

# Why Children Fail to Learn to Read: Identifying the Cognitive, Neural, and Environmental Precursors

A dissertation

submitted by

Ola Ozernov-Palchik

In partial fulfilment for the degree of Doctor of Philosophy in Child  
Study and Human Development: Cognitive Science

Tufts University

Graduate School of Arts and Sciences

May 2018

Advisor: Maryanne Wolf, Ed.D.



## Abstract

Reading development utilizes and repurposes multiple cognitive and neural systems that support innate functions such as vision, hearing, language, and learning. In most children, these systems become seamlessly integrated through the experience of reading to form a reading circuit. In a subset of children with developmental dyslexia, however, neural specialization for reading proceeds atypically and reading impairment ensues. Because of the complexity of this circuit and the heterogeneity of dyslexia-related deficits, forming a cohesive theory of dyslexia etiology has been challenging and multiple hypotheses have been proposed. To establish the etiological causes of dyslexia, it is important to demonstrate that a particular mechanism of deficit precedes reading impairment and is present in pre-reading children. This dissertation reports findings from four studies that investigated the cognitive, neural and environmental substrates of reading failure in pre-reading children. Specifically, *Study 1* applied latent profile analysis to identify heterogeneous profiles of dyslexia risk in 1,215 kindergarten students and to demonstrate the high longitudinal stability of these profiles. Using voxel based morphometry and diffusion weighted imaging, *Study 2* demonstrated the grey matter and white matter characteristics of each of the risk profiles identified in *Study 1* in a subset of the children (n=100). The study also showed that some neuroanatomical substrates of risk predict longitudinal reading outcomes above behavioral measures. In support of the regularity extraction deficit in dyslexia, *Study 3* demonstrated unique association between the processing of temporal regularity in musical rhythm and reading development in 74 kindergarten children. Furthermore, a mediation analysis revealed a causal path from rhythm performance to phonological awareness skills and to reading skills. Finally, using

diffusion methods, *Study 4* demonstrated links between socioeconomic status (SES) and the coherence of white matter tracts important for reading development. Additionally, SES modulated the relationship between dyslexia risk and longitudinal reading outcomes. The findings in these studies may contribute to a better theoretical understanding of the etiology of dyslexia and in addition, highlight the importance of early identification and individualized remediation.

## Acknowledgements

I was extremely fortunate to have an incredible team of mentors, colleagues, and teachers. Each one taught, challenged, and inspired me to become a better thinker, researcher, teacher, and writer. This was a long journey for me with many twists and turns. To the many people who made it possible, I offer my deepest gratitude:

To my advisor Maryanne Wolf, for broadening my horizons and introducing me to the richness of dyslexia research, and to the impact of such research on children's lives. Maryanne's groundbreaking work paved the way for me and many others to explore developmental dyslexia from a multi-deficit approach. As Maryanne's student I feel myself standing on the shoulders of a giant. I couldn't have asked for a more supportive and inspiring advisor.

To Nadine Gaab, my mentor at Boston Children's Hospital, for opening the doors to the world of cognitive neuroscience for me. Nadine has mentored me from my first scientific presentation, to the research presented in this thesis. Much of what I have learned about the cognitive neuroscience of dyslexia, and academia at large, I owe to her. She demonstrated to me that there are no limits to what an ambitious and dedicated female scientist can accomplish while raising a family.

To my co-advisor Ani Patel, whose book *Music, Language, and the Brain* opened my eyes to the fascinating world of music cognition. Ani taught me how to be a better writer and thinker. He inspired me to ask more basic questions for the sake of advancement of knowledge and science, to follow ideas wherever they take us, and to form cross-disciplinary and cross-institutional collaborations in pursuit of these ideas.

To John Gabrieli, my mentor at MIT, for guiding my thinking and research through all our collaborations. I am excited to keep learning from him how to be a more rigorous scientist, collaborative colleague, and effective communicator.

To members of the Center for Reading and Language Research, Stephanie Gottwald and Melissa Orkin for motivating and supporting me through this journey and for reminding me always of the significance of our research for transforming the lives of children.

To members of the Gaab Lab past and present, including Nora Raschle, Jennifer Zuk, Bryce Becker, Meaghan Perdue, Yingying Wang, and Xi Yu, for always being there to assist, discuss, explain, and sometimes just listen to me vent. I learned so much from each of these incredible researchers.

To members of the Gabrieli Lab past and present. To Elizabeth Norton, whose journey inspired my own (for that, I will be forever grateful). I am truly fortunate to have had the chance to work with Elizabeth on the READ study and learn how to conduct research thoughtfully and methodically. To Sara Beach, for being a sounding board for my ideas, a scan buddy, and the ultimate proof reader. To Kelly Halverson, for her mastery in testing children, attention to detail, and hard work. To Wendy Georgan, for her excitement for research and for assisting me on my R programming endeavor.

To everyone else who contributed to all aspects of the studies presented here and without whom none of this research would be possible, including Abbie Cyr, Zeynep Saygin, Tyler Perrachione, Natalya Zaika, Garrett Fuselier, and Cathy Moritz.

To the many excellent professors I had at Tufts. To Gina Kuperberg, who transformed my thinking on human cognition. To Sara Johnson, who taught me many of the statistical methods I applied in my research and who was always available to answer questions. To Chip Gidney for his support and contribution to my dissertation. To Dan Dennett, Phil Holcomb, Ray Jackendoff, Ayanna Thomas, Henry David Feldman, Ariel Goldberg, and J.P. de Ruiter, thanks to whom I can proudly call myself a cognitive scientist.

To all my colleagues at SMU. Patricia Mathes, who gave me my first research position, introduced me to the world of reading research, and ignited my entire professional trajectory. To Sherril English, who taught me the skillful art of coordinating research in schools and who instilled confidence in me to pursue opportunities I would not have believed I could pursue. To David Chard, Reid Lyon, Dawn Levy, Elise Gallegos, Dylan Farmer, Beverly Weiser, and Diane Lyon, for the support, friendship, knowledge, and inspiration provided through my many years at SMU (and since).

