starting with the most inferior slice (FOV: 288 x 288 mm; matrix: 96 x 96; flip angle: 90 deg; phase encode direction: A/P; 4 dummy scans discarded at acquisition; slice thickness: 3 mm; inter-slice gap: 1 mm; reconstructed voxel size: 3 mm^3). The reading tasks acquired 156 images for each task (scan length: 5.2 minutes).

In this study, an orthographic spelling task and a phonological rhyming task were designed with the same timing parameters and were presented in counterbalanced order in the scanner. Each task had four conditions, which consisted of easy words, difficult words, false font symbols (Turkeltaub et al., 2003) and a fixation cross baseline (see Figure 3.1). On each trial, the participant indicated which word was spelled correctly ('top' or 'bottom') on a response pad for the spelling task, and whether the 2 words rhymed ('yes' or 'no') for the rhyming task. During the symbols condition, the participant indicated if the 2 symbol strings looked the same or not ('yes' or 'no'). For the fixation condition, the participant was asked to press the 'no' button several times, which eliminates potential brain activation differences between blocks due to button pressing. Each block lasted 26 seconds and each word or symbol string was presented for 4 seconds with a 1200 ms interstimulus interval. There were 5 trials of words or symbols per block. The 4 blocks were presented in random order within a cycle and each cycle was repeated 3 times, for a total of 12 presentation blocks. Reaction time and accuracy were recorded for each trial. Due to equipment failure, reaction time and accuracy were not obtained for 2 participants.

Figure 3.1. fMRI spelling and rhyming tasks



Note: Participants were asked to indicate (a) which word was spelled correctly on the spelling task, or (b) whether the two words rhymed on the rhyming task, on blocks of easy and difficult words. Participants also indicated whether symbols looked the same or not, and pressed a button randomly during the blank fixation block. In example (a), the correct answers are 'bottom' for easy words, 'top' for difficult words, and 'no' for symbols. In example (b), the correct answers are 'yes' for easy words, 'no' for difficult words, and 'yes' for symbols.

Both tasks included monosyllabic words of 4 to 6 letters, and blocks were designed to be relatively easy or difficult based on printed word frequency (Balota et al., 2007). Across the spelling and rhyming tasks, words were matched for length and frequency as well as estimates of word imageability and age of acquisition (Cortese & Fugett, 2004; Cortese & Khanna, 2008) to decrease the influence of these factors on brain activation. The spelling words pairs were chosen to sound the same to emphasize orthography (e.g., “gole” and “goal” sound the same, but only one is spelled correctly) and the rhyming word pairs had dissimilar endings to emphasize phonology and not orthography (e.g., “theme” and “dream” rhyme, but have different endings). Proper names and brand names were not included to reduce familiarity with these words (e.g., “Jane”, “Coke”). Word lists are included in Appendix A. Good and poor readers read the same word lists to minimize differences in brain activity based on differences in words. Furthermore, pilot testing was conducted with students not involved in the current study to ensure words were readable by children with reading impairments. Participants included a group of good readers (*n* = 12) and poor readers (*n* = 8) between 7 and 10 years old who were recruited from the community and a private school for learning disabilities. Assessments confirmed that they met the same inclusion/exclusion criteria as above. In a computerized version of the spelling and rhyming tasks, participant’s accuracy and reaction time were recorded. Word pairs were chosen for the fMRI task if participants read the words with at least 60% accuracy under 4 seconds (i.e., the presentation time during the fMRI tasks).

An Epson Powerlite Home Cinema 5010 projector was used to present all stimuli, which were projected onto a rear projection screen at the back of the scanner bore and viewed through a rear-facing mirror that was attached to the head coil. Words were presented in Arial font and were subtended 3.34 deg at a viewing distance of 60 cm. A MacBook Pro laptop with Matlab (The MathWorks, Inc.) and Psychtoolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) software was used to present the tasks and to collect behavioural responses.

MRI Analyses

FSL’s fMRI Expert Analysis Tool (FEAT version 6.00) was used for pre-processing and analyses (Jenkinson & Smith, 2001; Jenkinson, Bannister, Brady, & Smith, 2002; Smith, 2002; Smith et al., 2004). For each participant, pre-processing of the data was conducted to reduce the effects of motion and spatial-temporal factors, which is commonly done in functional neuroimaging. The pre-processing steps included: 3D motion correction using a six parameter rigid-body tri-linear interpolation aligned to the middle image in the time course (MCFLIRT), slice-timing correction using Fourier time series phase-shifting, spatial smoothing with a Gaussian kernel of 6 mm FWHM, grand-mean intensity normalization, high-pass temporal filtering (sigma = 50 s), and removal of non-brain tissues (BET). Translation (mm) and rotations (deg) were estimated for each scan and data were excluded if framewise displacement exceeded 1 mm. Data for one scan on the spelling task (one poor reader at pre-test) and for 4 scans on the rhyming task (2 poor readers at pre-test, 2 average readers at post-test) were excluded due to excessive motion. Using FLIRT, linear registration was conducted to align functional images to each participant's T1-weighted image and the NIH pediatric template for 7-11 year olds (Fonov et al., 2011). These 2 steps of linear registration are required for whole-brain group analyses.

In addition to whole-brain analyses, *a priori* regions of interest (ROI) were established on each participant's T1 image using Freesurfer version 5.3.0 (Dale, Fischl, & Sereno, 1999) and the Desikan-Killiany atlas (Desikan et al., 2006). Eight left hemisphere ROIs that are known to be involved with language processing were included in the analyses (Dick & Tremblay, 2012; Friederici & Gierhan, 2013): inferior frontal gyrus (subdivided into 3 regions of pars orbitalis, pars triangularis, and pars opercularis); insular cortex; superior temporal gyrus; supramarginal gyrus; angular gyrus; and fusiform gyrus. These regions have also been implicated in aspects of orthographic and phonological processing (e.g., Borowsky et al., 2006; Taylor et al., 2013). The same ROIs in the right hemisphere were also included in the analysis. Linear registration using FSL's FLIRT was conducted to transform the anatomical ROIs into functional space for each participant.

Analyses and Results

Demographic Information and Change in Literacy Skills

Table 3.1 includes descriptive information for each group. Group differences on the demographic variables were examined using *t*-tests for continuous variables and chi-square tests for categorical variables. Degrees of freedom were adjusted for unequal variances when necessary. There were no differences between groups in age at pre-test, $t(33) = -0.25$, $p = .806$, age at post-test, $t(27) = -0.03$, $p = .978$, nonverbal IQ, $t(33) = 1.69$, $p = .100$, or handedness, $t(33) = -0.39$, $p = .699$. Similar distributions between groups were also shown for participant's gender, $\chi^2(1) = 0.44$, $p = .508$, mother's education, $\chi^2(4) = 8.11$, $p = .088$, and father's education, $\chi^2(4) = 8.01$, $p = .091$. On the Conners 3 rating scale, students in the poor reader group had significantly higher levels of inattention, $t(17.71) = -2.59$, $p = .019$, and learning problems, $t(33) = -4.78$, $p < .001$, than good readers. No difference between groups were shown on ratings of hyperactivity, executive functioning, aggression, or peer relations (all $p > .10$).

Table 3.1. Descriptive information for MRI participants

Variable	Good reader	Poor reader
Age (M years, SD)		
Pre-test	8.54 (0.41)	8.57 (0.38)
Post-test	8.87 (0.42)	8.87 (0.39)
Screening measures (M, SD)		
Nonverbal IQ (Standard score) [a]	108.36 (8.18)	104.08 (5.19)
Handedness (HQ) [b]	72.59 (22.73)	76.03 (28.98)
Conners 3 Parent Rating Scale (M, SD) [c]		
Inattention *	50.91 (2.36)	62.15 (3.07)
Hyperactivity	55.86 (2.71)	62.54 (3.53)
Learning Problems *	48.23 (2.27)	66.00 (2.95)
Executive Functioning	53.18 (2.01)	56.23 (2.61)
Aggression	52.68 (2.70)	53.69 (3.51)
Peer Relations	52.23 (3.13)	57.00 (4.07)

Note: [a] Standardized mean = 100, SD = 15; [b] Positive scores indicate predominant right-handedness; [c] Standardized mean = 50, SD = 10; *Groups significantly different, $p < .05$;

In the current study, linear mixed models were conducted to determine differences between the 2 groups and 2 time points. These models included repeated effects of time as well as fixed effects of group and time along with an interaction between these factors. The results indicated that good readers had significantly higher scores than poor readers on all literacy tests (all $p < .001$; see Table 3.2). Unlike the large sample in Chapter 2, no significant group by time interactions were shown, which likely reflects small sample size. Therefore, paired sample *t*-tests were subsequently conducted for each group separately to increase power and to examine changes from pre-test to post-test. These results showed that both groups improved in word recognition fluency over time (good readers: $t(16) = -3.62$, $p = .002$; poor readers: $t(11) = -3.34$, $p = .007$), while the poor readers also had increased decoding skills, $t(11) = -2.56$, $p = .026$. There were no other significant changes over time (all $p > .05$).

Table 3.2. Change in literacy skills over time

Variable (M, SD) [a]	Good reader		Poor reader	
	Pre-test	Post-test	Pre-test	Post-test
Word Recognition	112.54 (6.57)	115.24 (10.52)	89.33 (7.24)	90.67 (6.40)
Decoding [b]	112.12 (9.29)	113.53 (7.63)	92.83 (5.18)	96.00 (7.11)
Word Recognition Fluency [b, c]	111.82 (11.61)	115.94 (11.27)	84.50 (9.15)	88.75 (7.30)
Decoding Fluency	108.76 (9.58)	112.12 (9.03)	85.42 (5.11)	86.92 (7.31)
Reading Comprehension	115.94 (10.18)	116.24 (9.88)	89.17 (7.91)	93.42 (8.37)
Spelling	109.71 (10.88)	110.82 (13.28)	90.17 (5.46)	91.08 (6.27)

Note: [a] Standardized mean = 100, SD = 15; [b] Increased scores for poor readers, $p < .05$, [c] Increased scores for good readers, $p < .05$.

Accuracy and Reaction Time on fMRI Tasks

Linear mixed models were conducted to examine differences in accuracy and reaction time between groups, conditions, and time point for each fMRI task. These models included repeated effects of time and condition as well as fixed effects of group, time, and condition along with interactions between these factors. As expected, good readers had significantly higher accuracy and faster reaction time than poor readers on spelling (accuracy: $F(1, 141.29) = 94.03$, $p < .001$; reaction time: $F(1, 150.29) = 17.21$, $p < .001$) and rhyming (accuracy: $F(1, 160.26) = 129.53$, $p < .001$; reaction time: $F(1, 164.95) = 28.54$, $p < .001$) tasks (see Table 3.3). Group differences on the symbols condition were not expected, but may have occurred because some poor readers were slower or had difficulty switching between task instructions for each condition.

Table 3.3. Accuracy on fMRI tasks

Task and Condition	Good reader		Poor reader	
	Pre-test	Post-test	Pre-test	Post-test
Spelling				
Easy Words	93.67 (9.15)	92.55 (9.68)	66.15 (16.43)	71.52 (19.11)
Difficult Words	89.33 (8.49)	93.33 (9.42)	57.44 (17.54)	63.64 (28.50)
Symbols	94.00 (7.46)	94.12 (7.78)	75.38 (23.79)	83.03 (27.06)
Rhyming				
Easy Words	86.00 (9.88)	90.59 (7.48)	58.46 (23.59)	61.11 (22.35)
Difficult Words	86.00 (10.79)	85.88 (9.39)	55.38 (17.92)	56.67 (15.95)
Symbols	91.33 (9.45)	94.90 (4.43)	82.56 (13.20)	81.11 (17.25)

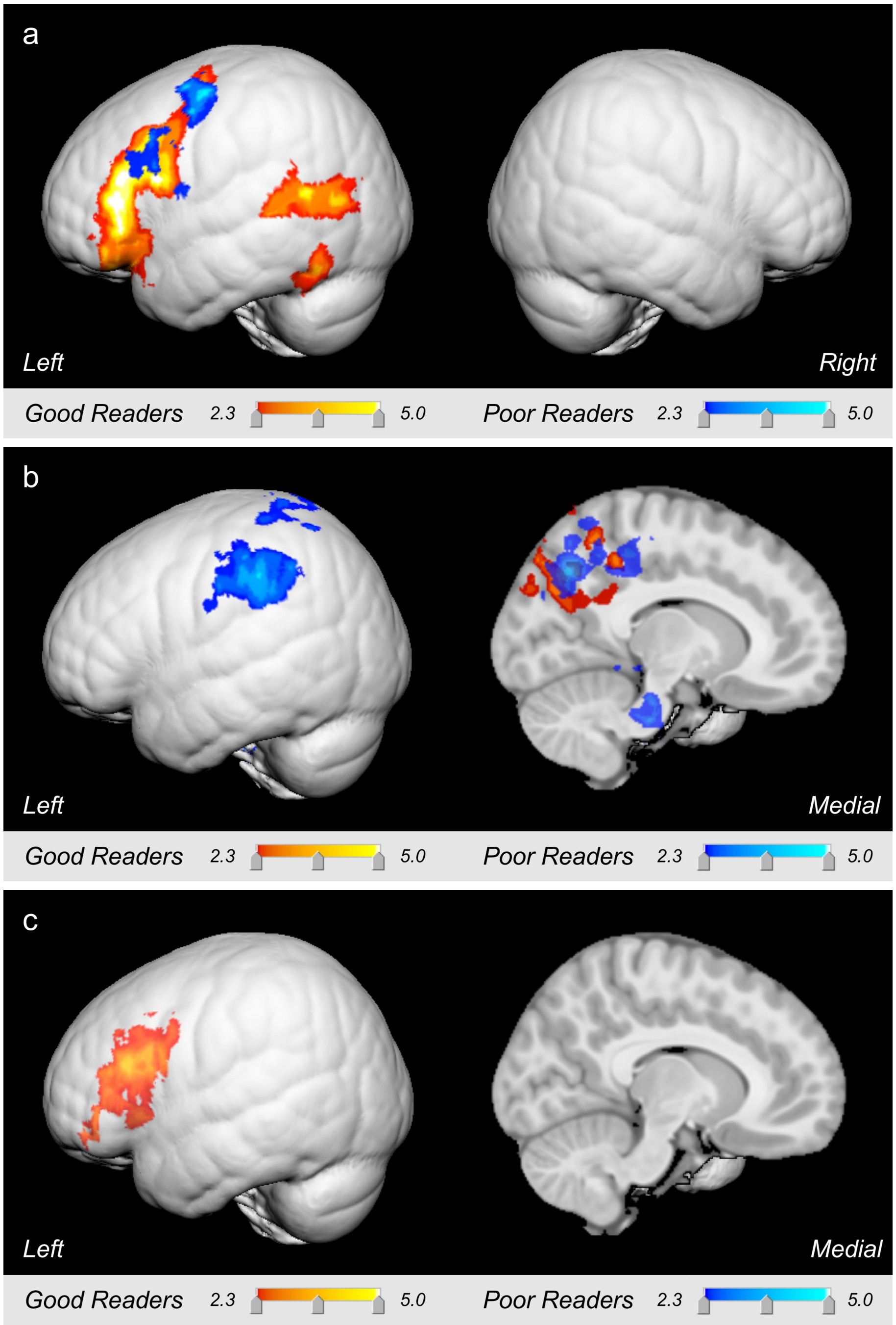
There were also significant group by condition interactions in accuracy (spelling: $F(2, 96.15) = 3.28$, $p = .042$; rhyming: $F(2, 128.84) = 10.51$, $p < .001$). Pairwise comparisons showed that good readers had similar accuracy across all conditions on the spelling task ($p > .05$), and lower accuracy on difficult words compared to symbols on the rhyming task ($p = .014$). For the poor reader group, participants had similar accuracy between easy and difficult words ($p > .05$), which was lower than the symbols condition for both spelling and rhyming ($p < .05$). There were no interactions found in reaction time, but a significant effect of condition showed that participants were faster to respond to the symbols condition than to easy and difficult words (spelling: $F(2, 112.44) = 16.96$, $p < .001$; rhyming: $F(2, 119.19) = 48.84$, $p < .001$). Finally, there was no change in accuracy or reaction time from pre-test to post-test (all $p > .10$).

Whole-Brain Activation from fMRI Tasks

Following pre-processing, each participant's whole-brain fMRI data were analyzed with FMRIB's Improved Linear Model (FILM; Woolrich, Ripley, Brady, & Smith, 2001), which corrects for autocorrelation. These models included 3 predictors of easy words, difficult words, and symbols convolved with a double-gamma hemodynamic response function and temporal derivative. The blank fixation condition was used as the baseline. There were 3 contrasts of interest for this study, which included words > symbols, easy words > difficult words, and difficult words > easy words. For the spelling and rhyming tasks separately, a multi-level linear model was used to determine whole-brain activity and differences between groups (Beckmann, Jenkinson, & Smith, 2003). This included each participant's fMRI data at the first level, which were then pooled across time points at the second level and examined for group differences at the third level. All participants were included in this analysis as the multi-level model can have missing data. It cannot, however, examine time or interaction effects because the variances will not be properly modelled (i.e., FSL does not have this capacity yet). Rather, time and interaction effects were examined using a general linear model with participants who had 2 time points of data. Both the multi-level and general linear models used mixed effects (FSL's FLAME1) with a cluster threshold of $z > 2.3$ and $p < .05$ family-wise error rate. Activation maps of the z-statistic were created on the NIH pediatric brain template for presentation purposes.

The results for the spelling task are illustrated in Figure 3.2, and clusters of significant activation are described in Appendix B. This analysis showed increased activity for words compared to symbols in left inferior frontal, middle frontal, parietal-temporal, inferior temporal, fusiform, and hippocampal gyri (Figure 3.2, panel a). There was greater activity in precuneus, medial posterior cingulate, right inferior parietal, and bilateral pons/cerebellum for easy words (Figure 3.2, panel b) and greater activity in left inferior frontal gyrus for difficult words (Figure 3.2, panel c). One region showed a significant time effect, such that there was increased activity from pre-test to post-test in bilateral cerebellum. There were no significant group or interaction effects found in the whole-brain analyses.

Figure 3.2. Whole-brain activation maps for the fMRI spelling task

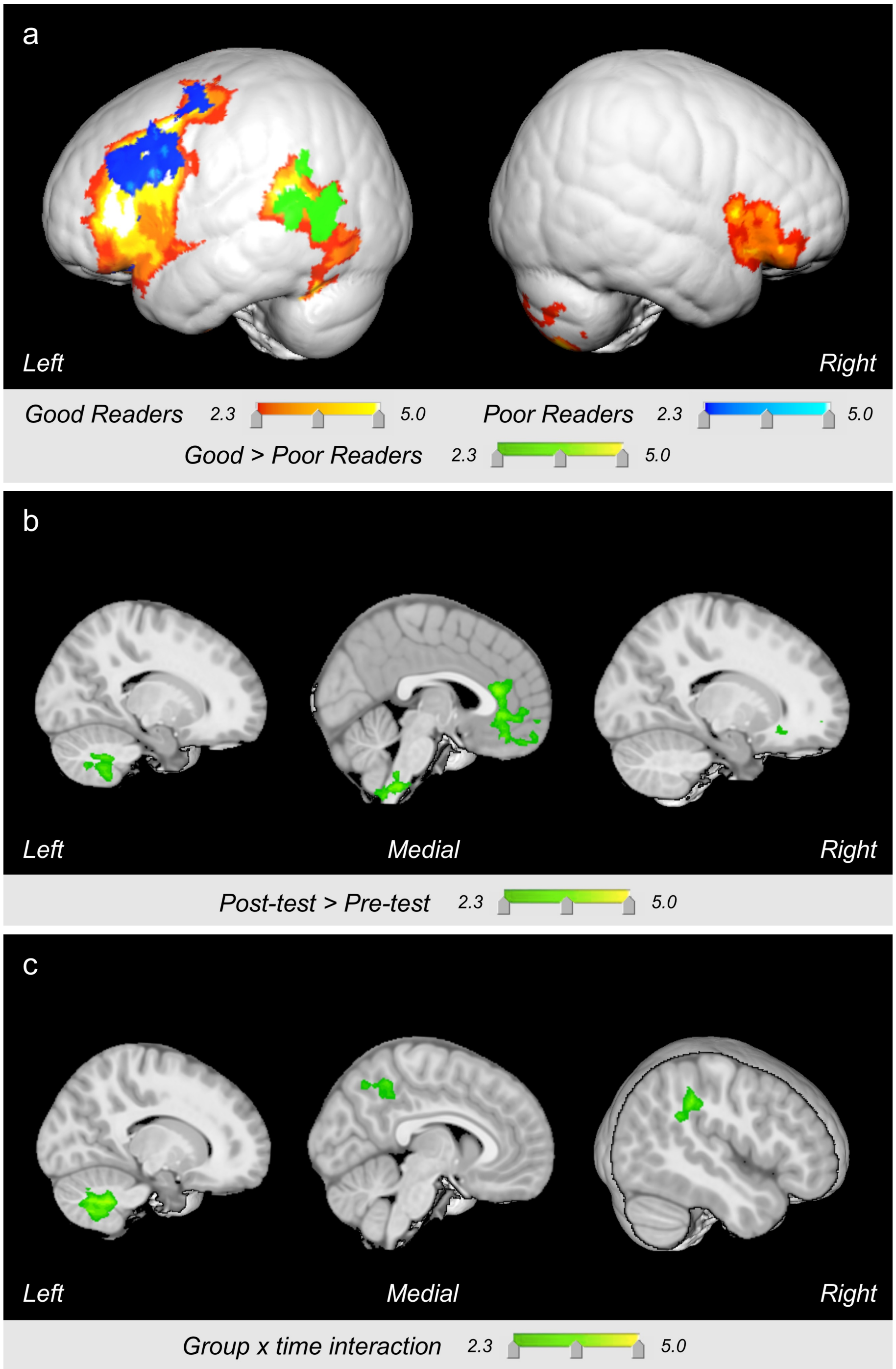


Note: Panel (a) shows greater activity for words compared to symbols condition, panel (b) shows greater activity for easy words compared to difficult words, and panel (c) shows greater activity for difficult words compared to easy words. Activation for good readers is shown in orange and poor readers in blue. Maps indicate clusters of $z > 2.3$, $p < .05$ corrected.