To Scott Baker, Hank Fien, and Chris Doabler from University of Oregon for teaching me rigorous educational research practices and for continuously encouraging and supporting me in my academic enterprise.

To all the schools and families who participated in our studies. I am incredibly grateful to the administrators and teachers who welcomed us into their schools and supported us through each step of our research. Our participants were constant reminders for why this research is important. None of this work would be possible without their commitment.

To the Dyslexia Foundation, Will Baker, Ben Powers and others. I am so fortunate to be a part of this incredible family of researchers, practitioners, and visionaries who strive every day to improve the lives of children with dyslexia.

To Nancy Duggan and others at Decoding Dyslexia, whose tireless advocacy and remarkable progress inspired me to work harder to advance early identification of dyslexia research.

To my family, whose love sustained me and fueled me.

I come from generations of scholars and from a heritage of strong and educated women. Education was everything in my family, even when they had nothing. My grandmother made shoes out of rags to walk to school during cold winters in the midst of World War II in Russia. My mother pursued her PhD after immigrating for a second time and in her third language. My father supported my mother, my sister, and me in our studies with great enthusiasm and interest. He carefully read all three of our dissertations, one on John Dewey, another one on bioengineering, and the current one on dyslexia, providing detailed and poignant feedback. Everything I am, I owe to my family.

To my sister Katia, who I have been following around since I could crawl. I am lucky to have such a wonderful work friend, laugh friend, and best friend.

To my husband Boris, the love of my life, for his (mostly) patient feedback on my writing, for his pride in my accomplishments, for his boundless optimism and sense of humor, and for always being a solid rock to lean on. I am truly grateful for having the best husband anyone could ever wish for.

To my sweetest Oziki, the happiness and meaning in my life, for showing me the true wonders of development. Ozzy is the true linguist and scientist in our family. I am forever inspired by his endless curiosity and passion for learning. My love for him is all encompassing and his existence is what I am most grateful for.

Funding: Evans Family Literacy Fellowship, NICHD grant 1R01HD067312-01 to co-PIs John Gabrieli and Nadine Gaab, Tufts Collaborates grant to Maryanne Wolf and Aniruddh Patel.

## Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>Table of Contents .....</b>	<b>vii</b>
<b>List of Tables .....</b>	<b>xii</b>
<b>List of Figures.....</b>	<b>xiii</b>
<b>1 Introduction .....</b>	<b>1</b>
1.1 Typical Reading Development.....	3
1.1.1 Phonological system.....	3
1.1.2 Orthographic system.....	4
1.1.3 Semantic/executive system.....	5
1.1.4 Connectivity across the reading circuit .....	7
1.1.5 Neuronal recycling and interactive specialization.....	8
1.2 Developmental Dyslexia .....	9
1.2.1 Genetics of dyslexia.....	10
1.2.2 Brain basis of dyslexia.....	11
1.2.3 Environmental factors.....	12
1.3 Studies in Pre-Readers .....	13
1.3.1 Differences in brain structure .....	14
1.3.2 Differences in white matter connectivity.....	14
1.3.3 Differences in brain function .....	15
1.3.4 Longitudinal evidence .....	16
1.4 Etiological Theories of Dyslexia.....	18
1.4.1 Phonological deficit.....	18
1.4.1.1..... Phonological processing	19
1.4.1.2..... Auditory processing	20
1.4.1.3..... Phonological awareness	21
1.4.2 RAN/Phonological retrieval .....	22
1.4.3 Visual.....	25



1.4.3.1.....	Low-level	26
1.4.3.2.....	Visual attention deficits	28
1.4.4	Temporal/motor coordination.....	29
1.4.5	Musical rhythm.....	31
1.4.6	Cerebellum.....	34
1.4.7	Statistical learning/perceptual adaptation .....	36
1.4.8	General language .....	37
<b>2</b>	<b>Study 1: Longitudinal Stability of Pre-Reading Skill Profiles of Kindergarten Children: Implications for Early Screening and Theories of Reading .....</b>	<b>42</b>
2.1	Abstract .....	42
2.2	Introduction .....	43
2.2.1	Early identification of dyslexia risk.....	45
2.2.2	Theories of dyslexia and implications for diagnosis and treatment ...	47
2.2.3	Identification of dyslexia risk .....	50
2.2.4	Longitudinal stability of risk classification .....	50
2.2.5	Cognitive and environmental influences .....	51
2.2.6	Current study .....	52
2.3	Methods .....	53
2.3.1	Participants and data collection procedures.....	53
2.3.2	Measures .....	54
2.3.2.1.....	Classification variables (the pre-kindergarten and kindergarten screening battery).....	54
2.3.2.2.....	1 <sup>st</sup> Grade (Year 2) assessment included the same measures as in kindergarten and additional measures:.....	55
2.3.3	Latent profile analysis .....	56
2.4	Results .....	59
2.4.1	Subtypes of early reading profiles .....	59
2.4.2	Profile descriptions based on kindergarten data .....	61
2.4.3	Longitudinal prediction based on latent class membership.....	64
2.4.4	Cognitive and environmental factors.....	69
2.5	Discussion .....	71
2.5.1	Implications for dyslexia risk subtypes .....	71
2.5.2	Longitudinal stability of risk classifications.....	75
2.5.3	Cognitive and environmental factors.....	77
2.5.4	Summary.....	79
<b>3</b>	<b>Study 2: Distinct Neural Alterations of Heterogeneous Dyslexia Risk Profiles .....</b>	<b>80</b>
3.1	Abstract .....	80
3.2	Introduction .....	81
3.3	Methods.....	86
3.3.1	Participants .....	86
3.3.2	KG behavioral measures.....	87
3.3.3	Background measures.....	88

3.3.4	Imaging procedures .....	88
3.3.5	2 <sup>nd</sup> Grade behavioral measures .....	89
3.3.6	Behavioral analysis and statistics .....	90
3.3.7	VBM analysis .....	90
3.3.8	DWI analysis .....	93
3.3.9	Power analysis .....	94
3.4	Results .....	94
3.4.1	KG behavioral results .....	94
3.4.2	Background measures results .....	95
3.4.3	2 <sup>nd</sup> Grade behavioral results .....	95
3.4.4	KG VBM results .....	98
3.4.5	KG DWI results .....	100
3.4.6	Longitudinal prediction analysis .....	101
3.4.7	Bias analysis .....	104
3.5	Discussion .....	104
3.5.1	Behavioral and grey matter group characteristics .....	105
3.5.1.1	..... LSK risk profile .....	105
3.5.1.2	..... Multi-risk profile .....	107
3.5.1.3	..... RAN risk profile .....	110
3.5.2	PA risk profile .....	112
3.5.3	White matter alterations.....	113
3.5.4	Contribution of neural measures to prediction of reading outcomes	116
3.5.5	Limitations.....	117
3.5.6	Theoretical implications of findings.....	118
<b>4</b>	<b>Study 3: Relationships Between Early Literacy and Non-Linguistic Rhythmic Processes in Kindergarten Children .....</b>	<b>121</b>
4.1	Abstract .....	121
4.2	Introduction .....	122
4.3	Methods.....	129
4.3.1	Participants and general procedures .....	129
4.3.2	Language and cognitive measures .....	130
4.3.3	Rhythm perception measures.....	131
4.4	Results .....	134
4.4.1	Exploratory factor analysis .....	134
4.4.2	Rhythm and literacy association.....	135
4.4.3	Mediation analysis.....	139
4.4.4	Regression analysis.....	140
4.5	Discussion .....	141
4.5.1	Relationship between rhythm and phonological awareness .....	142
4.5.2	Relationship between metrical processing and early literacy .....	144
4.5.3	Considerations and implications.....	145

<b>5</b>	<b>Study 4: The Relationship Between Socioeconomic Status and White Matter Coherence in Pre-Reading Children: A Longitudinal Investigation</b>	<b>148</b>
5.1	Abstract .....	148
5.2	Introduction .....	149
5.3	Methods .....	155
5.3.1	Overview .....	155
5.3.2	Participants .....	156
5.3.3	Kindergarten behavioral measures .....	157
5.3.4	2 <sup>nd</sup> Grade behavioral measures .....	158
5.3.5	Measures of reading outcomes .....	159
5.3.6	Socioeconomic status .....	159
5.3.7	Home literacy environment .....	160
5.3.8	Parental history of reading difficulties .....	160
5.3.9	Imaging procedures .....	160
5.3.10	Identification of key white matter tracts .....	162
5.4	Statistical analyses.....	163
5.4.1	Pre-literacy and literacy performance by SES.....	163
5.4.2	Differences between SES groups in tract FA .....	164
5.4.3	Differences between SES groups in AD and RD .....	164
5.4.4	Modulating effects of SES on FA/risk-reading relationship (longitudinal) .....	165
5.4.5	Classification of reading outcomes in lower- SES .....	165
5.5	Results .....	166
5.5.1	Differences by SES in (pre)literacy performance.....	167
5.5.2	Differences between SES groups in tract FA .....	168
5.5.3	Differences between SES groups in AD and RD .....	169
5.5.4	Modulating effects of SES on the risk-reading relationship (longitudinal) .....	169
5.5.5	Modulating effects of SES on the FA-reading relationship (longitudinal) .....	170
5.5.6	Classification of reading outcomes in low SES group .....	171
5.6	Discussion .....	174
5.6.1	Association between SES and (pre-) literacy performance .....	175
5.6.2	Left ILF tract coherence and SES.....	175
5.6.3	Modulating effects of SES on left ILF-reading relationship .....	179
5.6.4	Classification of reading outcomes in the lower-SES group.....	180
5.6.5	Limitations and future directions.....	182
5.7	Significance.....	182
<b>6</b>	<b>Conclusions and Future Directions.....</b>	<b>185</b>
6.1	Theories of Dyslexia Etiology.....	185
6.2	Multi-Factorial Model of Risk .....	188
6.3	Early Identification of Risk .....	190
6.4	Socioeconomic Disparities in Reading .....	192
6.5	Bridging Research and Practice .....	194
	<b>References .....</b>	<b>195</b>



## List of Tables

Table 2.1: Screening and follow-up sample characteristics.....	52
Table 2.2: Model comparison for optimum latent class solution.....	57
Table 2.3: Point estimates of each of the six latent classes across IQ, PA, NWR, RAN and LSK predictors.....	57
Table 2.4: Reading performance of the members of each profile.....	57
Table 2.5: Mean point estimates and significance tests for longitudinal prediction of reading outcomes from latent class formation at kindergarten.....	58
Table 3.1: Comparison of behavioral scores across the LPA groups .....	85
Table 3.2: Brain regions showing significant differences in gray matter volume index between LPA groups. ....	86
Table 3.3. Results of stepwise multiple regression analyses predicting reading comprehension and fluency .....	89
Supplemental Table 3.4: Comparison of family and home environment measures across LPA groups .....	104
Table 4.1: Temporal intervals between note onsets .....	116
Table 4.2 Language (standard scores) and rhythm (percentage accuracy) performance. ....	119
Table 4.3: Correlation coefficients between early literacy measures and rhythm variables .....	121
Table 4.4: Hierarchical regression predicting Graphemes scores.....	122
Table 5.1: Model with the highest reading-outcome classification performance in lower-SES group.....	150
Supplemental Table 5.1. ....	157
Supplemental Table 5.2 .....	158