The results for the rhyming task are shown in Figure 3.3, and clusters of significant activation are described in Appendix C. Increased activity for words compared to symbols was shown in: bilateral inferior frontal gyrus, basal ganglia, and hippocampal gyrus; medial frontal; left middle frontal, insula, parietal-temporal, and inferior temporal gyrus; and right cerebellum (Figure 3.3, panel a). Group differences were shown in the left parietal-temporal cortex with more activity for the good readers. Furthermore, a significant time effect was shown in anterior cingulate, orbitofrontal, and cerebellar areas, which indicates that there was greater activity in these areas for difficult words over time (i.e., greater difference between difficult and easy words from pre-test to post-test) (Figure 3.3, panel b).

On the rhyming task, significant group by time interactions were shown in right parietal cortex for words compared to symbols, and in left cerebellum for easy words compared to difficult words (Figure 3.3, panel c). The right parietal cluster included precuneus, anterior intraparietal sulcus, and supramarginal gyrus. These interaction effects were further examined using FSL's featquery tool, which extracted mean percent signal change for each participant and region. In FSL, percent signal change is calculated as the difference between condition (easy words, difficult words, symbols) and the overall mean signal. Analyses with independent *t*-tests were conducted to determine group differences at each time point, as including both group and time in the model would be considered "double-dipping" (i.e., an interaction effect would be found since this is what defined the region). At pre-test, there were no activation differences between groups in parietal (easy words: $t(22) = -0.03$, $p = .977$; difficult words: $t(22) = -0.43$, $p = .675$), or cerebellar (easy words: $t(22) = 0.75$, $p = .461$; difficult words: $t(22) = 1.14$, $p = .267$) regions. Following intervention, poor readers had significantly greater activation in parietal cortex but on the easy trials only (easy words: $t(22) = -2.28$, $p = .033$; difficult words: $t(22) = -1.25$, $p = .223$). No significant effects were found in left cerebellum at post-test (easy words: $t(22) = -0.25$, $p = .802$; difficult words: $t(22) = -1.09$, $p = .289$).

Figure 3.3. Whole-brain activation maps for the fMRI rhyming task

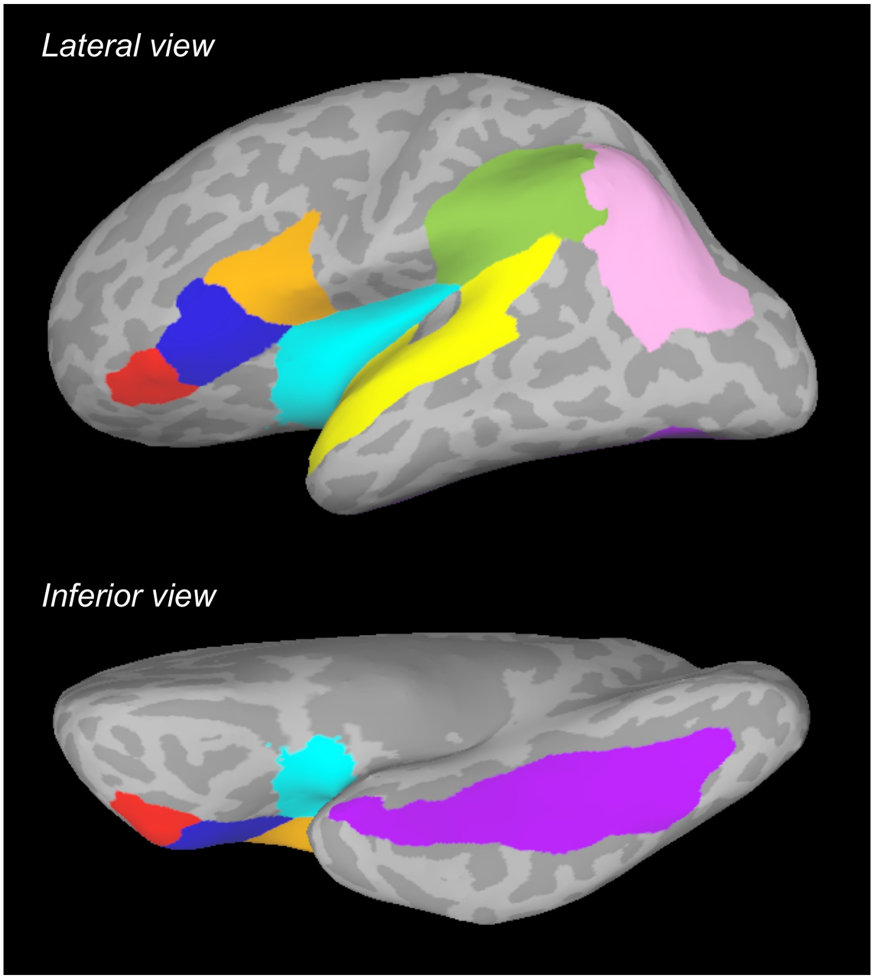


Note: Panel (a) shows greater activity for words compared to symbols condition. Activation for good readers is shown in orange, poor readers in blue, and the difference between the 2 groups in green. Panel (b) illustrates changes from pre-test to post-test in green, and panel (c) illustrates regions that had a significant group by time interaction in green. Maps indicate clusters of $z >$ 2.3, $p < .05$ corrected.

Region of Interest Analyses

It is possible that group or interaction effects were not shown in whole-brain analyses due to small sample size and power. Therefore, the purpose of the ROI analysis was to evaluate differences between groups, conditions, and time points in brain activity using regions that were defined independently from the whole-brain analyses. As described above, there were 8 regions in each hemisphere (see Figure 3.4 for an example). Mean percent signal change was extracted from each ROI using FSL's featquery tool, which was used in subsequent analyses.

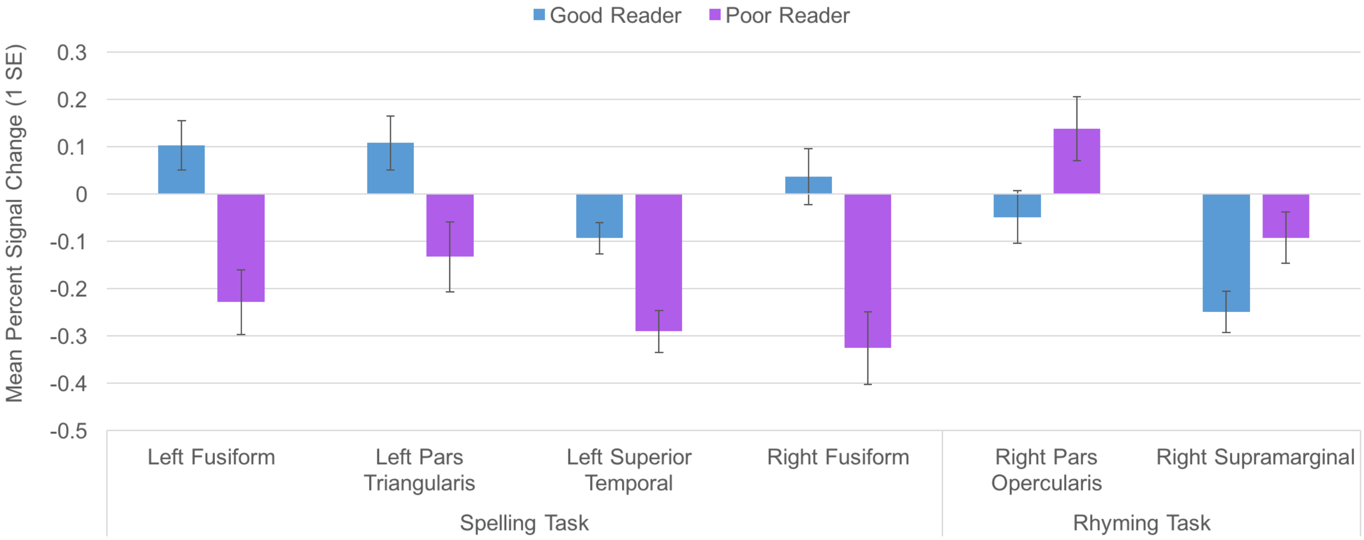
Figure 3.4. Regions of interest



Note: left hemisphere ROIs are shown for one participant. Each ROI is represented by a different colour: red (pars orbitalis), dark blue (pars triangularis), orange (pars opercularis), light blue (insula), yellow (superior temporal gyrus), green (supramarginal gyrus), pink (angular gyrus), and purple (fusiform gyrus).

Linear mixed models were conducted for the spelling and rhyming tasks separately. The models included repeated effects of time and condition as well as fixed effects of group, time, and condition with interactions between these factors. On the spelling task, there were significant group differences in functional activity in left fusiform, *F*(1, 159.37) = 15.02, *p* < .001, left pars triangularis, *F*(1, 156.86) = 6.67, *p* = .011, left superior temporal, *F*(1, 156.60) = 12.89 = 12.89, *p* < .001, and right fusiform, *F*(1, 153.99) = 13.97, *p* < .001 (see Figure 3.5). The graphs illustrate that poor readers had less activation in these regions compared to the overall mean, which suggests that these areas were de-activated when they were reading. The good readers, however, did not differ significantly from zero indicating that the spelling task did not activate these regions in the good readers. A significant time effect was found in right supramarginal gyrus, such that there was increased activity from pre-test to post-test in this region, *F*(1, 171.10) = 4.41, *p* = .037. Also, an effect of task condition was shown in 2 regions: there was less signal for easy and difficult words compared to symbols in right angular gyrus, *F*(2, 108.46) = 3.75, *p* = .027, and right supramarginal gyrus, *F*(2, 114.39) = 3.46, *p* = .035. Finally, there was a group by time interaction found in right superior temporal cortex, *F*(1, 147.62) = 4.61, *p* = .033, which was accounted for by increased activity over time for the good reader group.

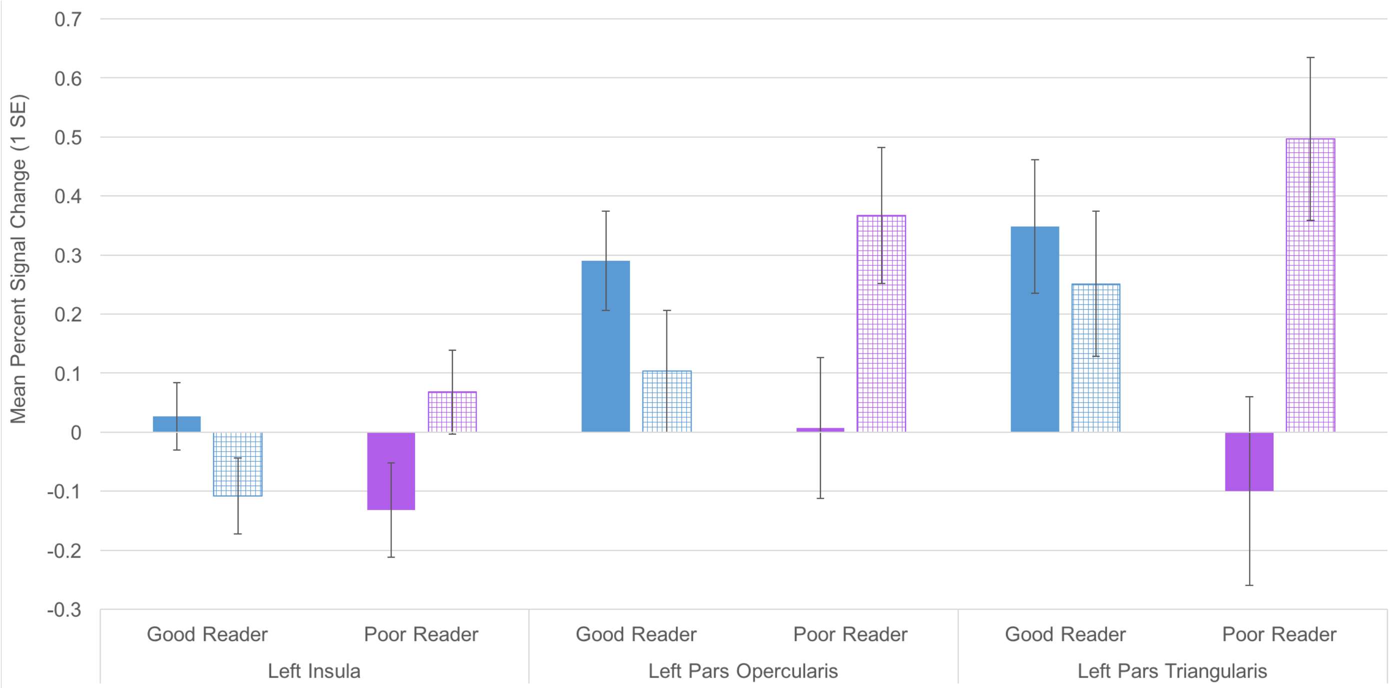
Figure 3.5. Main effect of group on the fMRI spelling and rhyming tasks in regions of interest



On the rhyming task, group differences were shown in right pars opercularis, *F*(1, 155.05) = 4.56, *p* = .034, and right supramarginal gyrus, *F*(1, 145.74) = 5.16, *p* = .025 (see Figure 3.5). In comparison to the spelling task, different trends were shown for the rhyming task.

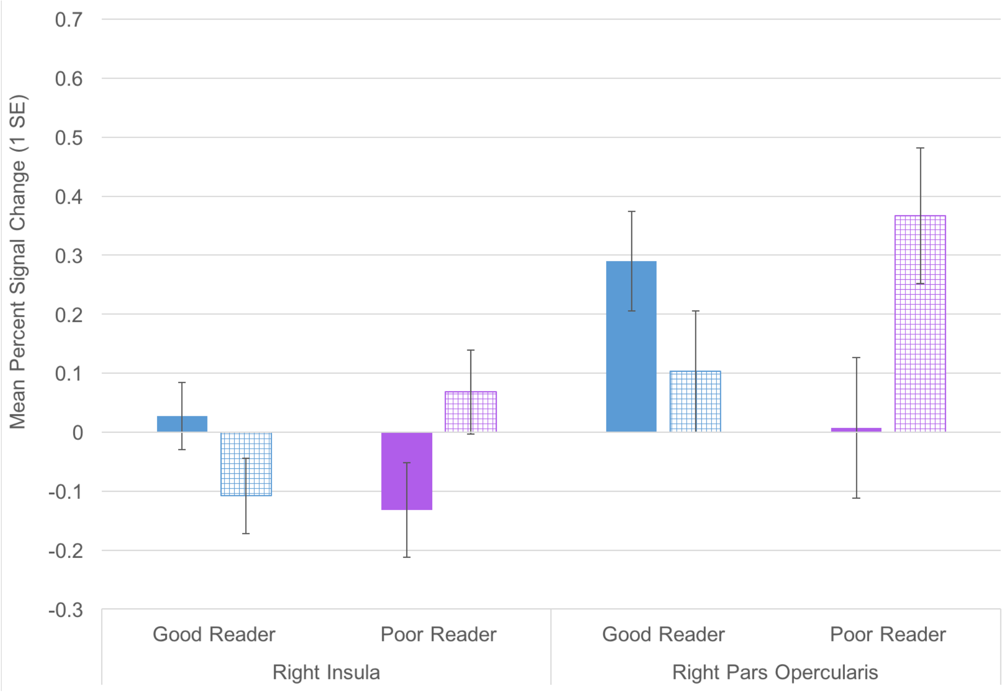
Poor readers showed greater activity in right pars opercularis whereas good readers showed de-activation in right supramarginal gyrus. The results also found that there was increased activity from pre-test to post-test in left angular gyrus, *F*(1, 165.38) = 4.04, *p* = .046, and right pars opercularis, *F*(1, 155.05) = 3.98, *p* = .048. Finally, significant group by time interactions were found in 5 regions, including left insula, *F*(1, 159.09) = 5.95, *p* = .016, left pars opercularis, *F*(1, 154.65) = 6.61, *p* = .011, left pars triangularis, *F*(1, 163.42) = 6.64, *p* = .011, right insula, *F*(1, 154.12) = 3.91, *p* = .050, and right pars opercularis, *F*(1, 155.05) = 4.75, *p* = .031 (see Figures 3.6 and 3.7). These graphs illustrate that there was increased activity for the poor readers in these regions after intervention, whereas the good readers had similar activation over time. Furthermore, there was significant activity in frontal regions but activity in insular cortex was not different from zero.

Figure 3.6. Group by time interactions in the left hemisphere on the fMRI rhyming task



Note: solid bars indicate pre-test (time 1) and checkered bars indicate post-test (time 2).

Figure 3.7. Group by time interactions in the right hemisphere on the fMRI rhyming task



Note: solid bars indicate pre-test (time 1) and checkered bars indicate post-test (time 2).

fMRI and Behavioural Predictors of Improved Literacy Skills

One objective of this study was to examine the fMRI and behavioural predictors of improvement in literacy skills. Difference scores from pre-test to post-test were calculated for the KTEA-II decoding and word recognition fluency measures, as the analyses above indicated significant time differences on these measures only (see Table 3.2). These difference scores were used as the dependent variables. Predictors included pre-test measures of: fMRI percent signal change from regions that showed significant group by time interaction effects (i.e., on the rhyming task, this included left cerebellum and right parietal cortex from whole-brain analyses; left insula, left pars opercularis, left pars triangularis, right insula, and right pars opercularis from ROI analyses); KTEA-II and WJ III COG language, short-term memory, and processing speed measures; and ADHD ratings from the Conners 3 Parent scale. Due to small sample size, the effect of each group of predictors (i.e., fMRI, language, short-term memory, processing speed, and Conners 3) were examined in separate linear regression models.

The results indicated that activation from the rhyming task in left pars triangularis was associated with improved decoding skills ($\beta = -1.02$, $p = .048$). Specifically, lower activity at pre-test was related to greater reading skills over time. The overall model with fMRI predictors, however, was not significant (adjusted $R^2 = .08$, $F(5, 22) = 1.48$, $p = .238$). Also, improved word recognition fluency over time was related to higher scores on listening comprehension, $\beta = 0.45$, adjusted $R^2 = 0.14$, $F(2, 26) = 4.24$, $p = .05$, and memory span, $\beta = 0.52$, adjusted $R^2 = 0.19$, $F(2, 26) = 4.24$, $p = .025$, at pre-test. The other pre-test measures of phonological awareness, processing speed, working memory, and ADHD ratings were not significant predictors of improvement in reading (all models $p > .05$).

Finally, analyses were conducted to determine whether differences in brain activity over time were related to improved reading skills. Difference scores from pre-test to post-test were calculated for brain regions that showed significant group by time interaction effects on the rhyming task (left cerebellum and right parietal cortex from whole-brain analyses; left insula, left pars opercularis, left pars triangularis, right insula, and right pars opercularis from ROI analyses). These difference scores were included as predictors in linear regression models with the same dependent variables described above. The results showed that increased activity in the right parietal cortex over time was a significant predictor of gains in decoding, $\beta = 0.49$, adjusted $R^2 = 0.21$, $F(2, 25) = 6.03$, $p = .007$. Changes in activation over time in left cerebellum, pars opercularis, or pars triangularis were not associated with improved decoding or word recognition fluency skills ($p > .50$).

Discussion

Brain Regions Involved for Orthographic and Phonological Tasks

This study showed that the brain regions activated for orthographic (spelling) and phonological (rhyming) fMRI tasks were largely similar, and this included a left hemisphere network of inferior frontal, middle frontal, parietal-temporal, inferior temporal, fusiform, and hippocampal gyri. This result replicates previous research showing that reading and language tasks predominately activate the left hemisphere, particularly in average readers (Shaywitz et al., 2006). It also emphasizes that similar brain regions may be involved for phonological and orthographic processing, although direct comparisons between these tasks were not conducted. For the rhyming task, additional regions of left insula, medial frontal, right inferior frontal, right basal ganglia, right hippocampal gyrus, and right cerebellum were activated. Activation of right hemisphere regions may be due to several reasons, including compensatory mechanisms and level of difficulty for poor readers (Shaywitz et al., 2003). However, the right hemisphere is also activated in typically-developing readers (e.g., Bitan et al., 2007; Booth et al., 2004; Church et al., 2008; Olulade, Napoliello, & Eden, 2013) and therefore both good and poor readers may recruit the right hemisphere due to increased processing demands of the rhyming task in comparison to the spelling task (e.g., matching of rhyming words while keeping phonemes in working memory). In addition, areas of activation in medial frontal cortex may be related to cognitive processes such as response selection or error monitoring (Carter & van Veen, 2007; Rushworth, Buckley, Behrens, Walton, & Bannerman, 2007). Finally, involvement of the cerebellum is consistent with previous studies showing that this structure is activated during various cognitive and language tasks, including reading (Stoodley & Schmahmann, 2009), and the right cerebellum is specifically activated possibly due to the structural connectivity between cortical regions and the contralateral cerebellar hemisphere (Catani & Thiebaut de Schotten, 2008).

Group differences on the spelling task were not shown in whole-brain analyses, which may be due to power and variability between participants. In regions of interest, however, poor readers had less activation than good readers in left fusiform, left pars triangularis, left superior temporal, and right fusiform regions, which was primarily related to "de-activation" relative to a baseline condition when they were completing the task. The task was quite difficult for the poor readers and therefore de-activation may reflect that these regions were not engaged during a difficult task (Pugh et al., 2008). Furthermore, the poor readers may have activated other brain regions when reading, but these results were not significant in whole-brain or region of interest analyses.