## List of Figures

Figure 2.1: Latent profile analysis model for the identification of reading subgroups: Optimal solution. ....	55
Figure 2.2 Latent class regression predicting reading outcomes from the earlier-formed kindergarten latent profiles.....	59
Figure 2.3: Latent class distribution across SES groups.....	61
Figure 3.1: Grey matter volume differences between each risk group and the average group.....	87
Figure 3.2. Average fractional anisotropy of left SLF, AF, and ILF by LPA group. ....	88
Figure 3.3. KG grey matter volume of the VOT predicts unique variance in GORT Comprehension .....	90
Figure 4.1: Screenshot of the rhythm school app .....	115
Figure 4.2 Scatterplots showing the relationship between the PAT Graphemes scores and rhythm performance .....	120
Figure 4.3: Rhythm-phonology-literacy mediation model .....	121
Figure 5.1: Fractional anisotropy differences between lower- and higher-SES children .....	146
Figure 5.2: Differences in frequencies by KG risk, 2nd grade reading outcomes, and SES.....	147
Figure 5.3: Association between tract FA (nodes 1-19) and reading by SES.....	148
Figure 5.4: Variables ranked by relative importance.....	149



# 1 Introduction

**All fluent readers are alike; each dysfluent reader is dysfluent in a different way.**

Lev Tolstoy opened his timeless novel *Anna Karenina* with the often-quoted line: “All happy families are alike; each unhappy family is unhappy in its own way.” Similarly in reading: fluent reading develops in like manner across most readers; in dyslexia, each reader is challenged in his or her own way, based on patterns of deficits, strengths, and environments.

A child learning to read is a child uncovering the world. As children acquire the ability to decode the symbols of their own language, they begin to realize that the places around them have names and billboards have meaning. They discover the key to accessing worlds beyond their reality and they can use it whenever they want. The more they read, the better they get. The better they get, the more they read. Reading develops their vocabulary, world knowledge, imagination, and critical thinking. Reading lets the child experience the world through the eyes of others, from *Harry Potter* to *Huckleberry Finn*, and develop the art of empathy. Most amazing of all, reading transforms a child’s brain from one that can see, hear, hold things in memory, and process language to one that utilizes all those functions to construct meaning letter by letter, word by word.



What happens, however, when an otherwise typically developing child struggles to read without any discernable causes? Without *appropriate intervention*, such child will embark on a different path, a path of a snowballing reading failure. This child might remain silent, while his or her peers enthusiastically identify names and sounds of letters. This child might begin avoiding activities related to early reading. Such avoidance would lead to an increasingly growing gap in reading between them and their peers. Such a gap has been estimated at 1,000 words a week as early as 1<sup>st</sup> grade (Stanovich, 1986). This discrepancy in reading, would lead to other gaps in grammar and vocabulary skills. The child may start losing his motivation for learning and confidence in her overall ability, while struggling to complete class-and homework activities that will become increasingly more challenging. In 2<sup>nd</sup> or 3<sup>rd</sup> grade, when classroom instruction transitions from learning to read to reading to learn, the gap between that child and his or her classmates will start spilling over into other subjects such as math, social sciences, and science. The likelihood of the child graduating from high school is lower and the likelihood of ending up in the criminal justice system is higher, particularly if he or she is from an underprivileged background (Grigorenko, 2006).

This thesis investigates the possible underlying mechanisms of reading failure in such children, most of whom have developmental dyslexia. By investigating the cognitive, neural, and environmental substrates of reading-related deficits in pre-reading children, I aim to promote the understanding of possible core causes of dyslexia. I hope that this work will aid in early identification and effective

remediation of children, *prior* to formal reading instruction -- and that it can thus help prevent the negative psychosocial consequences of reading failure.

## **1.1 Typical Reading Development**

Reading is a complex neurobiological process that involves multiple systems including language, vision, speech, hearing, attention, and learning systems (Dehaene, 2009; Wolf, 2008). Through the experience of reading, these systems come together to form the functional and structural brain reading network (Dehaene et al., 2010). Although the components of this network have been consistently identified, reliance on these components shifts across reading development and varies across orthographies (Black, Xia, & Hoefft, 2017). Additionally, since there is no one-to-one correspondence between a particular brain structure and its function, each of the network's components can support a multitude of reading-related functions, a range of which may not be comprehensively characterized yet.

### **1.1.1 *Phonological system***

Reading development starts as early as *in utero*, with repeated exposure to the low-pass filtered and muffled sounds of the language a child will be born into. Newborns at 2 days old already show a preference for native language versus foreign language (Moon, Cooper, & Fifer, 1993). By 6 months of age, infants are consistently able to distinguish the speech sounds of their native language from those of other languages (Kuhl et al., 2006). The ability to recognize and distinguish among native speech sounds is called *phonological processing*. As children gain expertise in processing the sounds of their language, they begin to acquire an

awareness of the sound structure and understand that words are comprised of smaller, discernible units. This metacognitive understanding of the intricacies of speech is called *phonological awareness*. In the educational context it is often measured by testing child's ability to manipulate speech by forming rhymes or taking words apart to form other words.

Neuroimaging studies have implicated the left temporo-parietal and inferior frontal regions in phonological skills (Brennan, Cao, Pedroarena-Leal, McNorgan, & Booth, 2013; Graves, Desai, Humphries, Seidenberg, & Binder, 2009; Martin, Schurz, Kronbichler, & Richlan, 2015; Martin, 2015; Vigneau et al., 2006). These regions form the *dorsal reading network* and include the posterior superior temporal gyrus, supramarginal gyrus and dorsal inferior frontal gyrus.

### **1.1.2 *Orthographic system***

When children enter school, or shortly beforehand, they begin learning that speech sounds have visual representations called letters and start gaining *orthographic awareness*. This awareness allows children to *decode*, that is, translate words from print to sound, and retrieve the meaning of the written word. With increased experience and following several successful decoding attempts (Ehri, 2005), children learn the orthographic representations of words and develop the ability to efficiently retrieve the meaning of words in future encounters with them. As children become more efficient with orthographic processing, they rely less and less on the phonological system, and become faster and more fluent readers (Ehri, 1995).

Learning the form of letter strings and words is supported by higher-order visual recognition systems in the *ventral reading network* that encompasses the occipito-temporal region including *the visual word form area* (VWFA-Dehaene, 2011) and the inferior temporal gyrus. As children master the letter to sound correspondences, the link between visual recognition regions and the temporo-parietal phonological awareness regions strengthens (Pugh et al., 2000). Through repeated experiences with a given word, children learn to identify the word as a whole automatically, and reading proceeds more fluently from the visual word recognition areas to meaning areas commonly located in the middle temporal lobe, angular gyrus, in the inferior frontal gyrus (Binder et al., 1999; Gabrieli, Poldrack, & Desmond, 1998; Price, 2011; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001).

### **1.1.3 *Semantic/executive system***

Fluent reading relies on automatic identification of familiar words and the ability to efficiently decode unfamiliar words. Fluent reading is additionally governed by global cognitive mechanisms, such as attention and executive control (Breznitz, 2006; Wolf & Katzir-Cohen, 2001). Some of these mechanisms are located in the prefrontal brain regions, including the inferior frontal gyrus, which is also implicated in phonology and semantics (Adleman et al., 2002; Beneventi, Tonnessen, Ersland, & Hugdahl, 2010; Blumenfeld, Booth, & Burman, 2006; Devlin, Matthews, & Rushworth, 2003; Kovelman et al., 2011; Poldrack et al., 1999; Price, 1998). Reading comprehension requires, among other things, lexical and background knowledge, correct utilization of linguistic cues, and inference and reasoning skills (Just, 2013). Semantics and comprehension are supported by many

brain regions, especially the inferior frontal, angular gyrus, anterior fusiform gyrus, and middle temporal areas (Badre & Wagner, 2007; Binder, Desai, Graves, & Conant, 2009; Booth, Bebko, Burman, & Bitan, 2007; Chou et al., 2006; Cutting et al., 2006; Dobbins, Foley, Schacter, & Wagner, 2002; Friederici & Weissenborn, 2007; Kuperberg, Sitnikova, & Lakshmanan, 2008; Poldrack et al., 1999; Price, 1998).

To summarize, two major networks and a third, more general network have been proposed as involved in reading: the left ventral system for orthographic processing, the left dorsal system for phonological processing, and a more distributed system across the brain for semantic and sentence/syntactic processing. Several researchers (Pugh et al., 2001; Pugh, 1996) have proposed that the development sequence of the reading network closely tracks the behavioral developmental reading stages. The dorsal network for decoding emerges early in development, followed by the ventral system, which through its connectivity to the dorsal system becomes specialized for the automatic recognition of print. As the ventral system strengthens and word recognition becomes increasingly automatic, the dorsal system becomes less involved. A recent meta-analysis of functional MRI studies further confirmed this developmental shift (Martin et al., 2015). Additionally, while only the left hemispheric regions have been discussed here, reading development is thought to involve the initial engagement of the left and right regions simultaneously, followed by disengagement of right hemisphere regions (Brem et al., 2010; Turkeltaub, Gareau, Flowers, Zeffiro, & Eden, 2003; Yamada et al., 2011).

#### 1.1.4 *Connectivity across the reading circuit*

The integration across the distributed and distant brain networks is made possible by the long-range white matter connections in the brain. These connections are formed by myelinated axons of neurons that communicate with other neurons and together form large bundles called *tracts* or *fasciculi*. These bundles form pathways in the brain through which neuronal signals are transmitted and orchestrate distant brain regions into networks. The dorsal system includes the *left arcuate fasciculus* (AF), connecting the superior temporal lobe with the inferior frontal gyrus (IFG), and the *left superior longitudinal fasciculus* (SLF), connecting the inferior parietal with the inferior frontal/premotor regions (Catani & Jones, 2005; Shinoura et al., 2013; Thiebaut de Schotten, Cohen, Amemiya, Braga, & Dehaene, 2014; Vandermosten, Boets, Poelmans, et al., 2012; Wong, Chandrasekaran, Garibaldi, & Wong, 2011; Yeatman et al., 2011). The ventral system includes the *left inferior longitudinal fasciculus* (ILF), connecting the posterior inferior temporal gyrus with the ventral anterior and medial temporal lobe and the *inferior fronto occipital fasciculus* (IFOF), connecting the occipito-temporal region to the prefrontal cortex (Catani, Howard, Pajevic, & Jones, 2002; Catani & Jones, 2005; Martino, Brogna, Robles, Vergani, & Duffau, 2010; Sarubbo, De Benedictis, Maldonado, Basso, & Duffau, 2013; Vandermosten, Boets, Poelmans, et al., 2012). Both ILF and IFOF are thought to pass in close proximity and connect to the visual word form area (Yeatman, Rauschecker, & Wandell, 2013).

### **1.1.5 *Neuronal recycling and interactive specialization***

Reading is a recent invention in the human evolution, having evolved approximately 5,000 years ago (Wolf, 2008). It has been proposed that brain specialization for reading involves the recycling of existing circuits to support novel functions (Neuronal Recycling Hypothesis—Dehaene, 2007). For example, visual word identification relies on modification of cortical visual circuits that evolved for generic object recognition. In turn, the visual symbols for reading developed across societies follow common geometric patterns present in nature. There is experimental and computational modeling evidence for this hypothesis. It has been demonstrated that brain areas responsive to faces, became responsive to words after receiving reading instruction (Dehaene et al., 2010). In another study, greater specialization for letters in the left fusiform gyrus and smaller left fusiform face area were associated with increased reading ability in young children (Centanni et al., 2018). A computational study demonstrated that a model learned letter identification by extracting domain-general visual features from natural scenes in lower-level processing layers, and domain specific features in higher-level layers in response to printer letters exposure (Testolin, Stoianov, & Zorzi, 2017).