On the rhyming task, there was significantly less activation for the poor readers compared to good readers in left parietal-temporal cortex. This result is consistent with previous research (Shaywitz et al., 2002) and suggests that the left parietal-temporal cortex is less well developed in children with reading impairments. Region of interest analyses also indicated that poor readers significantly activated the right pars opercularis, but the good readers did not, suggesting a compensatory recruitment of this area related to increased processing demands (Shaywitz et al., 2003). Furthermore, good readers had de-activation in the right supramarginal gyrus. In contrast to the poor readers who showed de-activation that may be related to task difficulty, the good readers had accuracies that were near-perfect. Therefore, de-activation in this case could be because the task was relatively easy and this may have elicited some brain regions to be less activated in comparison to children with reading impairments. Research has shown that some brain regions are less activated following a task, including lateral parietal regions, and this has been termed the "default mode network" (reviewed in Raichle, 2015). It is possible that de-activation for the good readers is a reflection of the default mode network.

Effect of Intervention on Behavioural and fMRI Reading Tasks

In the current sample, poor readers improved in decoding skills and word recognition fluency after 3 months of intervention, indicating positive effects of school-based reading programs that are consistent with published literature (Al Otaiba & Torgesen, 2007; Cavanaugh, Kim, Wanzek, & Vaughn, 2004). However, the good readers also improved in word recognition fluency, which may be the result of practice effects on this task. While students were assessed with alternate forms to mitigate this effect, it is possible that students received practice with reading words more quickly and therefore improvements were shown on this timed reading task.

Across groups, increased brain activity over time was found in bilateral cerebellum and right supramarginal gyrus on the spelling task. There was also increased activity in left angular gyrus, right pars opercularis, medial anterior cingulate, medial orbitofrontal, and bilateral cerebellar areas on the rhyming task over time. These results indicate that both good and poor readers, on average, showed significant changes in brain activity after 3 months of instruction. Anterior cingulate gyrus has been implicated with error monitoring (Carter & van Veen, 2007), orbitofrontal cortex in decision making and default mode network (Bechara, Damasio, & Damasio, 2000; Raichle, 2015), and angular/supramarginal gyri in mapping orthography to phonology (Booth et al., 2004) as well as processing of semantics (Binder, Desai, Graves, & Conant, 2009). The region of the cerebellum that was activated on these tasks is likely related to motor movement (Stoodley & Schmahmann, 2009), such as button presses. Therefore, results suggest that activation changes in these regions may be due to various cognitive, language, or motor processes that changed with normal development or practice. For example, studies have shown areas of increased activity in medial frontal, precuneus, and lateral occipital regions with practice (Jolles, Grol, Van Buchem, Rombouts, & Crone, 2010; Weissman, Woldorff, Hazlett, & Mangun, 2002), whereas other studies have shown the opposite effect with regions of decreased activity (Bitan, Manor, Morocz, & Karni, 2005; Chein & Schneider, 2005; Tomasi, Ernst, Caparelli, & Chang, 2004). One review showed that learning is associated with both increases and decreases in brain activity and these may be related to task-specific and attentional processes that occur during learning (Kelly & Garavan, 2004). Furthermore, although good and poor readers differed in accuracy and reaction time, there was no change between time points. This result indicates that changes in brain activity were not related to practice with the fMRI tasks per se, but rather with the development of reading skills and strategies or increased comfort with the MRI environment. Therefore, increased brain activity over time may be because of additional recruitment of brain regions related to increased learning and practice.

Intervention effects, as evidenced by group by time interactions, were shown on the fMRI rhyming task. This included regions of bilateral pars opercularis, bilateral insula, left pars triangularis, left cerebellum, and right parietal cortex that encompassed the precuneus, anterior intraparietal sulcus, and supramarginal gyrus. After intervention, poor readers had increased activity in frontal regions, which is consistent and replicates other treatment studies (Shaywitz et al., 2004). Changes over time in right parietal regions were related to poor readers having significantly greater activity than good readers after intervention, but on easy words only. The regions activated in parietal cortex are associated with language skills as discussed above, but also orienting of attention in anterior intraparietal sulcus (Geng & Mangun, 2009) and default mode in precuneus (Raichle, 2015). Therefore, these results suggest that cognitive or language processes in the brain may be involved with improved reading skills. On the spelling task, there was a group by time interaction in right superior temporal gyrus, which was related to increased activity for the good readers over time. This result may be related to practice or typical development and not to reading treatment because the good reader group did not receive additional reading supports. Furthermore, most of the regions that changed for the poor readers were different from the good readers, but also different from the regions that showed increased activity over time for both groups. Therefore, there were unique and non-overlapping regions of the left inferior frontal, bilateral insula, and right parietal cortex that had increased activation for the poor readers only, suggesting effects of reading treatment in these brain regions. These results indicate that the poor readers were using different areas of the brain to process words following 3 months of intervention.

Finally, the predictors of improved reading skills were analyzed across groups. These results showed that students who had better attention span and listening comprehension had greater gains in word recognition fluency, and brain activity in left pars triangularis and right parietal cortex was associated with gains in decoding skills. It is reasonable that activation from the fMRI rhyming task was related to changes in nonsense word decoding, as both require phonological processing (i.e., knowledge of sounds in words) and are consistent with other imaging studies (Temple et al., 2003). Furthermore, changes in parietal regions may be due to increased attentional or default mode networks (Raichle, 2015), as described above. These results illustrate that both functional neuroimaging and cognitive testing results are related to gains in reading skills outside of the MRI scanner, which has been shown in other intervention studies (Hoeft et al., 2007b). It is unlikely, however, that neuroimaging on its own will be used to predict reading success for all students in the classroom. Rather, the results demonstrate that functional brain activity is predictive of future reading gains, which provides further evidence for the neurobiological bases of reading development.

Effect of Task Difficulty

Task difficulty was measured by accuracy and reaction time on the fMRI tasks as well as by differences in brain activity between easy and difficult blocks. On the spelling task, there were similar accuracies and reaction times between easy and difficult word blocks, however, there was greater cortical activity in left inferior frontal gyrus when reading difficult words. This result is consistent with Cattinelli et al. (2013) who showed that the inferior frontal gyrus, particularly the pars opercularis, is involved with task difficulty. Furthermore, subregions of the inferior frontal gyrus may be sensitive to different components of word processing, including phonological awareness in pars opercularis and semantics in pars triangularis (Jobard et al., 2003). Therefore, differences in processing relatively easy and difficult words may be due to the underlying roles of these regions. The spelling task also showed greater activity for easy words compared to difficult words in precuneus, medial posterior cingulate, right inferior parietal, and bilateral pons/cerebellum. As described above, these regions have been implicated in default mode or attentional networks (Raichle, 2015), and it is possible that easy word trials activated these networks.

There were also effects of difficulty on the rhyming task, such that the good readers had lower accuracies on the difficult trials when compared to symbols. This effect was not shown for the poor reader group – these students had similar accuracy between easy and difficult words, which was lower than the symbols condition. Furthermore, there was increased activity from pre-test to post-test in several brain regions, but on difficult trials only. This included anterior cingulate, orbitofrontal, and cerebellar areas, which may be related to increased recruitment of attention (Geng & Mangun, 2009), decision making (Bechara et al., 2000), error monitoring (Carter & van Veen, 2007), and default mode (Raichle, 2015) systems. This result suggests that, across groups, these systems were involved when reading difficult words at the second MRI scan.

Limitations

This study had a relatively small sample size overall and in comparison to the intervention study in Chapter 2. While most students from Chapter 2 were invited to the MRI study, fewer were interested in participating. Therefore, results may have been biased towards a certain sample of students, however, results in the current study were generally consistent with the previous literature. It will be important in future studies to have larger samples that are more representative of the school population.

The fMRI tasks were more difficult for the poor readers in this study, despite piloting these tasks with a different group of poor readers. This result is similar to previous studies that showed an overall low accuracy for children with reading disabilities (e.g., Simos et al., 2002, had 51% accuracy before intervention and 60% after intervention on a non-word rhyming task). This result implies that the tasks were too difficult for our group of poor readers, however, differences in relative difficulty were still shown in whole-brain and region of interest analyses.

Conclusions

The current study showed that the effects of school-based interventions were more evident on a phonological rhyming task, rather than on an orthographic spelling task. The poor readers showed a pattern of increased activation in bilateral inferior frontal, bilateral insula, right parietal, and left cerebellum following intervention, and activity in left pars triangularis and right parietal regions were associated with decoding skills but not word recognition fluency. As the target of these interventions included both phonological and orthographic skills, this implies that further emphasis on orthographic awareness is needed.

There were also several brain regions that showed activation differences between easy and difficult words, including left inferior frontal gyrus, anterior and posterior cingulate, medial orbitofrontal, right parietal, and bilateral cerebellum. It is possible that task difficulty was related to language processing in inferior frontal and supramarginal gyri (Binder et al., 2009; Booth et al., 2004; Jobard et al., 2003), error monitoring in anterior cingulate cortex (Carter & van Veen, 2007), decision making in orbitofrontal cortex (Bechara et al., 2000), orienting of attention in intraparietal sulcus (Geng & Mangun, 2009), and default mode network in orbitofrontal, precuneus, posterior cingulate, and inferior parietal regions (Fransson & Marrelec, 2008; Raichle, 2015). The location of cerebellar activation was likely related to motor movements (Stoodley & Schmahmann, 2009), including button presses. Furthermore, intervention effects appeared across easy and difficult trials, except for areas of the right parietal cortex. In this region, poor readers showed greater activity after intervention but only on the easy trials, implying greater recruitment of the right parietal cortex for relatively easier words. Overall, this study showed that there are multiple brain regions involved with relative task difficulty, and intervention effects are shown for both easy and difficult words.

CHAPTER 4. DIFFERENCES IN WHITE AND GREY MATTER STRUCTURES BEFORE AND AFTER READING INTERVENTION

Introduction

In the General Introduction, it was discussed that children and adults with reading impairments have differences in grey matter volume (Eckert et al., 2016; Richlan, 2012), cortical thickness (Ma et al., 2015), and surface area (Frye et al., 2010) as well as in white matter indices (Deutsch et al., 2005; Klingberg et al., 2000) when compared to typically developing readers. The loci of these differences have predominantly been in the grey matter of bilateral superior temporal gyri (Richlan et al., 2012) and in left parietal-temporal white matter (Deutsch et al., 2005; Klingberg et al., 2000), implicating deficits in the arcuate fasciculus/superior longitudinal fasciculus and corona radiata (Vandermosten et al., 2012b). Overall, these previous studies indicate that poor readers have differences in both white and grey matter structures in comparison to good readers.

The effect of reading intervention on brain structure, however, has received less empirical review than the effects on brain function. In a prospective design, Krafnick and colleagues (2011) examined the Lindamood-Bell Seeing Stars program with 11 children with reading impairments. There were changes in grey matter volume in fusiform gyrus, hippocampus, precuneus, and cerebellum following this 8-week program, and there were fewer changes after a period of no intervention. Another study showed that differences in cortical thickness between poor and good readers persisted regardless of whether the child had remediated reading skills or not (Ma et al., 2015). These studies, however, did not include a control group (Krafnick et al., 2011) or a prospective design (Ma et al., 2015) and therefore grey matter structures before and after intervention in comparison to typically developing readers could not be evaluated.

In studies of white matter microstructure, one research group examined an intensive reading program that was delivered for 100 hours over 6 months (Keller & Just, 2009). Poor readers had increased fractional anisotropy in left anterior frontal regions after treatment, and this same region had lower fractional anisotropy compared to good readers before the intervention started. Another study examined the outcomes of children who were considered treatment "responders" or "non-responders" to reading intervention (Farris et al., 2011). Results showed that increased fractional anisotropy in the corpus callosum was associated with higher functional connectivity between the inferior frontal lobes, however, "responders" did not differ from "non-responders" in fractional anisotropy in the corpus callosum. Both studies used a whole-brain analytic approach, whereas tractography is typically used to examine white matter tracts based on *a priori* knowledge of anatomy or results from a functional imaging study. Several intervention studies have used tractography, and results showed that fractional anisotropy in right superior longitudinal fasciculus was positively correlated with later reading skills after non-specific intervention (Hoeft et al., 2011) and that white matter connectivity was correlated with reading skills after a tier 2 program (Davis et al., 2010). These effects were analyzed before (Hoeft et al., 2011) or after (Davis et al., 2010) treatment, and a control group of typically developing readers was included in only one study (Hoeft et al., 2011). Similarly to the grey matter results, an evaluation before and after treatment with a comparison to a control group has not been conducted with tractography.

These previous studies of grey and white matter structures are important and illustrate that brain structures may change after treatment or may be predictive of later reading skills. However, a prospective study of changes in various grey matter metrics (volume, area, thickness) and white matter tractography has not been completed, which is important because there may be aspects of grey or white matter that show subtle changes following reading treatment. Therefore, the purpose of the current study was to examine changes in grey matter volume, cortical thickness, surface area, and white matter microstructure before and after intervention in brain regions implicated in language and reading development. Furthermore, another purpose of this study was to evaluate whether grey and white matter indices before intervention were associated with gains in reading skills over time. These grey and white matter indices were obtained using magnetic resonance imaging (MRI) through T1-weighted images and diffusion tensor imaging (DTI).

Methods

Participants

The same participants from the functional MRI study in Chapter 3 took part in the current study, which included 22 good readers and 13 poor readers at pre-test as well as 17 good readers and 12 poor readers at post-test. The poor readers received small group instruction or intensive reading supports, and were analyzed together because of small sample size. Good readers received no additional reading supports at school. Further details on each group's performance on the screening and reading measures are provided in Chapter 3 and Tables 3.1 and 3.2.

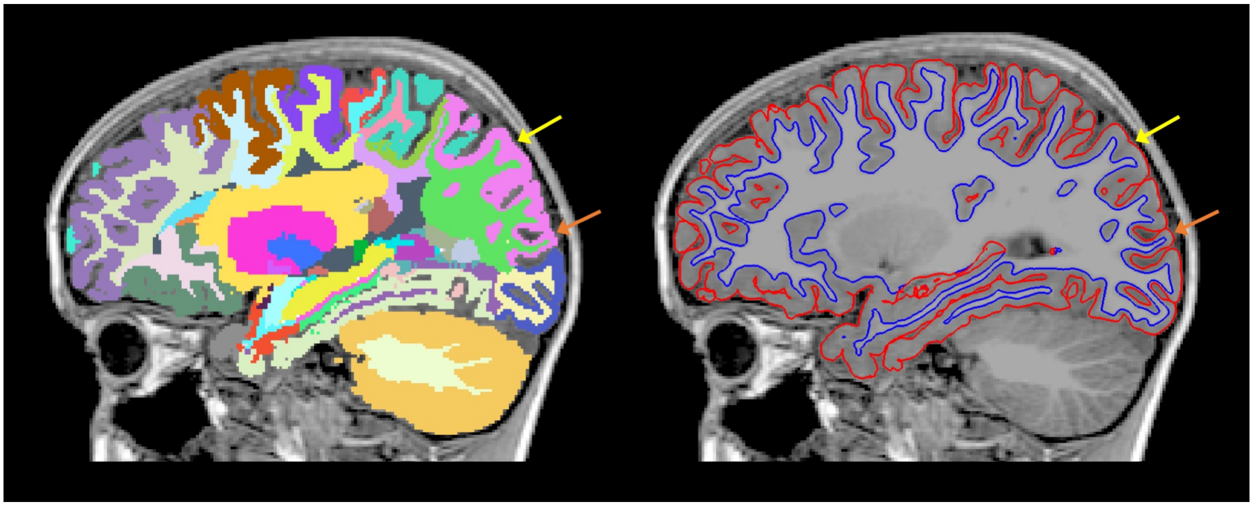
MRI Measures

All participants first practiced in a mock scanner prior to the real scan. A 3.0 Tesla GE Discovery MR 750 scanner with a 32-channel head coil was used to collect MRI data, and images of interest for the current study included 3D-T1 and diffusion tensor imaging (DTI) scans. Participants watched a movie of their choice during these structural scans. The same 3D-T1 sequence from Chapter 3 was used in the current study. This included images that were acquired in 170 sagittal slices, with the following parameters: field-of-view (FOV): 256 x 256 mm; matrix: 256 x 256; TR: 7.6 ms; TE: 3.5 ms; flip angle: 8 deg; slice thickness: 1 mm; voxel size: 1 mm^3; scan length: 5 minutes. A single-shot echo planar imaging (EPI) sequence with high angular diffusion imaging (HARDI) was used for the DTI scan. Sequence parameters included: FOV: 256 x 256 mm; matrix: 224 x 224; TR: 7000 ms; TE: 60 ms; flip angle: 90 deg; phase encode direction: A/P; 60 gradient directions; b = 1000 s/mm^2; 3 b=0 images; 60 axial slices; slice thickness: 2 mm; reconstructed voxel size: 0.875 x 0.875 x 2 mm^3; scan length: 7.5 minutes.

Grey Matter Analyses

This study used a region of interest (ROI) approach with same ROIs from Chapter 3 (see Figure 3.4), which included bilateral inferior frontal gyrus (subdivided into 3 regions of pars orbitalis, pars triangularis, and pars opercularis); insular cortex; superior temporal gyrus; supramarginal gyrus; angular gyrus; and fusiform gyrus. In Freesurfer version 5.3.0 (Dale et al., 1999), the following steps were conducted automatically for each participant: affine transformation to the MNI305 brain atlas, intensity normalization, skull stripping, white matter segmentation, removal of topological defects, inflation of the brain onto a sphere, registration to a spherical atlas, assignment of labels for each cortical surface, and computation of statistics for each region based on the Desikan-Killiany atlas (Desikan et al., 2006). For all participants, manual edits to the white matter mask were necessary as the program did not accurately identify the boundary between the pial surface and skull. Two data sets (1 good reader at pre-test and 1 poor reader at pre-test) were excluded at this point because of poor T1 image quality due to excessive motion. Total intracranial volume as well as total grey matter volume, total surface area, and average cortical thickness for each ROI were calculated by Freesurfer and used in subsequent analyses (see Figure 4.1).

Figure 4.1. Grey and white matter segmentation in Freesurfer



Note: In one participant, the left panel illustrates Freesurfer's parcellation into cortical and subcortical regions, and the right panel illustrates the segmentation between white and grey matter boundaries. The arrows point to the left angular gyrus, and highlights the difference between relatively thinner cortex (orange arrow) and relatively thicker cortex (yellow arrow).

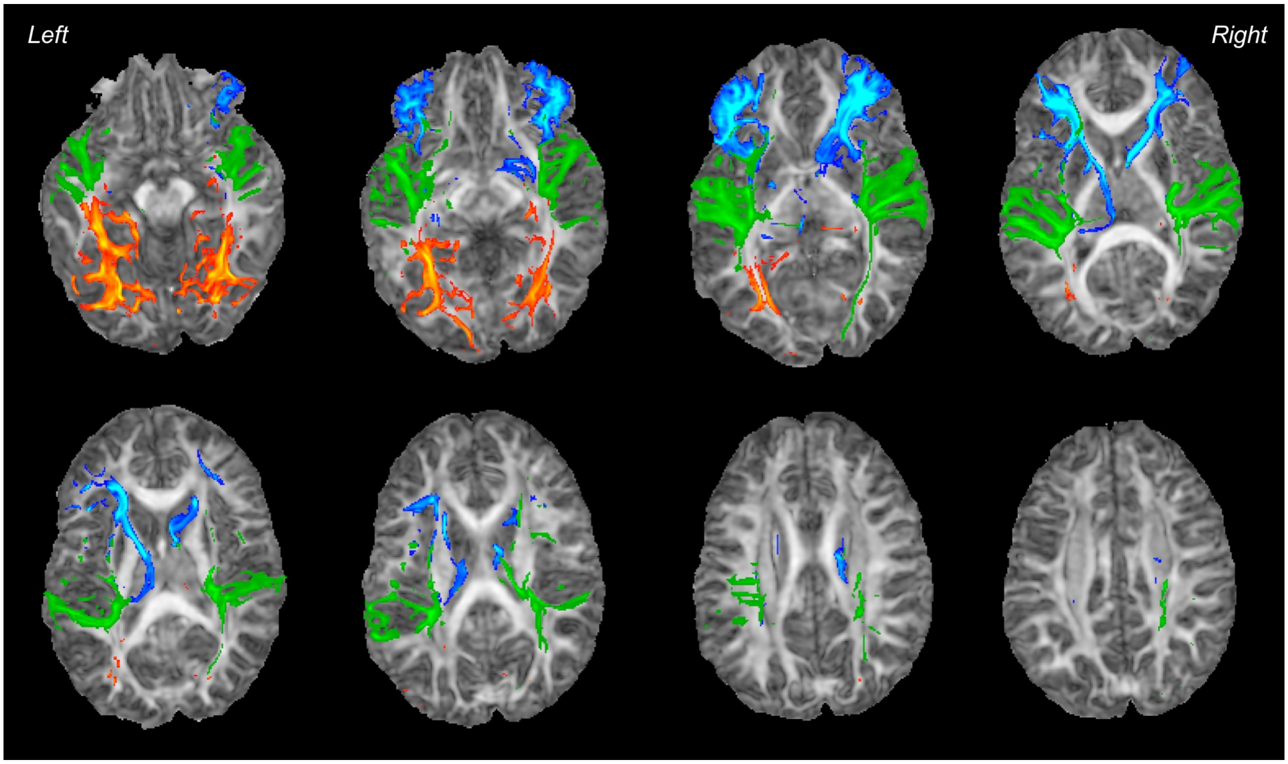
White Matter Analyses

TORTOISE version 2.5.1 (Pierpaoli et al., 2010) and FMRIB's Diffusion Toolbox version 3.0 (FDT; Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007; Behrens et al., 2003) were used for pre-processing and analysis of the DTI data. The following pre-processing steps were conducted in TORTOISE: motion and eddy current correction with reference to the first b0 image using mutual information-based registration and b-matrix reorientation (Rohde, Barnett, Basser, Marenco, & Pierpaoli, 2003) and identification and removal of directions that had excessive motion through visual inspection (i.e., signal drop out). Participants were included in the analyses if they had at least 30 directions of data to ensure robust estimation of diffusion tensors (Jones, 2004). With this criterion, there were 4 data sets excluded (2 good readers at pre-test, 1 poor reader at pre-test, and 1 poor reader at post-test).

Diffusion tensors were computed in FSL's DTIFIT program using least squares estimation. Tensors are the mathematical representation of a diffusion ellipsoid (i.e., cylindrical or circular shape of diffusion), which are represented by a matrix of 3 eigenvalues (Kingsley, 2006). These eigenvalues are then used to create maps of fractional anisotropy (FA) and mean diffusivity (MD). FA is the relative difference between the largest eigenvalue and the other 2 eigenvalues. It describes the amount of diffusion that is anisotropic, where higher FA indicates more anisotropy and this is expected in white matter tissues. Lower FA may occur in white matter when it is less developed or damaged. In contrast, MD is the mean of the 3 eigenvalues and it describes the amount of diffusion in the tensor. Higher MD values indicate more diffusion, which is expected in cerebrospinal fluid or damaged tissue. The purpose of examining both FA and MD is because each metric may be sensitive to different types of white matter changes, such as increased levels of water, cell proliferation, or inflammation (Alexander et al., 2007).

DTIFIT was followed by Bayesian estimation of diffusion parameters with FSL's BEDPOSTX and probabilistic tractography with FSL's PROBTRACKX (Behrens et al., 2003; 2007). Each participant's white matter image and ROIs from Freesurfer were converted into DTI space using linear registration and nearest neighbour interpolation in FSL (FLIRT; Jenkinson et al., 2002). In PROBTRACKX, each ROI was included as a seed region (starting point) with the remaining ipsilateral ROIs included as target regions (end point). The white matter mask on the ipsilateral side was used as a waypoint mask, which ensured that the paths went through white matter and not grey matter. Also, the contralateral hemisphere was used as an exclusion mask to ensure that paths remained in the same hemisphere (i.e., paths did not cross into the contralateral hemisphere). Tracking parameters included 5000 samples with a maximum of 2000 steps, 0.5 mm step length, and 0.2 curvature threshold (i.e., approximately 80 deg). The paths created by the program were thresholded at 0.25% of the probabilistic index of connectivity to remove outliers, and then mean FA and MD was calculated within each path (see Figure 4.2).

Figure 4.2. White matter pathways from probabilistic tractography



Note: An example of white matter pathways for one participant. Blue indicates paths from inferior frontal gyrus (pars orbitalis), green indicates paths from superior temporal gyrus, and orange indicates paths from fusiform gyrus. Lighter colours denote a higher probability that paths went through that region.

Results

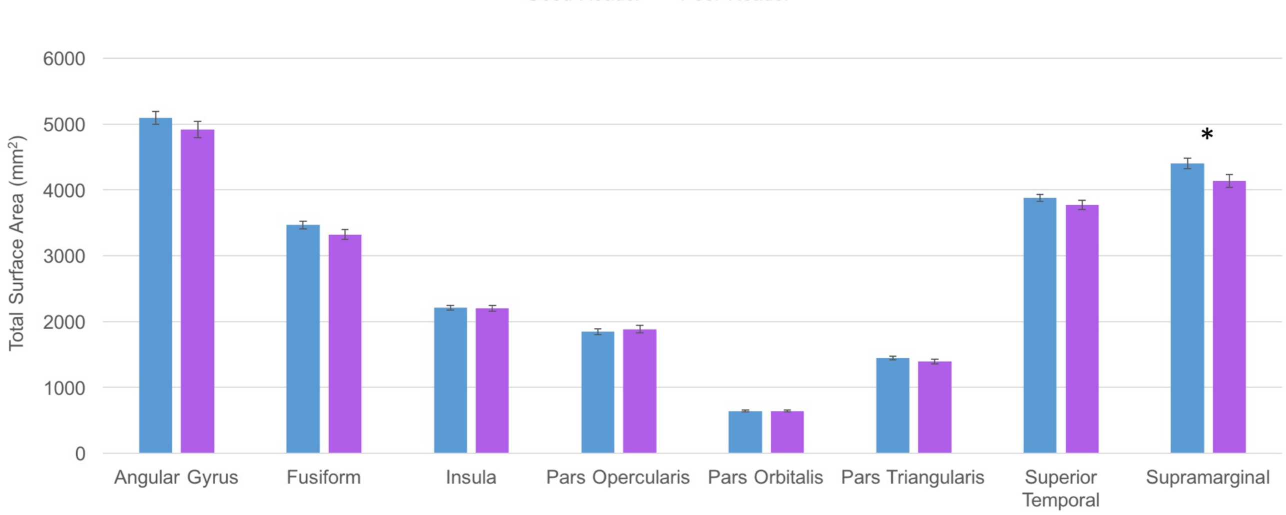
Grey Matter Structures

Linear mixed models were used to evaluate differences between reading groups and time points. Previous studies have shown that gender and brain volume may impact the results of grey matter analyses (Altarelli et al., 2013; Eckert et al., 2016; Evans et al., 2013) therefore, the current models included time as a repeated effect, reading group, time, and gender as fixed effects factors, and total intracranial volume as a fixed effect covariate. There was also a fixed effect interaction between reading group and time, however, an interaction between reading group and gender was not explored in this study due to small sample size (i.e., 5 girls in poor reader group). In each ROI, total grey matter volume, total surface area, and average cortical thickness were used as the dependent variables. Results are reported with an uncorrected *p*-value due to small sample size.

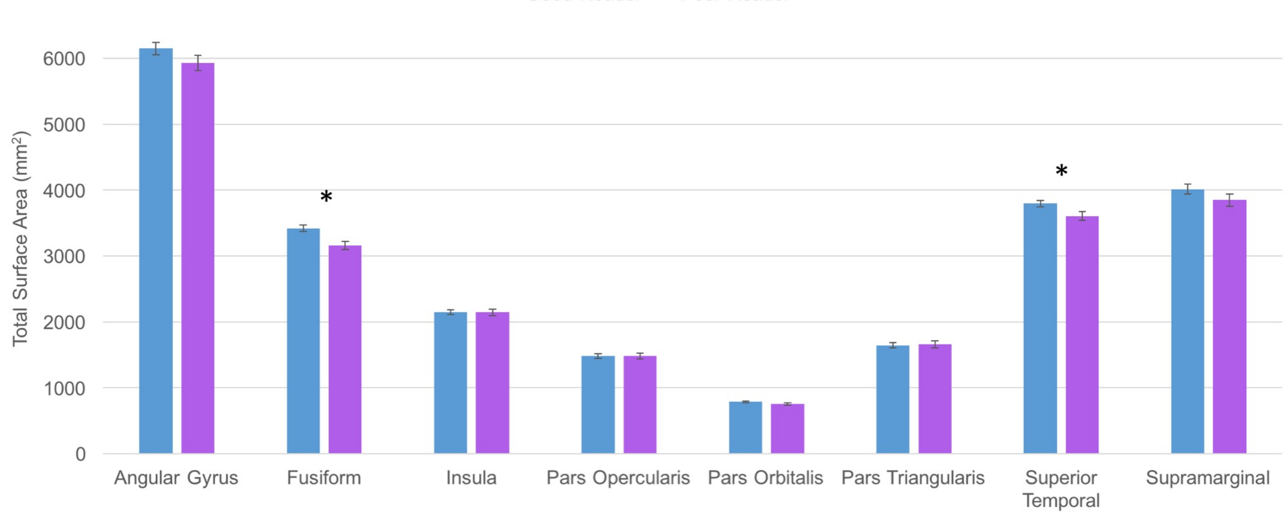
The results showed significant group differences in surface area and cortical thickness, but not in grey matter volume. Specifically, good readers had greater surface area than poor readers in left supramarginal gyrus, $F(1, 51.21) = 4.48$, $p = .039$, right fusiform gyrus, $F(1, 55.89) = 9.74$, $p = .003$, and right superior temporal gyrus, $F(1, 51.71) = 5.49$, $p = .023$ (see Figure 4.3). In contrast, poor readers had thicker cortex than good readers in left fusiform gyrus, $F(1, 55.89) = 7.79$, $p = .007$, left supramarginal gyrus, $F(1, 53.94) = 6.52$, $p = .014$, right pars opercularis, $F(1, 55.56) = 5.53$, $p = .022$, and right supramarginal gyrus, $F(1, 55.86) = 7.40$, $p = .009$ (see Figure 4.4). There were no significant changes from pre-test to post-test or group by time interactions in grey matter volume, surface area, or cortical thickness (all $p > .05$). In terms of the additional variables, total intracranial volume was positively associated with surface area (range in $F = 6.20 – 46.97$, $p < .05$) and grey matter volume (range in $F = 7.89 – 46.61$, $p < .05$), but not cortical thickness (all $p > .05$). Furthermore, gender differences were shown in 3 regions even when controlling for total cranial volume. In comparison to male participants, females had less surface area in right pars orbitalis, $F(1, 55.24) = 4.13$, $p = .047$, thinner cortex in left insula, $F(1, 55.81) = 7.36$, $p = .009$, and thicker cortex in left pars orbitalis, $F(1, 55.58) = 4.83$, $p = .032$.

Figure 4.3. Main effect of group in surface area

a. Left hemisphere



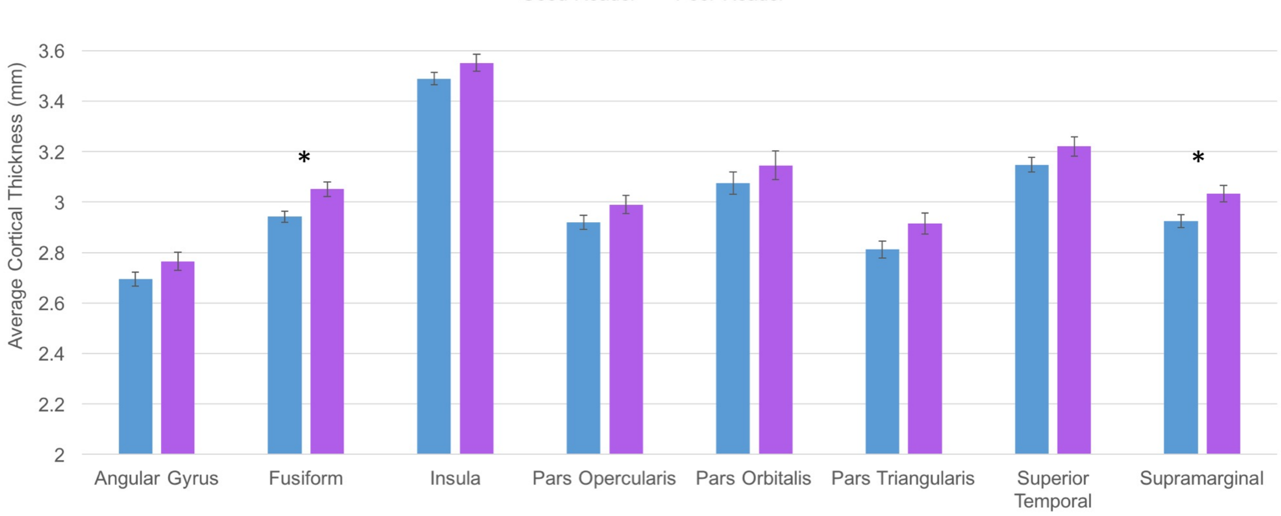
b. Right hemisphere



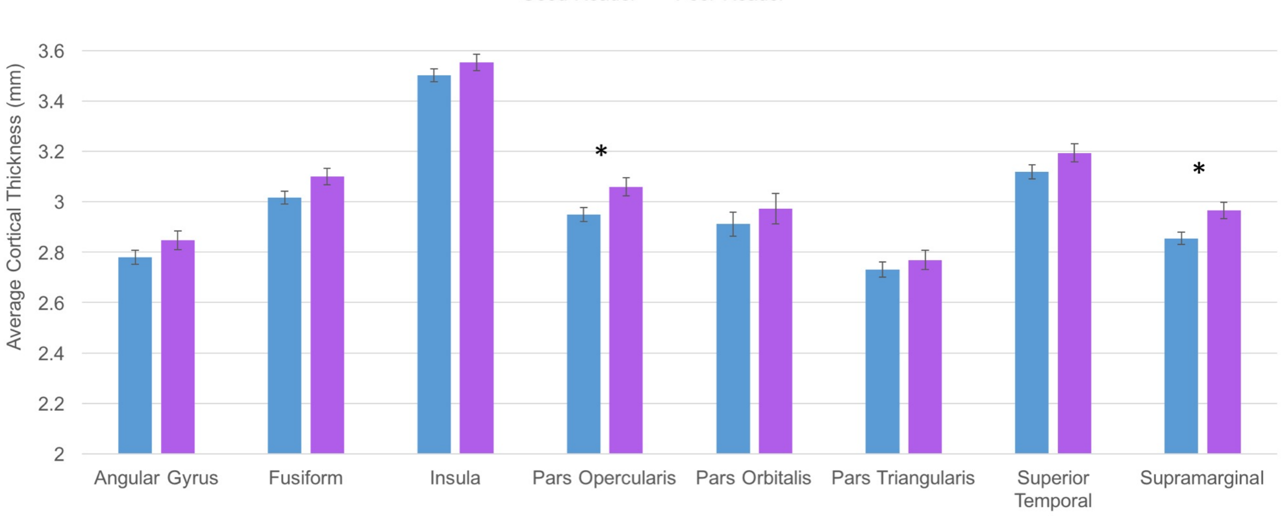
Note: Panel (a) illustrates the results for the left hemisphere, and panel (b) illustrates results for the right hemisphere. An asterisk (*) indicates significant group differences ($p < .05$), and bars denote 1 standard error (SE).

Figure 4.4. Main effect of group in cortical thickness

a. Left hemisphere



b. Right hemisphere



Note: Panel (a) illustrates the results for the left hemisphere, and panel (b) illustrates results for the right hemisphere. An asterisk (*) indicates significant group differences ($p < .05$), and bars denote 1 standard error (SE).

White Matter Structures

Independent *t*-tests showed that there were no differences between groups in the number of diffusion directions at pre-test, $t(30) = .40$, $p = .689$, or post-test, $t(26) = 1.20$, $p = .242$. This result indicates there was a similar amount of diffusion parameters for both groups and thus any group differences cannot be attributable to motion correction methods (i.e., removal of directions with large motion artifacts).

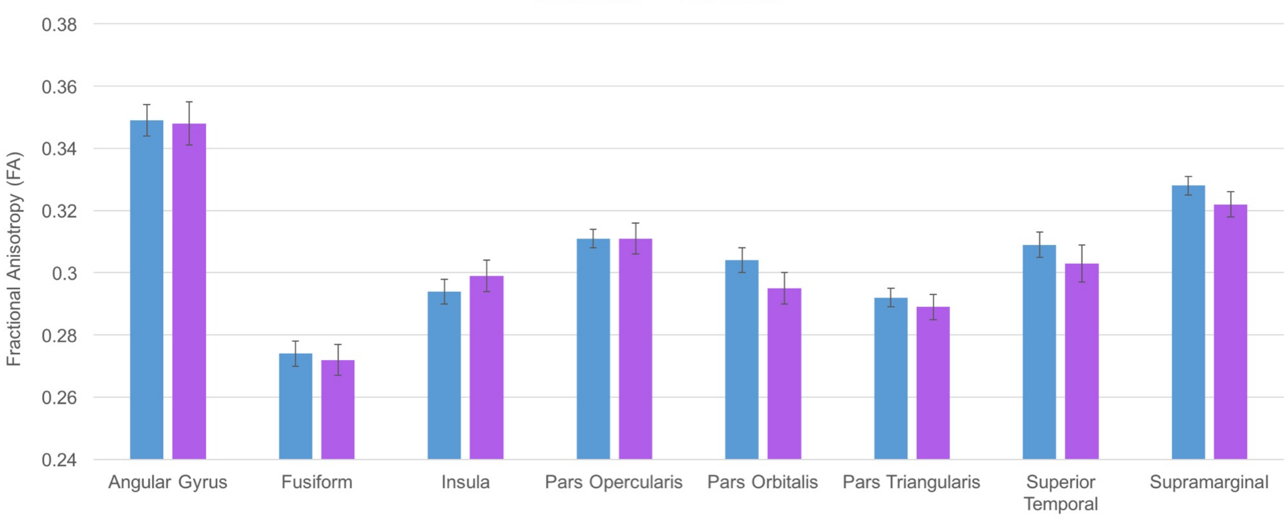
Similarly to above, linear mixed models were conducted with time as a repeated effect, group, time, and gender as fixed effects factors, and total intracranial volume as a fixed effect covariate. Fractional anisotropy and mean diffusivity within each ROI path were evaluated in separate models. Results are reported with an uncorrected *p*-value due to small sample size. The results showed that good readers had significantly higher fractional anisotropy compared to poor readers in the paths from right pars orbitalis, *F*(1, 50.74) = 4.81, *p* = .033 (see Figure 4.5). In contrast, poor readers had greater mean diffusivity than good readers in 10 of the 16 paths (see Figure 4.6), including bilateral insular cortex (left: *F*(1, 51.57) = 5.41, *p* = .024; right: *F*(1, 51.94) = 8.88, *p* = .004), bilateral pars opercularis (left: *F*(1, 48.89) = 4.53, *p* = .038; right: *F*(1, 49.64) = 4.95, *p* = .031), bilateral pars orbitalis (left: *F*(1, 51.54) = 4.80, *p* = .033; right: *F*(1, 43.62) = 5.13, *p* = .029), bilateral supramarginal gyrus (left: *F*(1, 50.80) = 4.18, *p* = .046; right: *F*(1, 50.87) = 6.18, *p* = .016), left pars triangularis, *F*(1, 46.95) = 5.06, *p* = .029, and left superior temporal gyrus, *F*(1, 51.07) = 6.83, *p* = .012.

A main effect of time was shown in mean diffusivity only, such that there was decreased diffusivity from pre-test to post-test in 8 of the paths (see Figure 4.7). These paths included bilateral insula (left: *F*(1, 51.34) = 5.52, *p* = .023; right: *F*(1, 51.97) = 6.30, *p* = .015), bilateral pars triangularis (left: *F*(1, 45.81) = 5.40, *p* = .025; right: *F*(1, 50.78) = 8.53, *p* = .005), right pars opercularis, *F*(1, 48.93) = 10.14, *p* = .003, right pars orbitalis, *F*(1, 42.40) = 4.62, *p* = .037, right superior temporal gyrus, *F*(1, 44.76) = 11.83, *p* = .001, and right supramarginal gyrus, *F*(1, 50.91) = 4.32, *p* = .043. Finally, there was one region that showed a significant group by time effect: the left pars orbitalis. This result was driven by increased fractional anisotropy for the good readers from pre-test to post-test in this region, *F*(1, 51.97) = 5.45, *p* = .023. There were no changes over time for the poor readers (*p* > .05).

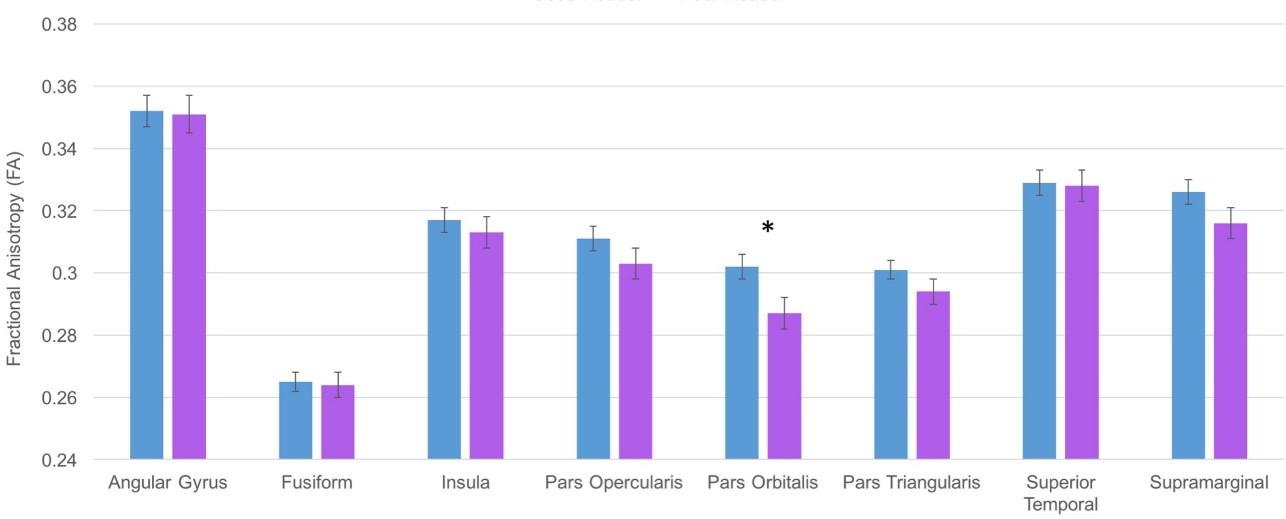
There were also effects of gender and total intracranial volume on fractional anisotropy and mean diffusivity. In comparison to males, female participants had lower fractional anisotropy in the paths from right fusiform gyrus, *F*(1, 50.90) = 5.46, *p* = .023, and right supramarginal gyrus, *F*(1, 48.16) = 4.44, *p* = .040. Females also had increased mean diffusivity in left fusiform paths compared to males, *F*(1, 49.90) = 6.00, *p* = .018. Finally, larger brain volumes were associated with increased fractional anisotropy (range in *F* = 4.36 – 10.31, *p* < .05), and increased mean diffusivity (range in *F* = 4.12 – 8.06, *p* < .05).