This functional specialization of the reading circuit is thought to be driven by patterns of extrinsic connectivity of the circuit (Schlaggar & McCandliss, 2007). In this Interactive Specialization view (Johnson, 2011) the emergence of functional specialization is driven by changes in interactions (connections) within and between brain regions. In support of this hypothesis, the connectivity of the VWFA in pre-readers, but not its responsiveness to print, has been shown to predict the

functional specificity the region three years later (Saygin et al., 2016). Thus, the brain specializes for reading through repurposing of evolutionary old networks into new networks, through the process of interactive specialization.

In sum, the reading network consists of multiple neural systems -- dorsal, ventral, prefrontal, and other -- distributed areas that support the different processes involved in reading: phonological, orthographic, semantic, and attentional. Understanding the complex multi componential nature of the reading network is important for appreciating the difficulty of identifying the causal mechanisms of the failure to develop reading: developmental dyslexia. Indeed, a deficit in each of these neural components can hinder the emergence of typical reading. Next, I will discuss developmental dyslexia and the various theories proposed to explain the etiological mechanisms of this disorder.

## **1.2 Developmental Dyslexia**

As discussed, for most children embarking on the journey of learning to read, the brain's specialization for reading progresses in accordance with the cognitive milestones characterizing typical reading development. In the United States, most children by the end of second grade are ready to transition from learning to read to reading to learn. However, there is a subset of children, 5-17%, who despite adequate instruction and otherwise typical development, fail to develop the expected reading skills (Lyon, Shaywitz, & Shaywitz, 2003; Siegel, 2006). These children, who have developmental dyslexia, require intensive instructional intervention and are likely to experience at least some difficulties with reading well into adulthood (Shaywitz et al., 1999; Wilson & Lesaux, 2001).



Research suggests that dyslexia represents the lower end of normal distribution of word reading skills, and that diagnosis requires setting a somewhat arbitrary threshold on a continuous variable (Peterson & Pennington, 2012). Neuroimaging studies of children and adults with dyslexia reveal structural and functional atypicalities in the systems comprising the neural reading network (Norton, Beach, & Gabrieli, 2015).

### **1.2.1 *Genetics of dyslexia***

The heritability of dyslexia has been estimated at 60% (Grigorenko, 2004). Several dyslexia susceptibility genes have been proposed: ROBO1, DCDC2, DYX1C1, KIAA0319 (Marino et al., 2014; Mascheretti et al., 2013). These genes are supported by a number of independent replication studies and are thought to be involved in cortical development. Neurodevelopment theories of dyslexia propose that functional genetic variants affect neuronal migration, neurite outgrowth, cortical morphogenesis and ciliary<sup>1</sup> structure and function. This alters the typical course of the development of brain regions and the structural connections among the regions that support the precursors of reading such as phonological processing (Galaburda, 2006). This, in turn, disrupts the emergence of the functional circuits specialized for reading. Alterations in these circuits may be associated with the range of sensorimotor, perceptual, and cognitive deficits reported in dyslexia (Goswami, 2015; Ozernov-Palchik, Yu, Wang, & Gaab, 2016).

---

<sup>1</sup>Cilia are organelles that project from the surface of neuronal cells and are involved in receiving signals and transmitting them into the cell (Green & Mykytyn, 2014).

Evidence for this developmental sequence comes from genetic studies. Polymorphisms in dyslexia genes have been linked to differences in cortical structure and white matter connectivity within the dorsal and ventral networks and to deficits in phonological processing (Cope et al., 2012; Czamara et al., 2011; Dennis et al., 2009; Eicher & Gruen, 2013; Ludwig et al., 2008; Marino et al., 2012; Newbury et al., 2011; Pinel et al., 2012; Scerri, Darki, Newbury, Whitehouse, Peyrard-Janvid, Matsson, Ang, Pennell, Ring, Stein, et al., 2012; Wilcke et al., 2012). Thus, evidence suggests a developmental pathway from dyslexia susceptibility genes to differential development of brain function and structure crucial for learning to read.

### **1.2.2 *Brain basis of dyslexia***

In line with behavioral research, the most consistently reported differences in dyslexia are in the dorsal phonological system and the ventral orthographic system. Indeed, reduced grey matter volume and cortical thickness (considered to be an index of the amount of grey matter, consisting of neuronal cell bodies, in a specific region) and degree of gyrification (index of cortical organization) in the temporo-parietal and occipito-temporal regions have been reported in children and adults with dyslexia (Richlan, Kronbichler, & Wimmer, 2009, 2013; Richlan, 2011, 2012; Schurz et al., 2014). Similarly, reduced coherence of the white matter tracts connecting these regions namely, left AF, SLF, ILF, and IFOF (Carter et al., 2009; Vandermosten, Boets, Poelmans, et al., 2012), have been reported. Importantly, studies have demonstrated that children with dyslexia exhibit brain differences within the reading network even when compared to reading-level-matched

children, i.e., younger children without dyslexia who read at the same level (Hoeft et al., 2006). This suggests that at least some of the brain characteristics of dyslexia are possibly related to dyslexia etiology<sup>2</sup>, but not to reading failure.