Figure 4.5. Main effect of group in fractional anisotropy

a. Left hemisphere



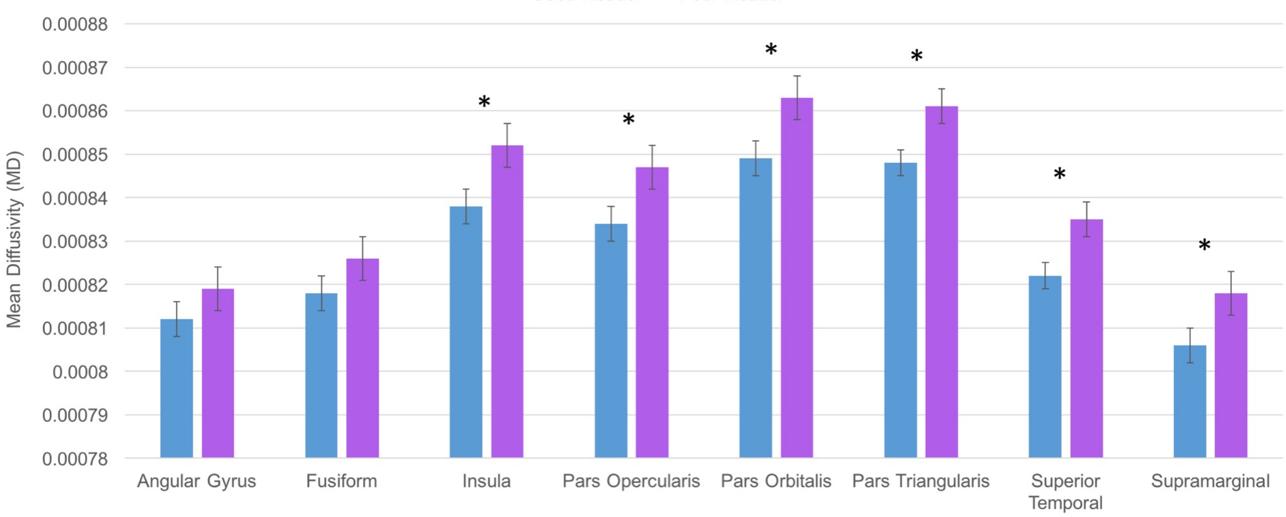
b. Right hemisphere



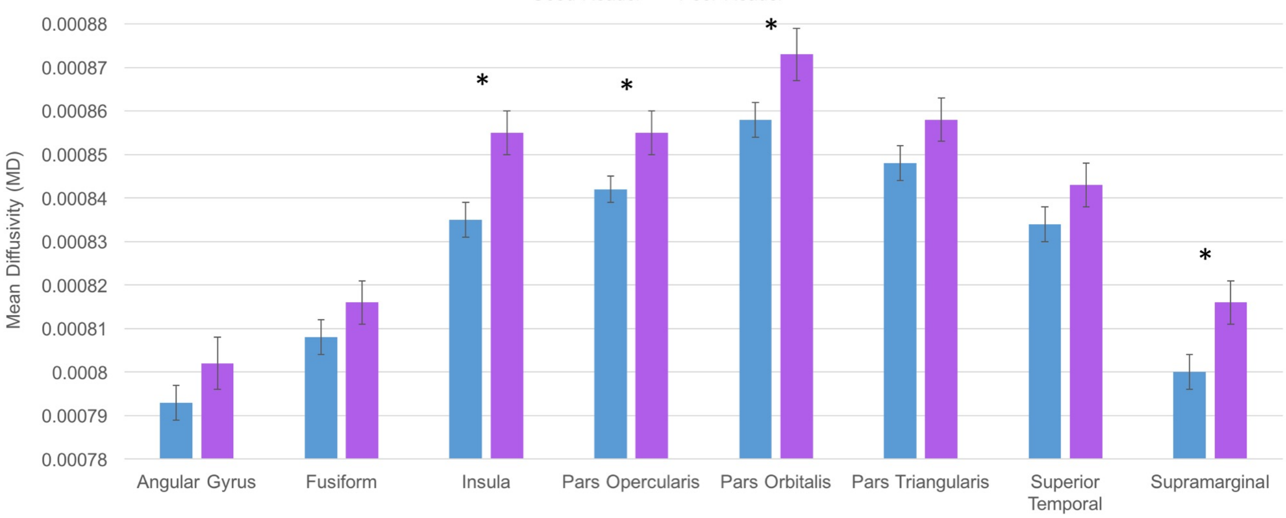
Note: Panel (a) illustrates the results for the left hemisphere, and panel (b) illustrates results for the right hemisphere. An asterisk (*) indicates significant group differences ($p < .05$), and bars denote 1 standard error (SE).

Figure 4.6. Main effect of group in mean diffusivity

a. Left hemisphere



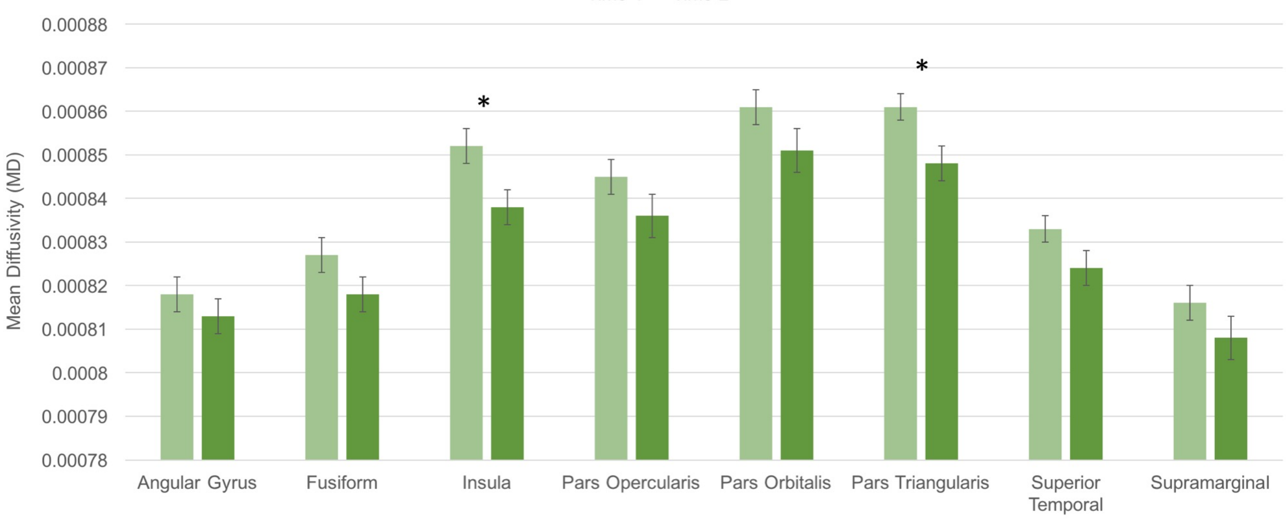
b. Right hemisphere



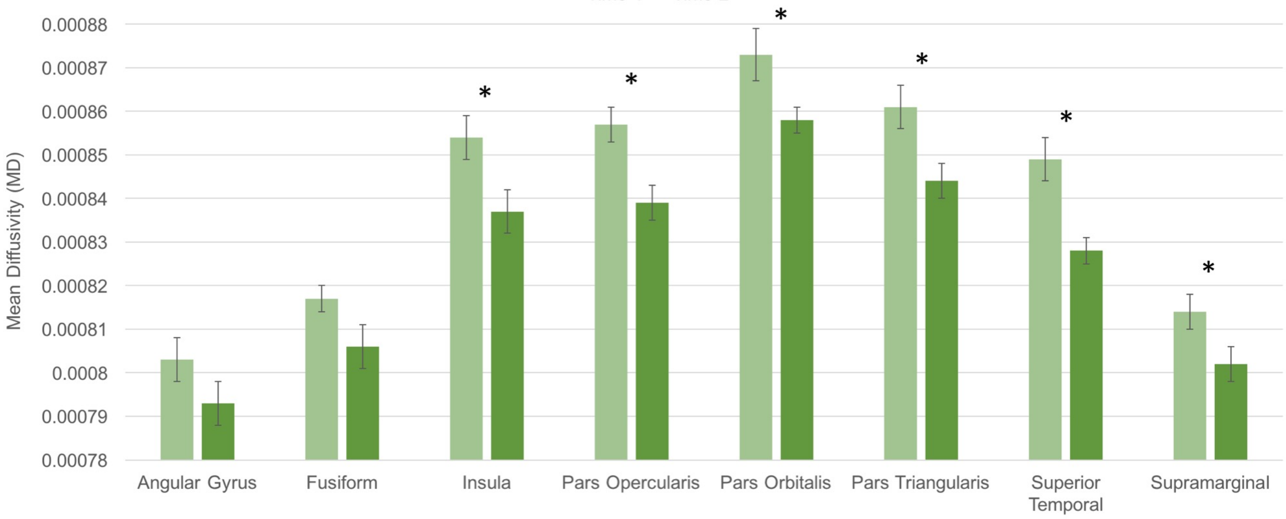
Note: Panel (a) illustrates the results for the left hemisphere, and panel (b) illustrates results for the right hemisphere. An asterisk (*) indicates significant group differences ($p < .05$), and bars denote 1 standard error (SE).

Figure 4.7. Change from pre-test to post-test in mean diffusivity

a. Left hemisphere



b. Right hemisphere



Note: Panel (a) illustrates the results for the left hemisphere, and panel (b) illustrates results for the right hemisphere. An asterisk (*) indicates significant change from time 1 to time 2 (*p* < .05), and bars denote 1 standard error (SE).

Grey and White Matter Predictors of Improved Reading

In the fMRI study in Chapter 3, the behavioural results showed that both good and poor readers improved in word recognition fluency from pre-test to post-test. The poor readers, however, also showed increased decoding skills over time. In order to evaluate the predictors of improved reading skills, the same method from Chapter 3 was used in the current analysis. This included calculating change scores from pre-test to post-test on word recognition fluency and decoding tasks, which were then included as dependent variables in linear regression analyses. The grey and white matter indices that showed group differences were included as predictors, and the left and right hemispheres were evaluated in separate models due to small sample size.

The results showed that white matter in the right hemisphere was associated with gains in word recognition fluency, adjusted $R^2 = .42$, $F(5, 19) = 4.49$, $p = .007$. Specifically, increased mean diffusivity in right insular paths ($\beta = 0.50$) and decreased mean diffusivity in right pars opercularis paths ($\beta = -0.64$) were related to improved word recognition fluency over time. White matter was not associated with gains in decoding skills (left: adjusted $R^2 = .00$, $F(6, 19) < 1$, $p = .619$; right: adjusted $R^2 = .12$, $F(5, 19) = 1.66$, $p = .193$), and grey matter measures were not a significant predictor of improved word recognition fluency (left: adjusted $R^2 = .00$, $F(3, 23) < 1$, $p = .879$; right: adjusted $R^2 = .00$, $F(4, 22) < 1$, $p = .995$) or decoding (left: adjusted $R^2 = .00$, $F(3, 23) < 1$, $p = .831$; right: adjusted $R^2 = .00$, $F(4, 22) < 1$, $p = .560$).

Discussion

Group Differences in Grey and White Matter Structures

Consistent with previous research, the current study found group differences in bilateral grey and white matter regions. Children with good reading skills had larger cortical surface area than poor readers in left supramarginal gyrus, right fusiform, and right superior temporal gyrus, whereas poor readers had thicker cortex than good readers in left fusiform, left supramarginal gyrus, right pars opercularis, and right supramarginal gyrus. Similar findings have been shown, such that typical readers had larger surface area in bilateral inferior frontal and fusiform gyri when compared to adults with a history of reading impairments (Frye et al., 2010). Furthermore, increased cortical thickness for poor readers has been shown in left fusiform (Ma et al., 2015), right superior temporal gyrus (Ma et al., 2015), and right supramarginal gyrus (Frye et al., 2010). Therefore, children with reading impairments have a different pattern of grey matter development when compared to similarly-aged typical readers, and these differences have been shown in bilateral regions of the brain that implicate both orthographic and phonological reading networks.

In contrast, other studies have shown that poor readers have a thinner cortex in left occipital-temporal cortex (Altarelli et al., 2013) and in left orbitofrontal, superior temporal, and middle temporal regions (Clark et al., 2014). The reason for conflicting results could be due to various factors such as age, remediation status, or analytical method. Therefore, recent studies have used multivariate approaches to address these concerns. The results from these studies showed differences between good and poor readers in surface area and cortical network properties (Qi et al., 2016), whereas classification into typical and reading-disabled groups was based on cortical folding and curvature of the cortex rather than volume, area, or thickness (Płoński et al., 2016). The reason for grey matter differences in poor readers continues to be an area of debate, such as differences that occur in neuronal migration and axonal growth (Galaburda, LoTurco, Ramus, Fitch, & Rosen, 2006).

The current study showed no differences between poor and good readers in grey matter volume, which is consistent with some studies (Ma et al., 2015) but inconsistent with others (meta-analysis by Richlan et al., 2012). This result could be due to several reasons, such as small sample size, inclusion of covariates such as gender and total intracranial volume, or the fact that grey matter volume is a combination of surface area and cortical thickness. For example, Eckert and colleagues recently showed that differences between typical readers and reading disabled groups in left frontal and temporal grey matter volume were not significant after correcting for total grey matter volume (Eckert et al., 2016). Other studies have shown that differences between reading groups may be different for boys and girls (Evans et al., 2013), and there were fewer differences when children were matched for reading ability (i.e., poor readers were compared to younger good readers; Krafnick, Flowers, Luetje, Napoliello, & Eden, 2014). The current results showed significant associations between total cranial volume, grey matter volume, and surface area within ROIs, and there were some differences between males and females. It is possible that the participants in this study differed in the components that make up grey matter volume (surface area and cortical thickness) rather than grey matter volume because of relatively low power to detect these differences. Altogether, the results showed that poor readers had differences in multiple regions implicated in reading and language development, and this was shown in surface area and cortical thickness only.

In white matter structures, good readers had higher fractional anisotropy than poor readers in paths from right pars orbitalis. Furthermore, poor readers had greater mean diffusivity than good readers in paths from bilateral insula, bilateral inferior frontal gyrus, bilateral supramarginal gyrus, and left superior temporal gyrus. These paths do not directly correspond to white matter tracts because of the methods of probabilistic tractography. However, visual inspection indicates that dorsal/phonological white matter tracts are most likely involved, including superior longitudinal fasciculus, arcuate fasciculus, and corona radiata. Lower fractional anisotropy and higher mean diffusivity for the poor readers suggest axonal degeneration or demyelination, however, it could also be related to less coherent alignment or crossing fibres within the white matter paths (Jones, Knösche, & Turner, 2013). Therefore, the current results can only suggest that there are differences between good and poor readers in white matter microstructure, which is consistent with previous studies that used voxel-wise or tractography approaches (Deutsch et al., 2005; Klingberg et al., 2000; Odegard, Farris, Ring, McColl, & Black, 2009; Vandermosten et al., 2012a). There were also effects of total intracranial volume and gender in white matter microstructure, such that larger brain volumes were associated with increased fractional anisotropy and diffusivity and females had lower fractional anisotropy and higher diffusivity in several regions. The current results provide further evidence that these factors should be accounted for in analyses involving white matter development.

No Effect of Intervention in Grey and White Matter Structures

This study included the same group of participants from Chapter 3, and analyses from Chapter 3 indicated positive effects of intervention in word recognition fluency and decoding skills. The analyses in this study showed that mean diffusivity decreased over time in bilateral white matter regions, which is consistent with the typical developmental trajectory of decreased diffusivity and increased anisotropy with age (Qiu et al., 2008; Tamnes et al., 2010). There was one group by time interaction, which was driven by good readers having increased fractional anisotropy in left pars orbitalis paths over time. These results indicate that white matter microstructure changed over time, however, there were no effects of reading intervention in white or grey matter measures. These results are consistent with Ma and colleagues (2015) who showed that children with a history of reading impairments had grey matter abnormalities even if they had remediated reading and spelling skills. Furthermore, in the current study, there was a large gap in reading skills between the good and poor readers, indicating that further intervention would be needed. It is possible that 3 months of instruction was not long enough to demonstrate significant changes in grey and white matter structures.

In contrast, other studies that included a prospective design showed changes in grey matter volume in several regions (Krafnick et al., 2011) and white matter in left anterior frontal cortex (Keller & Just, 2009) after reading intervention. These studies included children of similar ages and program length to the current study, and the amount of improvement in reading skills after treatment was also comparable. There were, however, several methodological differences, including type of intervention program and analytical approach and thus it is unknown if these factors impacted the results. Also, it is possible that more intensive programming leads to changes in grey and white matter (i.e., intensive program versus small group supports), but this could not be ascertained in this study because of small sample size. Further prospective studies that include measures of both grey and white matter structures are needed to answer this question. Overall, the current results suggest that deficits in grey matter surface area and cortical thickness as well as in white matter fractional anisotropy and mean diffusivity exist for poor readers, possibly because these students continued to have significant reading impairments following intervention.

White Matter is Associated with Improved Reading Skills

Although there were no changes in grey and white matter following intervention, white matter diffusivity within the right hemisphere was related to improved word recognition fluency over time. This result is similar to Hoeft et al. (2011) who showed that increased fractional anisotropy in right superior longitudinal fasciculus was related to gains in single word reading over time. Furthermore, the relationship between diffusivity and gains in reading fluency differed according to the region: greater diffusivity in right insular paths and less diffusivity in right pars opercularis paths were associated with improved reading. Both of these paths involve the superior longitudinal fasciculus or corona radiata, which suggests that there are differences along the path that cannot be captured by averaging. For example, one group of researchers have shown that white matter microstructure differs along the length of tracts and this may vary according to a child's development (Johnson et al., 2014). Therefore, more specific methods that segment white matter tracts (e.g., Yeatman, Dougherty, Myall, Wandell, & Feldman, 2012) may be able to determine the specific location of group differences and it may be more sensitive to intervention effects. Altogether, the current study provides further evidence that white matter microstructure in the right hemisphere is associated with reading gains over time, and it may be localized to the superior longitudinal fasciculus and corona radiata.

Limitations

The sample size in the current study was relatively small to examine sub-groups and therefore these analyses were not possible (e.g., low versus adequate response to treatment, small group versus intensive supports). Furthermore, an ROI approach was used to examine areas that have been implicated in reading development and thus other grey or white matter regions may have shown changes following intervention. For example, the studies by Keller and Just (2009) and Krafnick et al. (2011) showed changes in frontal white matter and fusiform, precuneus, hippocampus, and cerebellar grey matter. Therefore, larger prospective studies that measure both grey and white matter indices may be able to determine the effects of intervention. Also, a multivariate approach that combines various methods may be an appropriate way to answer these questions (e.g., Hoeft et al., 2011; Płoński et al., 2016).

Conclusions

In conclusion, this study replicates previous studies and demonstrates that there are grey and white matter differences between typically developing readers and children with reading impairments. These differences existed before and after reading intervention, suggesting that poor readers continued to have deficits in reading skills as well as in grey and white matter structures when compared to similarly-aged typical readers. Furthermore, white matter organization in the right hemisphere at the first time point was associated with gains in word recognition fluency, which provides further evidence that neuroimaging measures are predictive of later reading ability.

CHAPTER 5. GENERAL DISCUSSION

Academic and Neuroimaging Outcomes of Reading Intervention

The primary goal of this dissertation was to examine the academic and neuroimaging outcomes of school-based reading interventions. I used a prospective and longitudinal design and evaluated the immediate and one-year results of an intensive reading program (Chapter 2) as well as the effects of intervention on brain function (Chapter 3) and brain structure (Chapter 4). I hypothesized that group differences would be shown across behavioural and neuroimaging measures, and that changes after treatment would be shown on measures of literacy, phonological awareness, functional brain activity during orthographic and phonological reading tasks, grey matter structure, and white matter microstructure. I further hypothesized that baseline literacy and cognitive scores as well as measures of brain function and structure would be associated with reading outcome.

Group Differences

The results in this dissertation showed some support for my hypotheses. Firstly, group differences were shown on behavioural tasks of literacy, language, and cognition. Students with reading impairments had poorer word recognition, word decoding, fluency, reading comprehension, spelling, listening comprehension, memory span, working memory, rapid naming, and processing speed when compared to good readers. There were no group differences on a measure of phonological awareness, which could have been related to task reliability or the fact that students had several years of phonological training in their school district. Overall, children who had difficulty learning to read had consistently weaker performance on literacy and literacy-related measures, which is consistent and replicates previous studies (reviewed in Norton & Wolf, 2012; Shaywitz, 1998; Siegel, 2003).

Furthermore, children with poor reading skills had differences in brain activity in comparison to typically developing readers. Using whole-brain and region of interest analyses, there was less activity for the poor readers in left parietal-temporal cortex during a fMRI rhyming task as well as in left fusiform, left pars triangularis, left superior temporal, and right fusiform on a fMRI spelling task. In contrast, poor readers had greater activity than good readers in right pars opercularis and right supramarginal gyrus on the rhyming task. These results are similar to the literature, which has shown patterns of decreased activation in left hemisphere regions in addition to the recruitment of the right hemisphere (Shaywitz et al., 2003). On the spelling task, significant effects were shown in the region of interest analysis only, which indicates that there was low power to detect differences in the whole-brain analysis.

There were also group differences found in grey matter and white matter microstructure. This effect was specifically shown in left fusiform, left supramarginal, right pars opercularis, and right supramarginal gyri for cortical thickness and in left supramarginal, right fusiform, and right superior temporal gyri for surface area. In white matter paths, good readers had higher fractional anisotropy in right pars orbitalis than poor readers, whereas poor readers had greater mean diffusivity in bilateral inferior frontal, bilateral insula, bilateral supramarginal, and left superior temporal gyri. Previous studies have similarly shown that children and adults with reading impairments have differences in cortical thickness (Ma et al., 2015), surface area (Frye et al., 2010), fractional anisotropy (Deutsch et al., 2005; Klingberg et al., 2000), and diffusivity (Fernandez et al., 2016; Vandermosten et al., 2012a), indicating that both grey and white matter structures are different for this population. However, there were no group differences in grey matter volume, as would be expected based on the literature (Richlan et al., 2012). This lack of a difference may be because grey matter volume is composed of surface area and thickness, and therefore these measures may be more sensitive to group differences.

Effects of Intervention

The studies in this dissertation showed that there were gains in word recognition and decoding fluency immediately after an intensive reading program and one year later. Furthermore, these students boosted their reading skills in word recognition and word recognition fluency so that they were similar to other students who were receiving small group supports. These gains after intervention are similar to previous studies of intensive and individualized programs, which were delivered for 30-45 minutes per day (Denton et al., 2013; O'Connor et al., 2005). Expanding on this previous research, the current study examined an all-day reading program that occurred for 3.75 hours per day over approximately 3 months. Effect sizes in the current study were in the medium to large range, which is similar to other studies, suggesting that an all-day program may not lead to greater improvement in reading compared to a less extensive daily treatment. This claim, however, needs empirical review through studies that specifically evaluate the effect of time in individualized and intensive programs.

The current results also showed that despite individualized and intensive interventions, poor readers continued to have much weaker reading skills in comparison to the good readers in their class. This is consistent with previous studies that have similarly shown a large gap between students who received intervention and those with typical reading skills (Simmons et al., 2008; Torgesen et al., 2001; Vaughn et al., 2009). Longitudinal studies have found that there is a group of persistently poor readers over time (e.g., Partanen & Siegel, 2014) and thus ongoing supports are needed. However, these children may never “catch up” to their peers and thus it will be important to emphasize functional reading skills so that these youths can attain a basic reading level. Furthermore, intervention effects are smaller for students in the upper elementary or secondary school grades in comparison to the primary grades (Vaughn & Wanzek, 2014). Therefore, it will be important to follow and evaluate the progress of these students over many years to determine the types of supports that will be beneficial for their academic achievement.

In the neuroimaging studies, there were changes in functional brain activity on a phonological rhyming task following intervention. In comparison to a group of typically developing readers, the students who received intervention had increased activity from pre-test to post-test in left inferior frontal, bilateral insula, and right parietal regions. This result shows that there were functional changes associated with reading remediation, and suggests that poor readers used different brain regions to read words following intervention. These results are consistent with previous imaging studies of intensive intervention (Eden et al., 2004; Gaab et al., 2007; Simos et al., 2002; Temple et al., 2003), and provides further support for the neurobiological changes that occur with intervention. The unique contribution of these studies is that functional changes in the brain also occurred for programs implemented by teachers in schools.

There were, however, no changes in brain activity during a spelling fMRI task or in brain structure after treatment. This result may be because poor readers had greater improvement in phonological decoding in comparison to spelling skills outside of the MRI scanner, and therefore activation changes on the fMRI spelling task may not be expected. Therefore, further emphasis on spelling skills may be required with future implementations of this intensive program. The reason for minimal changes in grey and white matter structures may be because poor readers continued to have significant reading impairments after intervention when compared to good readers. It may also be related to differences in intervention program or analytical method in comparison to previous studies that showed changes after treatment (Keller & Just, 2009; Krafnick et al., 2011). Furthermore, due to small sample size, the current neuroimaging studies combined students from small group and intensive programs. Larger samples may be able to detect differences between these types of programs over time. An inspection of individual data, however, indicates similar results for both groups of students. Overall, the current studies indicate that there were changes in functional brain activity, but not in brain structure, following 3 months of school-based reading intervention.

Predictors of Reading Outcome

The final objective of this dissertation was to determine the correlates or predictors of improved reading skills. Performance on measures of reading and spelling before intervention was related to classification of persistent reading deficits after intervention. The literacy tasks had better classification rates (84.6%) than cognitive tasks (70.7%), which indicates that baseline literacy skills have better predictive value than the other variables measured. Furthermore, parent ratings of ADHD symptoms were not associated with group classification. These results are consistent with other research groups that have shown that baseline reading scores can predict outcome more so than cognitive variables (Stuebing et al., 2015). The current study indicates similar results in a sample of students who were completing reading programs at their school.

There were also pre-test neuroimaging predictors of change in reading skills over time. The results showed that functional activity in the left hemisphere during a rhyming task was associated with gains in decoding skills, whereas white matter organization in the right hemisphere was related to improved word recognition fluency. Previous studies have similarly shown the predictive value of functional and structural neuroimaging (Hoeft et al., 2011). Therefore, prior to intervention, these results imply that brain function and structure may be able to identify those students who will show greater versus lower gains in reading over time. However, it is unlikely that neuroimaging will be used in the identification of students who will be "responders" to intervention. Rather, this dissertation and other studies illustrate that there are neurobiological bases of reading outcome.

Implications for Education and Future Research

Overall, this dissertation illustrated the importance of an intensive reading program that was delivered to third grade students. However, their reading skills continued to be much below other students who had typical reading ability, and students who received small group supports had minimal gains over time. These results indicate that adaptations to the intensive program and the service delivery model may be needed. For example, smaller student-to-teacher ratios and more time in the intensive program may lead to greater gains in reading (Fuchs, Fuchs, & Vaughn, 2014; Vaughn & Wanzek, 2014). Furthermore, interventions need to start early in Kindergarten or Grade 1 and intensive treatment may need to extend for years, and not months or weeks (Denton, 2012). These types of programs are expensive, but not as expensive as the long-term effects of reading disability including school drop-out, lower educational and vocational attainments, and increased risk for psychiatric disorders (Daniel et al., 2006; Undheim, 2003).

The response-to-intervention (RTI) model has elicited much empirical study, and it is possible that accelerated entry into intensive tier 3 programs may be beneficial for the lowest-performing readers (Compton et al., 2012). A well-documented process for how students move between tiers and frequent progress monitoring will ensure students receive timely intervention (Gersten et al., 2009). For any school district, it would be important to document any programmatic changes and to assess fidelity of intervention implementation. Furthermore, as these were individualized programs, the strategies that worked for each student should be shared with their home school to promote continuation of services.

The results from the neuroimaging studies indicated that intervention changes in the brain were illustrated on a phonological rhyming task only. As the target of these interventions included both phonological and orthographic skills, this result implies that further emphasis on orthographic awareness is needed (e.g., sight words, spelling). Similarly, there were no changes in grey or white matter organization after intervention and perhaps more time or greater intensity would lead to these structural changes. Further studies that specifically examine the effects of length and intensity are needed to answer this question. Also, it will be important to examine children as they progress through tiers of intervention to identify if there are greater changes in brain function and structure at different phases of the RTI model.

An important distinction in treatment programs is whether the child responded to the treatment or not, and also whether there are predictors of success prior to beginning the program. Several studies have shown that children who do not respond to reading intervention have different brain activation patterns (Davis et al., 2011; Odegard et al., 2008) and white matter connectivity (Davis et al., 2010) in comparison to those who showed improved readings scores. One longitudinal study showed that neuroimaging measures were better predictors of treatment success than were behavioural measures (Hoeft et al., 2011), which is promising and supports the hypothesis that neuroimaging measures may be used as indicators of reading improvement. The current study showed that both academic and neuroimaging measures were associated with reading outcomes and provides further support that baseline assessments, including neuroimaging, are predictive of later reading skills.

Strengths and Limitations

The primary strength of this dissertation is that I used a prospective and longitudinal design to assess the outcome of school-based reading programs. This type of design can answer questions about change within the same student, whereas a retrospective design would determine differences but in a separate sample of participants over time. Students were also recruited from all elementary schools in one school district and therefore this could be considered a population-based recruitment and may have less sampling bias. For example, if children were recruited from a clinic that specializes in learning disabilities, then the sample is already restricted to families who can afford to receive these services. However, as parents and students still had to opt into the research, there may be differences between those who decided to participate in the research versus those who did not. Finally, the studies in this dissertation were also unique, such that they were the first to evaluate an all-day reading program in a school context, first to examine effects of both orthographic and phonological reading and modulation of difficulty after reading intervention in functional brain activity, and first to examine multiple grey matter measures and white matter tractography following intervention. Therefore, there are multiple strengths of these studies and they provide further evidence for the outcomes of intensive reading treatment and the neurobiological bases of reading impairment.

There were, however, several limitations of these studies. There was a relatively small sample size for the neuroimaging studies and therefore conclusions are somewhat limited to this sample of students. However, small samples for neuroimaging are not uncommon given the difficulties with recruitment as well as large expense and time needed to collect and analyze the data. Furthermore, the current studies used magnetic resonance imaging (MRI) to evaluate changes in brain function and structure. The advantage of MRI is that it has good spatial resolution, but relatively poorer temporal resolution when compared to methods such as magnetoencephalography (MEG) or electroencephalography (EEG). Changes following intervention may be evident on a finer temporal scale (e.g., Hasko et al., 2014) than what can be measured with functional MRI and therefore future studies should include these types of methods. Finally, random assignment to small group or intensive programs would have enabled further control over the intervention design, and would have ensured that groups were similar at the first time point. However, the purpose of this study was to evaluate student's progress over time according to the type of supports they were already receiving in the school district, which meant that researchers were not involved in the decision-making nor progress monitoring of students. Although a randomized control trial would have greater experimental power than a quasi-experimental design, I examined the real-world implementation of interventions that were delivered by teachers and not researchers. The intensive and individualized programs discussed in the General Introduction included instructors who were part of (O'Connor et al., 2005) or selected, trained, and supervised by (Denton et al., 2013) the research team. Although the studies in this dissertation were not purely experimental, the inclusion of teachers as interventionists is a unique and noteworthy contribution to the literature.

General Conclusions

Overall, this dissertation showed the academic and neuroimaging outcomes of reading interventions in schools. There were gains in several reading skills after an intensive reading program, however, these students continued to have poorer reading skills than typical readers in their class. This gap between poor and good readers persisted throughout the third and fourth grades. There were also related changes in functional brain activity after intervention, and predictors of reading outcome included baseline literacy skills, activity in the left hemisphere, and white matter organization in the right hemisphere. These results indicate the need for ongoing supports so that students continue to show improved reading skills. Also, modification to the reading programs may be needed to assist with larger reading gains over time.

REFERENCES

Agnew, J. (2004). Effect of intensive training on auditory processing and reading skills. *Brain and Language*, *88*(1), 21–25. http://doi.org/10.1016/S0093-934X(03)00157-3

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: a review of the literature. *Remedial and Special Education*, *23*(5), 300–316. http://doi.org/10.1177/07419325020230050501

Al Otaiba, S., & Torgesen, J. (2007). Effects from intensive standardized kindergarten and first-grade interventions for the prevention of reading difficulties. In *Handbook of Response to Intervention* (pp. 212–222). Boston, MA: Springer US. http://doi.org/10.1007/978-0-387-49053-3_15

Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To wait in tier 1 or intervene immediately. *Exceptional Children*, *81*(1), 11–27. http://doi.org/10.1177/0014402914532234

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics*, *4*(3), 316–329. http://doi.org/10.1016/j.nurt.2007.05.011

Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., et al. (2014). Planum temporale asymmetry in developmental dyslexia: Revisiting an old question. *Human Brain Mapping*, *35*(12), 5717–5735. http://doi.org/10.1002/hbm.22579

Altarelli, I., Monzalvo, K., Iannuzzi, S., Fluss, J., Billard, C., Ramus, F., & Dehaene-Lambertz, G. (2013). A functionally guided approach to the morphometry of occipitotemporal regions in developmental dyslexia: evidence for differential effects in boys and girls. *Journal of Neuroscience*, *33*(27), 11296–11301. http://doi.org/10.1523/JNEUROSCI.5854-12.2013

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.). Washington, DC: American Psychiatric Publishing.

Ashby, F. G. (2011). Statistical analysis of fMRI data. Cambridge, MA: The MIT Press.

Aylward, E. H., Richards, T. L., Berninger, V. W., Nagy, W. E., Field, K. M., Grimme, A. C., et al. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology*, *61*(2), 212–219.

Backman, J., Bruck, M., Hebert, M., & Seidenberg, M. S. (1984). Acquisition and use of spelling-sound correspondences in reading. *Journal of Experimental Child Psychology*, *38*, 114–133.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. http://doi.org/10.1038/nrn1201

Badian, N. A. (1993). Phonemic awareness, naming, visual symbol processing, and reading. *Reading and Writing*, *5*(1), 87–100. http://doi.org/10.1007/BF01026920

Badian, N. A. (1995). Predicting reading ability over the long term: The changing roles of letter naming, phonological awareness and orthographic processing. *Annals of Dyslexia*, *45*(1), 79–96. http://doi.org/10.1007/BF02648213

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459.

Barwick, M. A., & Siegel, L. S. (1996). Learning difficulties in adolescent clients of a shelter for runaway and homeless street youths. *Journal of Research on Adolescence*, *6*(4), 649–670.

Beaulieu, C., Plewes, C., Paulson, L. A., Roy, D., Snook, L., Concha, L., & Phillips, L. (2005). Imaging brain connectivity in children with diverse reading ability. *NeuroImage*, *25*(4), 1266–1271. http://doi.org/10.1016/j.neuroimage.2004.12.053

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307.

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, *20*(2), 1052–1063. http://doi.org/10.1016/S1053-8119(03)00435-X

Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S., & Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, *34*(1), 144–155. http://doi.org/10.1016/j.neuroimage.2006.09.018

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., et al. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, *50*(5), 1077–1088. http://doi.org/10.1002/mrm.10609

Bell, N. (1997). Seeing stars: Symbol imagery for phonemic awareness, sight words and spelling. Avila Beach, CA: Gander Publishing.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507. http://doi.org/10.1093/biomet/93.3.491

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44–62. http://doi.org/10.1037/0022-0663.98.1.44

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. http://doi.org/10.1093/cercor/bhp055

Bitan, T., Burman, D. D., Chou, T.-L., Lu, D., Cone, N. E., Cao, F., et al. (2007). The interaction between orthographic and phonological information in children: An fMRI study. *Human Brain Mapping*, *28*(9), 880–891. http://doi.org/10.1002/hbm.20313

Bitan, T., Manor, D., Morocz, I. A., & Karni, A. (2005). Effects of alphabeticality, practice and type of instruction on reading an artificial script: An fMRI study. *Cognitive Brain Research*, *25*(1), 90–106. http://doi.org/10.1016/j.cogbrainres.2005.04.014

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Murray, M. S., Munger, K. A., & Vaughn, M. G. (2014). Intensive reading remediation in grade 2 or 3: Are there effects a decade later? *Journal of Educational Psychology*, *106*(1), 46–57. http://doi.org/10.1037/a0033663

Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R., & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing*, *11*(3), 239–273. http://doi.org/10.1023/A:1008050403932

Blackmon, K., Barr, W. B., Kuzniecky, R., DuBois, J., Carlson, C., Quinn, B. T., et al. (2010). Phonetically irregular word pronunciation and cortical thickness in the adult brain. *NeuroImage*, *51*(4), 1453–1458. http://doi.org/10.1016/j.neuroimage.2010.03.028

Bloom, J. S., Garcia-Barrera, M. A., Miller, C. J., Miller, S. R., & Hynd, G. W. (2013). Planum temporale morphology in children with developmental dyslexia. *Neuropsychologia*, *51*(9), 1684–1692. http://doi.org/10.1016/j.neuropsychologia.2013.05.012

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2004). Development of brain mechanisms for processing orthographic and phonologic representations. *Journal of Cognitive Neuroscience*, *16*(7), 1234–1249. http://doi.org/10.1162/0898929041920496

Borowsky, R., Cummine, J., Owen, W. J., Kelland Friesen, C., Shih, F., & Sarty, G. E. (2006). FMRI of ventral and dorsal processing streams in basic reading processes: insular sensitivity to phonology. *Brain Topography, 18,* 233-239.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Brambati, S. M., Termine, C., Ruffino, M., Danna, M., Lanzi, G., Stella, G., et al. (2006). Neuropsychological deficits and neural dysfunction in familial dyslexia. *Brain Research*, *1113*(1), 174–185. http://doi.org/10.1016/j.brainres.2006.06.099

Brandeis, D., Michel, M. C., & Amzica, F. (2009). From neuronal activity to scalp potential fields. In C. M. Michel, T. Koenig, D. Brandeis, L. R. R. Gianotti, & J. Wackermann (Eds.), *Dynamic spatiotemporal brain analyses using high-performance electrical neuroimaging, Part II: A step-by-step tutorial.* (pp. 1–24). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511596889.002

Breznitz, Z., Shaul, S., Horowitz-Kraus, T., Sela, I., Nevat, M., & Karni, A. (2013). Enhanced reading by training with imposed time constraint in typical and dyslexic adults. *Nature Communications*, *4*(1486), 1–6. http://doi.org/10.1038/ncomms2488

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2010). Test of Nonverbal Intelligence (4th ed.). Austin, TX: Pro-Ed.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology*, *28*(5), 874–886.

Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain*, *122*, 1901–1917.

Cao, F., Bitan, T., Chou, T.-L., Burman, D. D., & Booth, J. R. (2006). Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. *Journal of Child Psychology and Psychiatry*, *47*(10), 1041–1050. http://doi.org/10.1111/j.1469-7610.2006.01684.x

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 367–379.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*, 149–180.

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132. http://doi.org/10.1016/j.cortex.2008.05.004

Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics*, *26*(1), 214–238. http://doi.org/10.1016/j.jneuroling.2012.08.001

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading*, *3*(4), 331–361. http://doi.org/10.1207/s1532799xssr0304_2

Cavanaugh, C. L., Kim, A., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research, *2*(1), 9–21.

Center, Y., Wheldall, K., & Freeman, L. (1992). Evaluating the effectiveness of Reading Recovery: A critique. *Educational Psychology*, *12*(3-4), 263–274. http://doi.org/10.1080/0144341920120309

Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2001). Does success in the Reading Recovery program depend on developing proficiency in phonological-processing skills? A longitudinal study in a whole language instructional context. *Scientific Studies of Reading*, *5*(2), 141–176. http://doi.org/10.1207/S1532799Xssr0502_2

Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, *25*(3), 607–623. http://doi.org/10.1016/j.cogbrainres.2005.08.013

Cho, P., & Woo, G. C. (2004). Repeatability of the Waterloo Four-Contrast LogMAR Visual Acuity chart and Near Vision Test card on a group of normal young adults. *Ophthalmic & Physiological Optics*, *24*(5), 427–435. http://doi.org/10.1111/j.1475-1313.2004.00216.x

Christodoulou, J. A., Cyr, A., Murtagh, J., Chang, P., Lin, J., Guarino, A. J., et al. (2017). Impact of intensive summer reading intervention for children with reading disabilities and difficulties in early elementary school. *Journal of Learning Disabilities*, *50*(2), 115–127. http://doi.org/10.1177/0022219415617163

Church, J. A., Coalson, R. S., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2008). A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. *Cerebral Cortex*, *18*(9), 2054–2065. http://doi.org/10.1093/cercor/bhm228

Clark, K. A., Helland, T., Specht, K., Narr, K. L., Manis, F. R., Toga, A. W., & Hugdahl, K. (2014). Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain*, *137*(12), 3136–3141. http://doi.org/10.1093/brain/awu229

Clay, M. M. (1993). Reading Recovery: A guidebook for teachers in training. Auckland, New Zealand: Heinemann Education.

Clay, M. M. (2005). An observation survey of early literacy achievement (2nd ed.). Portsmouth, NH: Heinemann Education.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cole, D. M., Smith, S. M., & Beckmann, C. F. (2010). Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. *Frontiers in Systems Neuroscience*, *4*(8), 1–15. http://doi.org/10.3389/fnsys.2010.00008

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Educational Psychology*, *14*(3), 387–397.

Compton, D. L. (2003). Modeling the relationship between growth in rapid naming speed and growth in decoding skill in first-grade children. *Journal of Educational Psychology*, *95*(2), 225–239. http://doi.org/10.1037/0022-0663.95.2.225

Compton, D. L., Gilbert, J. K., Jenkins, J. R., Fuchs, D., Fuchs, L. S., Cho, E., et al. (2012). Accelerating chronically unresponsive children to tier 3 instruction: What level of data is necessary to ensure selection accuracy? *Journal of Learning Disabilities*, *45*(3), 204–216. http://doi.org/10.1177/0022219412442151

Conners, C. K. (2008). Conners (3rd ed.). North Tonawanda, NY: MHS.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384–387.

Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods*, *40*(3), 791–794. http://doi.org/10.3758/BRM.40.3.791

Curtin, S., Manis, F. R., & Seidenberg, M. S. (2001). Parallels between the reading and spelling deficits of two subgroups of developmental dyslexics. *Reading and Writing*, *14*(5-6), 515–

547. http://doi.org/10.1023/A:1011122219046

D'Agostino, J. V., & Harmey, S. J. (2016). An international meta-analysis of Reading Recovery. *Journal of Education for Students Placed at Risk*, *21*(1), 29–46. http://doi.org/10.1080/10824669.2015.1112746

D'Agostino, J. V., & Murphy, J. A. (2004). A meta-analysis of Reading Recovery in United States schools. *Educational Evaluation and Policy Analysis*, *26*(1), 23–38.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. http://doi.org/10.1006/nimg.1998.0395

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities*, *39*(6), 507–514. http://doi.org/10.1177/00222194060390060301

Davis, N., 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. http://doi.org/10.1080/87565641.2010.549875

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

de Jong, P. F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology*, *91*(3), 450–476.

Denckla, M. B., & Rudel, R. G. (1976). Rapid "automatized" naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*, 471–479.