### **1.2.3 *Environmental factors***

Despite the hereditary nature of reading ability and dyslexia, environmental influences play a significant role in shaping reading outcomes. Since it is difficult to dissociate heritable influences from those of post-and prenatal environment (which could also affect genetic expression through epigenetic mechanisms), the unique contribution of genetic and environmental factors, respectively, to dyslexia is not completely understood. It is fairly established, however, that brain regions that support reading and language are particularly affected by disadvantageous environments (Hackman, Farah, & Meaney, 2010). There has been some evidence for gene by environment interaction in dyslexia, with studies demonstrating greater genetic influences on reading in more socioeconomically advantageous environments (Friend, DeFries, & Olson, 2008). Only few studies to date have investigated the complex interplay between genes and environment in affecting brain development in dyslexia. It has been suggested that in children who are at genetic risk for dyslexia, poor environmental conditions can exert unfavorable influences increasing the likelihood of later reading failure (Snowling & Melby-Lervåg, 2016). However, another study demonstrated attenuated effects of environment, measured as home literacy practices, on brain specialization for

---

<sup>2</sup> The term “etiology” refers to initial or distal causes of dyslexia in the population. That is, the early factors that produce changes in the trajectory of reading development.

reading in children at familial risk for dyslexia as compared to controls (Powers, Wang, Beach, Sideridis, & Gaab, 2016). It is important to note that children in this study were primarily from higher socioeconomic backgrounds. In another study, brain activation patterns in children with dyslexia were similar, regardless of socioeconomic status (Monzalvo, Fluss, Billard, Dehaene, & Dehaene-Lambertz, 2012) and yet another study showed brain differences in response to intervention in children with dyslexia from lower socioeconomic families, but not from higher socioeconomic families (Romeo et al., 2017). Thus, there is a need for more studies to illuminate the complex process through which reading ability and dyslexia is transmitted from parents to children through genetic and environmental influences. The most convincing evidence for dissociating environmental factors from heritable factors comes from studies of young pre-reading children at genetic risk for dyslexia.

### **1.3 Studies in Pre-Readers**

In order to understand the etiological mechanisms of dyslexia, it is important to demonstrate that the brain substrates often identified in already-reading individuals with dyslexia are present prior to the onset of reading instruction. There is increasing evidence from neuroimaging studies of pre-readers that many of the brain atypicalities observed for dyslexia predate reading failure. Indeed, although dyslexia is often not diagnosed until the child fails to develop reading skills as expected in 2<sup>nd</sup> or 3<sup>rd</sup> grade, there is increasing evidence that the neurobiological pathway to reading failure starts *in utero* (Galaburda, LoTurco, Ramus, Fitch, & Rosen, 2006; Guttorm, Leppanen, Hamalainen, Eklund, &

Lyytinen, 2010; Koster et al., 2005; Langer et al., 2017; Leppanen, 2012). Because of the strong hereditary nature of dyslexia, most studies of pre-readers focused on comparing children who have familial history of dyslexia with children without familial history, in order to identify which brain substrates of dyslexia are also associated with dyslexia risk. Other studies determined risk based on performance on behavioral measures of early literacy.

### **1.3.1 *Differences in brain structure***

Studies in pre-reading children at risk for dyslexia have demonstrated that some differences in anatomical properties of the reading network are present even before children begin learning how to read. Several studies have reported reduced gray matter volume (Raschle, Chang, & Gaab, 2011) and cortical thickness (Black, 2012; Clark, 2014; Hosseini et al., 2013) in occipito-temporal and temporo-parietal regions in pre-readers at-risk for dyslexia. In support of the importance of the findings for early reading development, there was a significant association between structural indices in these regions and out-of-scanner pre-literacy performance (Raschle et al., 2011).

### **1.3.2 *Differences in white matter connectivity***

Several studies have demonstrated reduced coherence of the arcuate fasciculus and superior longitudinal fasciculus in pre-reading children at familial (Kraft et al., 2016; Vandermosten, 2015; Wang et al., 2016) and behavioral (Saygin et al., 2013) risk for dyslexia. Importantly, reduced coherence in the arcuate fasciculus was demonstrated in infants from families with a history of dyslexia and was related to worse language skills in all infants (with and without risk) (Langer

et al., 2017). Differences in at-risk pre-readers have also been reported in another pathways important for reading -- the inferior frontal occipital fasciculus that is involved in visual word form recognition (Vandermosten et al., 2016). Other studies, however, failed to find differences in this pathway (Kraft et al., 2016) or another pathway, inferior longitudinal fasciculus, also involved in word recognition (Saygin et al., 2013; Wang et al., 2016). While the findings of dorsal, but not ventral white matter atypicalities in pre-readers are inconsistent, they could suggest that differences in the ventral orthographical reading network develop as the result of reduced reading experience.

### **1.3.3 *Differences in brain function***

Functional brain alterations within the dorsal phonological system have been consistently observed in pre-reading children with a familial risk of dyslexia. For instance, at-risk preschoolers compared to no-risk preschoolers already show reduced neural activation during a first-sound matching and rapid auditory processing tasks (Raschle, Stering, Meissner, & Gaab, 2013; Raschle, Zuk, & Gaab, 2012). Brain activation in these tasks correlated with behavioral performance on pre-literacy measures (e.g., rapid automatized naming and phonological awareness).

Studies have demonstrated atypical orthographic processing in school-age children and adults with dyslexia (Temple et al., 2001), but it was suggested that these deficits are the consequence of a more limited reading experience in individuals with dyslexia (Olulade, Napoliello, & Eden, 2013). Nevertheless, atypical orthographic processing has been demonstrated in two studies of pre-

reading children at behavioral risk for dyslexia. In one study, at-risk kindergarten children demonstrated reduced activation for letters (Yamada et al., 2011). In another study of Norwegian 6-year-old children, the at-risk group demonstrated reduced activation to sight words in the occipito-temporal regions important for reading (Specht et al., 2009). These studies should be considered with caution. Behavioral risk could indicate, however, differences in environmental rather than hereditary influences (such as limited print exposure).

Overall, the literature in pre-reading children so far has demonstrated that some brain alterations in dyslexia seem to predate the onset of reading instruction. However, it is important to note that in these studies, it is still unclear which of the children will develop a reading disability. Therefore, it is important to conduct a longitudinal follow-up in later grades when a diagnosis of dyslexia can be made. Evidence from longitudinal studies will be discussed next.