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. http://doi.org/10.1177/0022219412442155

Denton, C. A., Fletcher, J. M., Anthony, J. L., & Francis, D. J. (2006). An evaluation of intensive intervention for students with persistent reading difficulties. *Journal of Learning Disabilities*, *39*(5), 447–466. http://doi.org/10.1177/00222194060390050601

Denton, C. A., Tolar, T. D., Fletcher, J. M., Barth, A. E., Vaughn, S., & Francis, D. J. (2013). Effects of tier 3 intervention for students with persistent reading difficulties and characteristics of inadequate responders. *Journal of Educational Psychology*, *105*(3), 633–648. http://doi.org/10.1037/a0032581

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. http://doi.org/10.1016/j.neuroimage.2006.01.021

Deutsch, G. K., Dougherty, R. F., Bammer, R., Siok, W. T., Gabrieli, J. D. E., & Wandell, B. (2005). Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. *Cortex*, *41*(3), 354–363. http://doi.org/10.1016/S0010-9452(08)70272-7

Dias, K., & Juniper, L. (2002). Phono-Graphix - who needs additional literacy support? An outline of research in Bristol schools. *Support for Learning*, *17*(1), 34–38. http://doi.org/10.1111/1467-9604.00230

Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. *Brain*, *135*(12), 3529–3550. http://doi.org/10.1093/brain/aws222

Dobson, V., Clifford-Donaldson, C. E., Green, T. K., Miller, J. M., & Harvey, E. M. (2009). Normative monocular visual acuity for early treatment diabetic retinopathy study charts in emmetropic children 5 to 12 years of age. *Ophthalmology*, *116*(7), 1397–1401. http://doi.org/10.1016/j.ophtha.2009.01.019

Eckert, M. A., Berninger, V. W., Vaden, K. I., Gebregziabher, M., Tsu, L., & Dyslexia Data Consortium. (2016). Gray matter features of reading disability: A combined meta-analytic and direct analysis approach. *eNeuro*, *3*(1), 1–15. http://doi.org/10.1523/ENEURO.0103-15.2015

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

Edwards, V. T., & Hogben, J. H. (1999). New norms for comparing children's lexical and nonlexical reading: A further look at subtyping dyslexia. *Australian Journal of Psychology*, *51*(1), 37–49. http://doi.org/10.1080/00049539908255333

Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology*, *92*(4), 605–619. http://doi.org/10.1037//0022-0663.92.4.605

Endress, S. A., Weston, H., Marchand-Martella, N. E., Martella, R. C., & Simmons, J. (2007). Examining the effects of Phono-Graphix on the remediation of reading skills of students with disabilities: A program evaluation. *Education and Treatment of Children*, *30*(2), 1–20. http://doi.org/10.1353/etc.2007.0006

Erickson, J., Derby, K. M., McLaughlin, T. F., & Fuehrer, K. (2015). An evaluation of Read Naturally on increasing reading fluency for three primary students with learning disabilities. *Educational Research Quarterly*, *39*(1), 3–20.

Evans, T. M., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2013). Sex-specific gray matter volume differences in females with developmental dyslexia. *Brain Structure and Function*, *219*(3), 1041–1054. http://doi.org/10.1007/s00429-013-0552-4

Fan, Q., Anderson, A. W., Davis, N., & Cutting, L. E. (2014a). Structural connectivity patterns associated with the putative visual word form area and children′s reading ability. *Brain Research*, *1586*, 118–129. http://doi.org/10.1016/j.brainres.2014.08.050

Fan, Q., Davis, N., Anderson, A. W., & Cutting, L. E. (2014b). Thalamo-cortical connectivity: What can diffusion tractography tell us about reading difficulties in children? *Brain Connectivity*, *4*(6), 428–439. http://doi.org/10.1089/brain.2013.0203

Farris, E. A., Odegard, T. N., Miller, H. L., Ring, J., Allen, G., & Black, J. (2011). Functional connectivity between the left and right inferior frontal lobes in a small sample of children with and without reading difficulties. *Neurocase*, *17*(5), 425–439. http://doi.org/10.1080/13554794.2010.532141

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. http://doi.org/10.1080/87565641.2016.1158264

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. http://doi.org/10.3758/BF03193146

Fernandez, V. G., Juranek, J., Romanowska-Pawliczek, A., Stuebing, K., Williams, V. J., & Fletcher, J. M. (2016). White matter integrity of cerebellar-cortical tracts in reading impaired children: A probabilistic tractography study. *Brain and Language*, *161*(C), 45–56. http://doi.org/10.1016/j.bandl.2015.07.006

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

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050–11055. http://doi.org/10.1073/pnas.200033797

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., & Brain Development Cooperative Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54*(1), 313–327. http://doi.org/10.1016/j.neuroimage.2010.07.033

Foorman, B. R., Francis, D. J., Winikates, D., Mehta, P., Schatschneider, C., & Fletcher, J. M. (1997). Early interventions for children with reading disabilities. *Scientific Studies of Reading*, *1*(3), 255–276. http://doi.org/10.1207/s1532799xssr0103_5

Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *NeuroImage*, *42*(3), 1178–1184. http://doi.org/10.1016/j.neuroimage.2008.05.059

Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Current Opinion in Neurobiology*, *23*(2), 250–254. http://doi.org/10.1016/j.conb.2012.10.002

Frye, R. E., Liederman, J., Malmberg, B., McLean, J., Strickland, D., & Beauchamp, M. S. (2010). Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. *Cerebral Cortex*, *20*(11), 2625–2635. http://doi.org/10.1093/cercor/bhq010

Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014). What is intensive instruction and why is it important? *Teaching Exceptional Children*, *46*(4), 13–18. http://doi.org/10.1177/0040059914522966

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 and Neuroscience*, *25*(3-4), 295–310.

Galaburda, A. M., LoTurco, J., Ramus, F., Fitch, R. H., & Rosen, G. D. (2006). From genes to behavior in developmental dyslexia. *Nature Neuroscience*, *9*(10), 1213–1217. http://doi.org/10.1038/nn1772

Geng, J. J., & Mangun, G. R. (2009). Anterior intraparietal sulcus is sensitive to bottom-up attention driven by stimulus salience. *Journal of Cognitive Neuroscience*, *21*(8), 1584–1601. http://doi.org/10.1162/jocn.2009.21103

Gennatas, E. D., Avants, B. B., Wolf, D. H., Satterthwaite, T. D., Ruparel, K., Ciric, R., et al. (2017). Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. *Journal of Neuroscience*, *37*(20), 5065–5073. http://doi.org/10.1523/JNEUROSCI.3550-16.2017

Georgiou, G. K., Das, J. P., & Hayward, D. V. (2008). Comparing the contribution of two tests of working memory to reading in relation to phonological awareness and rapid naming speed. *Journal of Research in Reading*, *31*(3), 302–318. http://doi.org/10.1111/j.1467-9817.2008.00373.x

Germanò, E., Gagliano, A., & Curatolo, P. (2010). Comorbidity of ADHD and dyslexia. *Developmental Neuropsychology*, *35*(5), 475–493. http://doi.org/10.1080/87565641.2010.494748

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2009). Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide. (pp. 1–60). Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gibson, L., Jr, Cartledge, G., Keyes, S. E., & Yawn, C. D. (2014). The effects of a supplementary computerized fluency intervention on the generalization of the oral reading fluency and comprehension of first-grade students. *Education and Treatment of Children*, *37*(1), 25–51. http://doi.org/10.1353/etc.2014.0003

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013). Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, *48*(2), 135–154. http://doi.org/10.1002/rrq.45

Gillam, R. B., Loeb, D. F., Hoffman, L. M., Bohman, T., Champlin, C. A., Thibodeau, L., et al. (2008). The efficacy of Fast ForWord Language intervention in school-age children with language impairment: a randomized controlled trial. *Journal of Speech, Language, and Hearing Research*, *51*(1), 97–119. http://doi.org/10.1044/1092-4388(2008/007)

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences*, *101*(21), 8174–8179. http://doi.org/10.1073/pnas.0402680101

Good, R. H., & Kaminski, R. A. (2002). Dynamic Indicators of Basic Early Literacy Skills (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Gottardo, A., Stanovich, K. E., & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology, 63*, 563–582.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106,* 491-528.

Hasbrouck, J. E., Ihnot, C., & Rogers, G. H. (1999). "Read naturally": A strategy to increase oral reading fluency. *Reading Research and Instruction*, *39*(1), 27–37. http://doi.org/10.1080/19388079909558310

Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Korne, G. (2014). What does the brain of children with developmental dyslexia tell us about reading improvement? ERP evidence from an intervention study. *Frontiers in Human Neuroscience*, *8*, 1–21. http://doi.org/10.3389/fnhum.2014.00441

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413–497. http://doi.org/10.1103/RevModPhys.65.413

Heikkilä, R., Aro, M., Närhi, V., Westerholm, J., & Ahonen, T. (2013). Does training in syllable recognition improve reading speed? A computer-based trial with poor readers from second and third grade. *Scientific Studies of Reading*, *17*(6), 398–414. http://doi.org/10.1080/10888438.2012.753452

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

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., et al. (2007a). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences*, *104*(10), 4234–4239. http://doi.org/10.1073/pnas.0609399104

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

Horowitz-Kraus, T., & Holland, S. K. (2015). Greater functional connectivity between reading and error-detection regions following training with the reading acceleration program in children with reading difficulties. *Annals of Dyslexia*, *65*(1), 1–23. http://doi.org/10.1007/s11881-015-0096-9

Horowitz-Kraus, T., DiFrancesco, M., Kay, B., Wang, Y., & Holland, S. K. (2015a). Increased resting-state functional connectivity of visual- and cognitive-control brain networks after training in children with reading difficulties. *NeuroImage: Clinical*, *8*(C), 619–630. http://doi.org/10.1016/j.nicl.2015.06.010

Horowitz-Kraus, T., Toro-Serey, C., & DiFrancesco, M. (2015b). Increased resting-state functional connectivity in the cingulo-opercular cognitive-control network after intervention in children with reading difficulties. *PLoS ONE*, *10*(7), e0133762–21. http://doi.org/10.1371/journal.pone.0133762

Horowitz-Kraus, T., Vannest, J. J., & Holland, S. K. (2013). Overlapping neural circuitry for narrative comprehension and proficient reading in children and adolescents. *Neuropsychologia*, *51*(13), 2651–2662. http://doi.org/10.1016/j.neuropsychologia.2013.09.002

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. http://doi.org/10.1002/brb3.281

Horwitz, B., Rumsey, J. M., & Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proceedings of the National Academy of Sciences*, *95*(15), 8939–8944.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156.

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

Jere-Folotiya, J., Chansa-Kabali, T., Munachaka, J. C., Sampa, F., Yalukanda, C., Westerholm, J., et al. (2014). The effect of using a mobile literacy game to improve literacy levels of grade one students in Zambian schools. *Educational Technology Research and Development*, *62*(4), 417–436. http://doi.org/10.1007/s11423-014-9342-9

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *NeuroImage*, *20*(2), 693–712. http://doi.org/10.1016/S1053-8119(03)00343-4

Johns, J. L. (2008). Basic Reading Inventory (10th ed.). Dubuque, IA: Kendall Hunt Publishing Co.

Johnson, R. T., Yeatman, J. D., Wandell, B. A., Buonocore, M. H., Amaral, D. G., & Nordahl, C. W. (2014). Diffusion properties of major white matter tracts in young, typically developing children. *NeuroImage*, *88*(C), 143–154. http://doi.org/10.1016/j.neuroimage.2013.11.025

Johnston, R. S., Thompson, G. B., Fletcher-Flinn, C. M., & Holligan, C. (1995). The functions of phonology in the acquisition of reading: Lexical and sentence processing. *Memory & Cognition*, *23*(6), 749–766.

Jolles, D. D., Grol, M. J., Van Buchem, M. A., Rombouts, S. A. R. B., & Crone, E. A. (2010). Practice effects in the brain: Changes in cerebral activation after working memory practice depend on task demands. *NeuroImage*, *52*(2), 658–668. http://doi.org/10.1016/j.neuroimage.2010.04.028

Jones, D. K. (2004). The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: A Monte Carlo study. *Magnetic Resonance in Medicine*, *51*(4), 807–815. http://doi.org/10.1002/mrm.20033

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, *73*, 239–254. http://doi.org/10.1016/j.neuroimage.2012.06.081

Jongejan, W., Verhoeven, L., & Siegel, L. S. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *Journal of Educational Psychology*, *99*(4), 835–851. http://doi.org/10.1037/0022-0663.99.4.835

Kail, R., Hall, L. K., & Caskey, B. J. (1999). Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics*, *20*, 303–314.

Katusic, S. K., Colligan, R. C., Barbaresi, W. J., Schaid, D. J., & Jacobsen, S. J. (2001). Incidence of reading disability in a population-based birth cohort, 1976-1982, Rochester, Minn. *Mayo Clinic Proceedings*, *76*, 1081–1092.

Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman Test of Educational Achievement (2nd ed.). San Antonio, TX: Pearson.

Keller, T. A., & Just, M. A. (2009). Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. *Neuron*, *64*(5), 624–631. http://doi.org/10.1016/j.neuron.2009.10.018

Kelly, A. M. C., & Garavan, H. (2004). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex*, *15*(8), 1089–1102. http://doi.org/10.1093/cercor/bhi005

Keyes, S. E., Cartledge, G., Gibson, L., Jr, & Robinson-Ervin, P. (2016). Programming for generalization of oral reading fluency using computer-assisted instruction and changing fluency criteria. *Education and Treatment of Children*, *39*(2), 141–172. http://doi.org/10.1353/etc.2016.0011

Kingsley, P. B. (2006). Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors. *Concepts in Magnetic Resonance Part A*, *28A*(2), 101–122. http://doi.org/10.1002/cmr.a.20048

Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, *95*(3), 453–464. http://doi.org/10.1037/0022-0663.95.3.453

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(ECVP Abtract Supplement).

Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J. D. E., Moseley, M. E., & Poldrack, R. A. (2000). Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. *Neuron*, *25*, 493–500.

Kovelman, I., Norton, E. S., Christodoulou, J. A., Gaab, N., Lieberman, D. A., Triantafyllou, C., et al. (2012). Brain basis of phonological awareness for spoken language in children and Its disruption in dyslexia. *Cerebral Cortex*, *22*(4), 754–764. http://doi.org/10.1093/cercor/bhr094

Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., et al. (2013). Cortical signatures of dyslexia and remediation: An intrinsic functional connectivity approach. *PLoS ONE*, *8*(2), e55454–14. http://doi.org/10.1371/journal.pone.0055454

Krafnick, A. J., Flowers, D. L., Luetje, M. M., Napoliello, E. M., & Eden, G. F. (2014). An investigation into the origin of anatomical differences in dyslexia. *Journal of Neuroscience*, *34*(3), 901–908. http://doi.org/10.1523/JNEUROSCI.2092-13.2013

Krafnick, A. J., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *NeuroImage*, *57*(3), 733–741. http://doi.org/10.1016/j.neuroimage.2010.10.062

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly*, *48*(1), 61–76. http://doi.org/10.1002/rrq.038

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150–161. http://doi.org/10.1037/0022-0663.100.1.150

Langer, N., Benjamin, C., Minas, J., & Gaab, N. (2015). The neural correlates of reading fluency deficits in children. *Cerebral Cortex*, *25*(6), 1441–1453. http://doi.org/10.1093/cercor/bht330

Lebel, C., Shaywitz, B., Holahan, J., Shaywitz, S., Marchione, K., & Beaulieu, C. (2013). Diffusion tensor imaging correlates of reading ability in dysfluent and non-impaired readers. *Brain and Language*, *125*(2), 215–222. http://doi.org/10.1016/j.bandl.2012.10.009

Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A., Weinberger, D. R., & Mattay, V. S. (2012). Normal age-related brain morphometric changes: Nonuniformity across cortical thickness, surface area and gray matter volume? *Neurobiology of Aging*, *33*(3), 617.e1–617.e9. http://doi.org/10.1016/j.neurobiolaging.2010.07.013

Lindamood, C. H., & Lindamood, P. C. (1975). Auditory Discrimination in Depth. Allen, TX: DLM Teaching Resources.

Lindamood, P., & Lindamood, P. (1998). The Lindamood Phoneme Sequencing Program for Reading, Spelling, and Speech LiPS. Teacher's Manual for the Classroom and Clinic. Austin, TX: Pro-Ed.

Lipka, O., Lesaux, N. K., & Siegel, L. S. (2006). Retrospective analyses of the reading development of grade 4 students with reading disabilities: Risk status and profiles over 5 years. *Journal of Learning Disabilities*, *39*(4), 364–378.

http://doi.org/10.1177/00222194060390040901

Lovio, R., Halttunen, A., Lyytinen, H., Näätänen, R., & Kujala, T. (2012). Reading skill and neural processing accuracy improvement after a 3-hour intervention in preschoolers with difficulties in reading-related skills. *Brain Research*, *1448*, 42–55. http://doi.org/10.1016/j.brainres.2012.01.071

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). Defining dyslexia, comorbidity, teachers knowledge of language and reading: A definition of dyslexia. *Annals of Dyslexia*, *53*, 1–14.

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology*, *50*(6), 668–675. http://doi.org/10.1111/j.1467-9450.2009.00791.x

Ma, Y., Koyama, M. S., Milham, M. P., Castellanos, F. X., Quinn, B. T., Pardoe, H., et al. (2015). Cortical thickness abnormalities associated with dyslexia, independent of remediation status. *NeuroImage: Clinical*, *7*, 177–186. http://doi.org/10.1016/j.nicl.2014.11.005

MacDonald, G. W., & Cornwall, A. (1995). The relationship between phonological awareness and reading and spelling achievement eleven years later. *Journal of Learning Disabilities*, *28*(8), 523–527. http://doi.org/10.1177/002221949502800807

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. http://doi.org/10.1002/hbm.22749

McCallum, R. S. (2006). What is the role of working memory in reading relative to the big three processing variables (orthography, phonology, and rapid naming)? *Journal of Psychoeducational Assessment*, *24*(3), 243–259. http://doi.org/10.1177/0734282906287938

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299. http://doi.org/10.1016/S1364-6613(03)00134-7

McGuinness, C., McGuinness, D., & McGuinness, G. (1996). Phono-Graphix: A new method for remediating reading difficulties. *Annals of Dyslexia*, *46*(1), 73–96. http://doi.org/10.1007/BF02648172

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, *271*(5245), 77–81. http://doi.org/10.1126/science.271.5245.77

Metsala, J. L., & David, M. D. (2017). The effects of age and sublexical automaticity on reading outcomes for students with reading disabilities. *Journal of Research in Reading*, *39*(4), 212–19. http://doi.org/10.1111/1467-9817.12097

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. http://doi.org/10.1016/j.neuropsychologia.2008.03.012

Meyler, A., Keller, T. A., Cherkassky, V. L., Lee, D., Hoeft, F., Whitfield-Gabrieli, S., et al. (2007). Brain activation during sentence comprehension among good and poor readers. *Cerebral Cortex*, *17*(12), 2780–2787. http://doi.org/10.1093/cercor/bhm006

Mugnaini, D., Lassi, S., La Malfa, G., & Albertini, G. (2009). Internalizing correlates of dyslexia. *World Journal of Pediatrics*, *5*(4), 255–264. http://doi.org/10.1007/s12519-009-0049-7

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665–681. http://doi.org/10.1037/0012-1649.40.5.665

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*(1), 134–147. http://doi.org/10.1037/0022-0663.98.1.134

Nagy, Z., Westerberg, H., & Klingberg, T. (2004). Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience*, *16*(7), 1227–1233. http://doi.org/10.1162/0898929041920441

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*, 85–101.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implication for reading instruction. Washington, D.C.: U.S. Government Printing Office.

Niedo, J., Lee, Y.-L., Breznitz, Z., & Berninger, V. W. (2014). Computerized silent reading rate and strategy instruction for fourth graders at risk in silent reading rate. *Learning Disability Quarterly*, *37*(2), 100–110. http://doi.org/10.1177/0731948713507263

North Vancouver School District. (1999). Reading 44: A Core Reading Framework. North Vancouver, BC: School District 44.

North Vancouver School District. (2001). Firm Foundations: Early literacy teaching and training. North Vancouver, BC: School District 44.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology*, *63*(1), 427–452. http://doi.org/10.1146/annurev-psych-120710-100431

O'Connor, R. E., Harty, K. R., & Fulmer, D. (2005). Tiers of intervention in kindergarten through third grade. *Journal of Learning Disabilities*, *38*(6), 532–538. http://doi.org/10.1177/00222194050380060901

Odegard, T. N., Farris, E. A., Ring, J., McColl, R., & Black, J. (2009). Brain connectivity in non-reading impaired children and children diagnosed with developmental dyslexia. *Neuropsychologia*, *47*(8-9), 1972–1977. http://doi.org/10.1016/j.neuropsychologia.2009.03.009

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. http://doi.org/10.1007/s11881-008-0014-5

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

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. http://doi.org/10.1016/j.neuron.2013.05.002

Partanen, M., & Siegel, L. S. (2014). Long-term outcome of the early identification and intervention of reading disabilities. *Reading and Writing*, *27*(4), 665–684. http://doi.org/10.1007/s11145-013-9472-1

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pierpaoli, C., Walker, L., Irfanoglu, M. O., Barnett, A., Basser, P., Chang, L.-C., et al. (2010). TORTOISE: An integrated software package for processing of diffusion MRI data. Paper presented at the International Society for Magnetic Resonance in Medicine, Stockholm, Sweden.

Pokorni, J. L., Worthington, C. K., & Jamison, P. J. (2004). Phonological awareness intervention: Comparison of Fast ForWord, Earobics, and LiPS. *The Journal of Educational Research*, *97*(3), 147–158. http://doi.org/10.3200/JOER.97.3.147-158

Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Rueckl, J. G., Constable, R. T., et al. (2008). Effects of stimulus difficulty and repetition on printed word identification: An fMRI comparison of nonimpaired and reading-disabled adolescent cohorts. *Journal of Cognitive Neuroscience*, *20*(7), 1146–1160. http://doi.org/10.1162/jocn.2008.20079

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain*, *119*, 1221–1238.

Płoński, P., Gradkowski, W., Altarelli, I., Monzalvo, K., van Ermingen-Marbach, M., Grande, M., et al. (2016). Multi-parameter machine learning approach to the neuroanatomical basis of developmental dyslexia. *Human Brain Mapping*, *38*(2), 900–908. http://doi.org/10.1002/hbm.23426

Qi, T., Bin Gu, Ding, G., Gong, G., Lu, C., Peng, D., et al. (2016). More bilateral, more anterior: Alterations of brain organization in the large-scale structural network in Chinese dyslexia. *NeuroImage*, *124*(Part A), 63–74. http://doi.org/10.1016/j.neuroimage.2015.09.011

Qiu, D., Tan, L.-H., Zhou, K., & Khong, P.-L. (2008). Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: Voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. *NeuroImage*, *41*(2), 223–232. http://doi.org/10.1016/j.neuroimage.2008.02.023

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*(1), 433–447. http://doi.org/10.1146/annurev-neuro-071013-014030

Rashotte, C. A., MacPhee, K., & Torgesen, J. K. (2001). The effectiveness of a group reading instruction program with poor readers in multiple grades. *Learning Disability Quarterly*, *24*(2), 119–134. http://doi.org/10.2307/1511068

Reilhac, C., Peyrin, C., Démonet, J.-F., & Valdois, S. (2013). Role of the superior parietal lobules in letter-identity processing within strings FMRI evidence from skilled and dyslexicreaders. *Neuropsychologia*, *51*(4), 601–612. http://doi.org/10.1016/j.neuropsychologia.2012.12.010

Reynolds, M., & Wheldall, K. (2007). Reading Recovery 20 years down the track: Looking forward, looking back. *International Journal of Disability, Development and Education*, *54*(2), 199–223. http://doi.org/10.1080/10349120701330503

Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., & Papanicolaou, A. C. (2011). Engagement of temporal lobe regions predicts response to educational interventions in adolescent struggling readers. *Developmental Neuropsychology*, *36*(7), 869–888. http://doi.org/10.1080/87565641.2011.606404

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. http://doi.org/10.1016/j.jneuroling.2007.07.002

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. http://doi.org/10.1016/j.jneuroling.2005.07.003

Richards, T. L., Berninger, V. W., Aylward, E. H., Richards, A. L., Thomson, J. B., Nagy, W. E., et al. (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. *American Journal of Neuroradiology*, *23*(10), 1678–1685.

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

Richlan, F. (2012). Developmental dyslexia: Dysfunction of a left hemisphere reading network. *Frontiers in Human Neuroscience*, *6*, 120. http://doi.org/10.3389/fnhum.2012.00120

Richlan, F., Kronbichler, M., & Wimmer, H. (2012). Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies. *Human Brain Mapping*, *34*(11), 3055–3065. http://doi.org/10.1002/hbm.22127

Ricketts, J., Nation, K., & Bishop, D. V. M. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235–257. http://doi.org/10.1080/10888430701344306

Rimrodt, S. L., Peterson, D. J., Denckla, M. B., Kaufmann, W. E., & Cutting, L. E. (2010). White matter microstructural differences linked to left perisylvian language network in children with dyslexia. *Cortex*, *46*(6), 739–749. http://doi.org/10.1016/j.cortex.2009.07.008

Ritchey, K., & Goeke, J. L. (2006). Orton-Gillingham and Orton-Gillingham-based reading instruction: A review of the literature. *The Journal of Special Education*, *40*(3), 171–183. http://doi.org/10.1177/00224669060400030501

Rohde, G. K., Barnett, A. S., Basser, P. J., Marenco, S., & Pierpaoli, C. (2003). Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magnetic Resonance in Medicine*, *51*(1), 103–114. http://doi.org/10.1002/mrm.10677

Rushworth, M. F., Buckley, M. J., Behrens, T. E., Walton, M. E., & Bannerman, D. M. (2007). Functional organization of the medial frontal cortex. *Current Opinion in Neurobiology*, *17*(2), 220–227. http://doi.org/10.1016/j.conb.2007.03.001

Saine, N. L., Lerkkanen, M.-K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2010). Predicting word-level reading fluency outcomes in three contrastive groups: Remedial and computer-assisted remedial reading intervention, and mainstream instruction. *Learning and Individual Differences*, *20*(5), 402–414. http://doi.org/10.1016/j.lindif.2010.06.004

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Computers in Biology and Medicine*, *41*(12), 1110–1117. http://doi.org/10.1016/j.compbiomed.2011.06.020

Scammacca, N., Vaughn, S., Roberts, G., Wanzek, J., & Torgesen, J. K. (2007). Extensive reading interventions in K-3: From research to practice (pp. 1–84). Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeney, J. M. (2005). Severe reading difficulties - can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality*, *13*(4), 209–227. http://doi.org/10.1207/s15327035ex1304_3

Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading.

Annual Review of Neuroscience, *30*(1), 475–503. http://doi.org/10.1146/annurev.neuro.28.061604.135645

Schlesinger, N. W., & Gray, S. (2017). The impact of multisensory instruction on learning letter names and sounds, word reading, and spelling. *Annals of Dyslexia*, 1–40. http://doi.org/10.1007/s11881-017-0140-z

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science, 14,* 238-242.

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., et al. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology*, *34*(5), 585–602. http://doi.org/10.1007/s10802-006-9037-8

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biological Psychiatry*, *55*(9), 926–933. http://doi.org/10.1016/j.biopsych.2003.12.019

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*(2), 101–110.

Shaywitz, S. E. (1998). Dyslexia. *The New England Journal of Medicine*, *338*(5), 307–312.

Shaywitz, S. E., Mody, M., & Shaywitz, B. A. (2006). Neural mechanisms in dyslexia. *Current Directions in Psychological Science*, *15*(6), 278–281. http://doi.org/10.1111/j.1467-8721.2006.00452.x

Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., et al. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biological Psychiatry*, *54*(1), 25–33. http://doi.org/10.1016/S0006-3223(02)01836-X

Siegel, L. (2008). Morphological awareness skills of English language learners and children with dyslexia. *Topics in Language Disorders*, *28*(1), 15–27. http://doi.org/10.1097/01.adt.0000311413.75804.60

Siegel, L. S. (2003). Basic cognitive processes and reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), (pp. 158–181). New York, NY: Handbook of learning disabilities.

Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology*, *24*(1), 28–37. http://doi.org/10.1037//0012-1649.24.1.28

Simmons, D. C., Coyne, M. D., Kwok, O. M., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities*, *41*(2), 158–173. http://doi.org/10.1177/0022219407313587

Simon, G., Lanoë, C., Poirel, N., Rossi, S., Lubin, A., Pineau, A., & Houdé, O. (2013). Dynamics of the anatomical changes that occur in the brains of schoolchildren as they learn to read. *PLoS ONE*, *8*(12), e81789–9. http://doi.org/10.1371/journal.pone.0081789

Simos, P. G., Breier, J. I., Fletcher, J. M., Foorman, B. R., Mouzaki, A., & Papanicolaou, A. C. (2001). Age-related changes in regional brain activation during phonological decoding and printed word recognition. *Developmental Neuropsychology*, *19*(2), 191–210. http://doi.org/10.1207/S15326942DN1902_4

Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., et al. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology*, *58*(8), 1203–1213.

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

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. *Journal of Learning Disabilities*, *40*(1), 37–48. http://doi.org/10.1177/00222194070400010301

Simos, P. G., Rezaie, R., Fletcher, J. M., Juranek, J., & Papanicolaou, A. C. (2011). Neural correlates of sentence reading in children with reading difficulties. *NeuroReport*, *22*(14), 674–678. http://doi.org/10.1097/WNR.0b013e328349ecf7

Slavin, R. E., Lake, C., Davis, S., & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*(1), 1–26. http://doi.org/10.1016/j.edurev.2010.07.002

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. http://doi.org/10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. http://doi.org/10.1016/j.neuroimage.2004.07.051

Snellings, P., van der Leij, A., de Jong, P. F., & Blok, H. (2009). Enhancing the reading fluency and comprehension of children with reading disabilities in an orthographically transparent language. *Journal of Learning Disabilities*, *42*(4), 291–305. http://doi.org/10.1177/0022219408331038

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research*, *43*(2), 219–234.

Sowell, E. R. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, *24*(38), 8223–8231. http://doi.org/10.1523/JNEUROSCI.1798-04.2004

Sprenger-Charolles, L., Siegel, L. S., Béchennec, D., & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, *84*(3), 194–217. http://doi.org/10.1016/S0022-0965(03)00024-9

Sprenger-Charolles, L., Siegel, L. S., Jiménez, J. E., & Ziegler, J. C. (2011). Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Studies of Reading*, *15*(6), 498–521. http://doi.org/10.1080/10888438.2010.524463

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, *21*(10), 590–604. http://doi.org/10.1177/002221948802101003

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, *86*(1), 24–53.

Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: A meta-analysis of neuroimaging studies. *NeuroImage*, *44*(2), 489–501. http://doi.org/10.1016/j.neuroimage.2008.08.039

Strong, G. K., Torgerson, C. J., Torgerson, D., & Hulme, C. (2011). A systematic meta-analytic review of evidence for the effectiveness of the "Fast ForWord" language intervention program. *Journal of Child Psychology and Psychiatry*, *52*(3), 224–235. http://doi.org/10.1111/j.1469-7610.2010.02329.x

Stuebing, K. K., Barth, A. E., Trahan, L. H., Reddy, R. R., Miciak, J., & Fletcher, J. M. (2015). Are child cognitive characteristics strong predictors of responses to intervention? A meta-analysis. *Review of Educational Research*, *85*(3), 395–429. http://doi.org/10.3102/0034654314555996

Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology*, *56*(1), 87–114. http://doi.org/10.1006/jecp.1993.1027

Swanson, H. L. (1994). Short-term memory and working memory: Do both contribute to our understanding of academic achievement in children and adults with learning disabilities? *Journal of Learning Disabilities*, *27*(1), 34–50. http://doi.org/10.1177/002221949402700107

Swanson, H. L., & Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *Journal of Educational Psychology*, *93*(4), 720–734. http://doi.org/10.1037//O022-O663.93.4.720

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182–198.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*(5245), 81–84.

Tamnes, C. K., Ostby, Y., Fjell, A. M., Westlye, L. T., Due-Tonnessen, P., & Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: Regional age-related changes in cortical thickness and white matter volume and microstructure. *Cerebral Cortex*, *20*(3), 534–548. http://doi.org/10.1093/cercor/bhp118

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. http://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. http://doi.org/10.1073/pnas.0030098100

Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. E. (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: An fMRI study. *NeuroReport*, *12*(2), 299–307.

Tomasi, D., Ernst, T., Caparelli, E. C., & Chang, L. (2004). Practice-induced changes of brain function during visual attention: A parametric fMRI study at 4 Tesla. *NeuroImage*, *23*(4), 1414–1421. http://doi.org/10.1016/j.neuroimage.2004.07.065

Torgesen, J. K., & Bryant, B. R. (2004). Test of Phonological Awareness (2nd ed.). Austin, TX: Pro-Ed.

Torgesen, J. 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), 33–58. http://doi.org/10.1177/002221940103400104

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second-to fifth-grade children. *Scientific Studies of Reading*, *1*(2), 161–185. http://doi.org/10.1207/s1532799xssr0102_4

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. http://doi.org/10.1037/0022-0663.91.4.579

Tunmer, W. E., & Chapman, J. W. (2003). The Reading Recovery approach to preventive early intervention: As good as it gets? *Reading Psychology*, *24*(3-4), 337–360. http://doi.org/10.1080/02702710390227378

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nature Neuroscience*, *6*(7), 767–773. http://doi.org/10.1038/nn1065

Undheim, A. M. (2003). Dyslexia and psychosocial factors. A follow-up study of young Norwegian adults with a history of dyslexia in childhood. *Nordic Journal of Psychiatry*, *57*(3), 221–226. http://doi.org/10.1080/08039480310001391

Vandermosten, M., Boets, B., Poelmans, H., Sunaert, S., Wouters, J., & Ghesquiere, P. (2012a). A tractography study in dyslexia: Neuroanatomic correlates of orthographic, phonological and speech processing. *Brain*, *135*(3), 935–948. http://doi.org/10.1093/brain/awr363

Vandermosten, M., Boets, B., Wouters, J., & Ghesquière, P. (2012b). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience and Biobehavioral Reviews*, *36*(6), 1532–1552. http://doi.org/10.1016/j.neubiorev.2012.04.002

Vaughn, S., & Wanzek, J. (2014). Intensive interventions in reading for students with reading disabilities: Meaningful impacts. *Learning Disabilities Research & Practice*, *29*(2), 46–53. http://doi.org/10.1111/ldrp.12031

Vaughn, S., Wanzek, J., & Murray, C. S. (2009). Response to early reading intervention: examining higher and lower responders. *Exceptional Children*, *75*(2), 165–183. http://doi.org/10.1177/001440290907500203

Vaughn, S., Wexler, J., Roberts, G., Barth, A. A., Cirino, P. T., Romain, M. A., et al. (2011). Effects of individualized and standardized interventions on middle school students with reading disabilities. *Exceptional Children*, *77*(4), 391–407. http://doi.org/10.1177/001440291107700401

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities*, *39*(2), 157–169. http://doi.org/10.1177/00222194060390020401

Vukovic, R. K., Lesaux, N. K., & Siegel, L. S. (2010). The mathematics skills of children with reading difficulties. *Learning and Individual Differences*, *20*(6), 639–643. http://doi.org/10.1016/j.lindif.2010.08.004

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, *33*(3), 468–479.

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Childrens and adults use of spelling-sound information in three reading tasks. *Memory & Cognition*, *12*(3), 293–305.

Weissman, D. H., Woldorff, M. G., Hazlett, C. J., & Mangun, G. R. (2002). Effects of practice on executive control investigated with fMRI. *Brain Research. Cognitive Brain Research*, *15*(1), 47–60.

Willows, D. M., & Ryan, E. B. (1986). The development of grammatical sensitivity and its relationship to early reading achievement. *Reading Research Quarterly*, *21*(3), 253. http://doi.org/10.2307/747708

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research*, *50*(4), 1093–1109. http://doi.org/10.1044/1092-4388(2007/076)

Wolf, M. (1986). Rapid alternating stimulus naming in the developmental dyslexias. *Brain and Language*, *27*(2), 360–379.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001a, 2007). Woodcock Johnson III Tests of Achievement. Rolling Meadows, IL: Riverside.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001b, 2007). Woodcock Johnson III Tests of Cognitive Abilities. Rolling Meadows, IL: Riverside.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, *14*(6), 1370–1386. http://doi.org/10.1006/nimg.2001.0931

Wright, M., & Mullan, F. (2006). Dyslexia and the Phono-Graphix reading programme. *Support for Learning*, *21*(2), 77–84. http://doi.org/10.1111/j.1467-9604.2006.00408.x

Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., & Feldman, H. M. (2012). Tract profiles of white matter properties: Automating fiber-tract quantification. *PLoS ONE*, *7*(11), e49790–15. http://doi.org/10.1371/journal.pone.0049790

Yeatman, J. D., Dougherty, R. F., Rykhlevskaia, E., Sherbondy, A. J., Deutsch, G. K., Wandell, B. A., & Ben-Shachar, M. (2011). Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. *Journal of Cognitive Neuroscience*, *23*(11), 3304–3317. http://doi.org/10.1162/jocn_a_00061

APPENDICES

Appendix A: Word Lists Used for fMRI Tasks

Task	Block	Word 1	Word 2	Average Word Frequency [a]	Average Imageability [b]	Average Age of Acquisition [c]
Spelling	Easy	coar	core	19755	4.30	4.22
Spelling	Easy	deep	deap	42695	4.30	3.23
Spelling	Easy	surch	search	78523	3.30	3.84
Spelling	Easy	keap	keep	183906	2.80	3.00
Spelling	Easy	first	ferst	518924	4.10	2.59
Spelling	Easy	chane	chain	23315	6.50	3.48
Spelling	Easy	feade	feed	20084	4.40	3.10
Spelling	Easy	sleep	sleap	25606	5.50	2.27
Spelling	Easy	vote	voat	58103	4.30	4.34
Spelling	Easy	game	gaime	197271	4.90	2.61
Spelling	Easy	raite	rate	87417	2.60	4.16
Spelling	Easy	shape	shaipe	23161	4.20	3.10
Spelling	Easy	gole	goal	26030	4.00	3.52
Spelling	Easy	leave	leeve	76506	3.90	3.31
Spelling	Easy	green	grean	90773	6.20	2.19
Spelling	Difficult	bleech	bleach	1319	5.70	4.41
Spelling	Difficult	scare	scair	4294	3.80	2.90
Spelling	Difficult	skait	skate	10328	5.90	3.34
Spelling	Difficult	ghost	goast	10849	6.60	2.97
Spelling	Difficult	waike	wake	12064	3.40	3.03
Spelling	Difficult	cheer	chear	1904	4.60	3.55
Spelling	Difficult	fake	faike	7853	3.20	3.78
Spelling	Difficult	spair	spare	10605	4.20	4.72
Spelling	Difficult	trale	trail	11055	5.60	3.86
Spelling	Difficult	laugh	laff	13449	4.90	2.53
Spelling	Difficult	baike	bake	2528	5.10	3.22
Spelling	Difficult	sye	sigh	8526	3.60	4.09
Spelling	Difficult	scope	scoap	10774	4.80	4.81
Spelling	Difficult	rack	rakk	11655	5.30	4.00
Spelling	Difficult	smoke	smoake	16648	6.40	3.81
Rhyming	Easy	walk	rare	34676	5.30	2.66
Rhyming	Easy	load	mode	53273	5.15	3.26
Rhyming	Easy	school	flows	61443	5.65	3.61

Task	Block	Word 1	Word 2	Average Word Frequency [a]	Average Imageability [b]	Average Age of Acquisition [c]
Rhyming	Easy	foam	home	97047	6.55	2.97
Rhyming	Easy	mail	mane	130429	6.40	2.36
Rhyming	Easy	talk	hail	47387	4.50	4.00
Rhyming	Easy	bread	red	56441	4.05	3.62
Rhyming	Easy	seem	steam	63762	4.65	3.75
Rhyming	Easy	call	camp	128950	4.30	4.90
Rhyming	Easy	fell	help	176155	3.65	4.33
Rhyming	Easy	light	bite	52484	5.25	2.95
Rhyming	Easy	bait	space	58683	4.05	4.12
Rhyming	Easy	write	fight	86465	5.80	3.22
Rhyming	Easy	case	pass	124040	4.00	3.95
Rhyming	Easy	please	freeze	293192	3.85	3.41
Rhyming	Difficult	clap	pail	714	5.05	3.56
Rhyming	Difficult	grew	rude	9009	5.45	3.17
Rhyming	Difficult	blocks	stroke	10237	3.55	3.07
Rhyming	Difficult	snail	scale	17436	4.95	3.72
Rhyming	Difficult	theme	dream	23384	4.85	3.30
Rhyming	Difficult	bore	soar	2231	5.10	3.60
Rhyming	Difficult	cruise	drum	7908	5.60	2.84
Rhyming	Difficult	nose	blows	9318	6.15	3.06
Rhyming	Difficult	roof	soft	13930	5.70	2.91
Rhyming	Difficult	frame	fade	17476	5.00	4.32
Rhyming	Difficult	creek	knee	7821	6.10	2.24
Rhyming	Difficult	tail	whale	6769	3.40	4.35
Rhyming	Difficult	shade	laid	8013	4.30	3.42
Rhyming	Difficult	thumb	blues	9588	3.85	3.83
Rhyming	Difficult	crew	true	15253	4.75	3.00

[a] Words and frequencies obtained from English Lexicon Project, http://elexicon.wustl.edu/ (Balota et al., 2007)

[b] Higher ratings indicate words that more easily elicit a mental image (e.g., “bread” vs. “true”; Cortese & Fugett, 2004)

[c] Higher ratings indicate words that are acquired later in development (i.e., a rating of 3 indicates an average age of 4-6 years old; Cortese & Khanna, 2008)

Appendix B: Clusters of Significant Activation for fMRI Spelling Task

Contrast	Region	Cluster Size (N voxels)	Max Z value	Coordinates		
				x	y	z
Words > Symbols						
Good Readers	Left inferior frontal gyrus, middle frontal gyrus	40867	5.76	-51	33	8
	Left parietal-temporal, inferior temporal gyrus	18820	5.55	-58	-50	10
	Left fusiform, hippocampal gyrus	9063	4.49	-21	-18	-17
Poor Readers	Left inferior frontal gyrus, middle frontal gyrus	10052	4.50	-48	-5	50
Group Difference	--	--	--	--	--	--
Time Effect	Bilateral cerebellum	6740	3.77	-9	-48	-54
Group by Time Interaction	--	--	--	--	--	--
Easy Words > Difficult Words						
Good Readers	Precuneus, medial posterior cingulate	23711	4.65	-8	-73	31
Poor Readers	Precuneus, medial posterior cingulate, right inferior parietal	37071	3.99	-14	-68	40
	Bilateral cerebellum	19517	3.68	16	-40	-42
Group Difference	--	--	--	--	--	--
Time Effect	--	--	--	--	--	--
Group by Time Interaction	--	--	--	--	--	--
Difficult Words > Easy Words						
Good Readers	Left inferior frontal	8444	3.94	-57	27	19
Poor Readers	--	--	--	--	--	--
Group Difference	--	--	--	--	--	--
Time Effect	--	--	--	--	--	--
Group by Time Interaction	--	--	--	--	--	--

-- No regions activated

Appendix C: Clusters of Significant Activation for fMRI Rhyming Task

Contrast	Region	Cluster Size (N voxels)	Max Z value	Coordinates		
				x	y	z
Words > Symbols						
Good Readers	Bilateral inferior frontal, basal ganglia, hippocampal gyrus; medial frontal; left middle frontal gyrus, insula, parietal-temporal, inferior temporal	170254	6.30	-44	30	4
	Right cerebellum	12019	4.59	33	-66	-52
Poor Readers	Left inferior frontal, middle frontal	12293	4.31	-46	27	10
Group Difference	Left parietal-temporal cortex	8407	3.78	-55	-50	23
Time Effect	--	--	--	--	--	--
Group by Time Interaction	Right precuneus, anterior intraparietal sulcus, supramarginal gyrus	6429	3.69	2	-50	48
Easy Words > Difficult Words						
Good Readers	--	--	--	--	--	--
Poor Readers	--	--	--	--	--	--
Group Difference	--	--	--	--	--	--
Time Effect	--	--	--	--	--	--
Group by Time Interaction	Left cerebellum	8039	4.06	-16	-53	-39
Difficult Words > Easy Words						
Good Readers	--	--	--	--	--	--
Poor Readers	--	--	--	--	--	--
Group Difference	--	--	--	--	--	--
Time Effect	Anterior cingulate, orbitofrontal cortex	10136	3.93	1	32	16
	Left cerebellum	8446	3.83	-4	-41	-57
Group by Time Interaction	--	--	--	--	--	--

-- No regions activated