#### **1.3.4 *Longitudinal evidence***

Longitudinal investigations have demonstrated that the brain differences reported in studies of pre-readers are not merely epiphenomenal to dyslexia risk, but have important implications for reading outcomes. In one study in German children, more gray matter volume in the left superior temporal gyrus at 1<sup>st</sup> grade was associated with greater gains in reading proficiency between 1<sup>st</sup> and 2<sup>nd</sup> grade (Linkersdörfer et al., 2014). This suggests that variation in brain structure can impact the development potential of reading early in schooling. Another longitudinal study in a group of Norwegian pre-reading children with and without a familial risk of dyslexia showed that children who eventually received a diagnosis

of dyslexia in 6<sup>th</sup> grade had a significantly thinner cortex in several low-level auditory, visual, and executive functioning regions in 1<sup>st</sup> grade. A study in Dutch pre-readers found reduced left-lateralization of the planum temporale, a superior-temporal region that has bigger left-to-right hemispheric volume and is involved in higher-order auditory processing, in children with familial risk of dyslexia. Interestingly, reduced lateralization was not related to reading outcomes in at-risk children, suggesting it could be epiphenomenal to familial risk. A Swiss fMRI study demonstrated that activation in response to words in the visual word form region (occipito-temporal region), together with an ERP component and RAN scores, discriminated with 94% accuracy between a small group of kindergarten children who became good readers from those who became poor readers in 2<sup>nd</sup> grade (Bach, Richardson, Brandeis, Martin, & Brem, 2013). In a study of German kindergartners with and without risk for dyslexia, baseline behavioral measures and left AF coherence predicted with 80% accuracy which children will develop dyslexia after two years of schooling (Kraft et al., 2016).

Taken together, these research studies, while still limited, suggest that neural alterations in dyslexia predate reading onset and reflect the differential developmental trajectory of reading brain networks as the result of genetic predisposition for dyslexia. While these studies reveal some clues on possible etiology of dyslexia, because of their dichotomous nature (i.e. comparing children with and without risk), they do not reflect the heterogeneity in dyslexia deficits that emerges in studies of adults.



## **1.4 Etiological Theories of Dyslexia**

Reading development involves multiple genetic, neural, cognitive, and environmental factors. Deficits in individuals with dyslexia as compared to typical readers have been found on a large number of diverse tasks motivating the development of multiple theories of dyslexia etiology. Next, some of these theories will be reviewed and evaluated.

### **1.4.1 *Phonological deficit***

Most compelling evidence for possible etiological mechanisms of dyslexia comes from literature on phonology (Brady, Shankweiler, & Mann, 1983; Katz, 1986; Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979; Snowling, 2000; Stanovich & Siegel, 1994; Vellutino et al., 1996; Wagner et al., 1997). The phonological theory of dyslexia suggests that children with dyslexia fail to learn how to read due to difficulties with storing, retrieving, and/or manipulating phonological representations. This difficulty impairs their development of phonological awareness and their acquisition of decoding skills. Decoding deficits affect fluent word recognition, which is an important prerequisite for adequate reading comprehension (Perfetti, Beck, Bell, & Hughes, 1987). Inquiry into this deficit spanned multiple cognitive levels from basic auditory, to speech, and to highest level of phonological awareness. Since phonological awareness is one of the closest cognitive process to reading, it has been most consistently linked to dyslexia. Yet, precisely because of this proximity, it is not clear how much it reveals about the underlying mechanisms of reading deficits in dyslexia.

#### 1.4.1.1 Phonological processing

It has been suggested that phonological deficits in dyslexia stem from less precise categorical perception. In natural speech, phonemes are embedded in syllables that form a continuous speech stream. Consistent identification of phonemes requires both generalization and precision. Instantiations of phonemes tend to vary across contexts (e.g., across speakers and words), requiring generalization across different instances of the same phoneme. At the same time, phoneme identification requires precision in distinguishing between phonemes that differ in subtle acoustic properties (e.g., /b/ from /d/). This ability to “draw” perceptual boundaries between phonemic units, one that is robust enough to ignore within-category differences, is called categorical speech perception (e.g., Liberman, Harris, Hoffman, & Griffith, 1957).

Studies of categorical perception of speech have demonstrated worst performance on such tasks in individuals with dyslexia (e.g., Hazan, Messaoud-Galusi, Rosen, Nouwens, & Shakespeare, 2009; Liberman et al., 1957; Maassen, Groenen, Crul, Assman-Hulsmans, & Gabreëls, 2001; Messaoud-Galusi, Hazan, & Rosen, 2011; Vandermosten et al., 2010). These findings indicate that individuals with dyslexia have a diminished capacity to identify and discriminate between phonemes, resulting in poor speech perception. Interestingly, such studies also demonstrated increased sensitivity for within category differences in dyslexia, suggesting that allophonic perception may interfere with correct phonemic categorization in this disorder (Bogliotti, Serniclaes, Messaoud-Galusi, & Sprenger-Charolles, 2008; Breier et al., 2001; McArthur, Atkinson, & Ellis, 2009;

Noordenbos, Segers, Serniclaes, Mitterer, & Verhoeven, 2012; Serniclaes, Sprenger-Charolles, Carré, & Demonet, 2001). Importantly, poor categorical perception of speech has also been reported in pre-readers at risk for dyslexia (Boets, Vandermosten, Poelmans, Luts, Wouters, & Ghesquière, 2011; de Bree, Wijnen, & Gerrits, 2010; Hakvoort et al., 2016; Noordenbos et al., 2012).

The path from categorical speech perception to phonology, and subsequently to reading, has been supported by studies that found correlations among these skills (Boets, Vandermosten, Poelmans, Luts, Wouters, & Ghesquiere, 2011; but see Hakvoort et al., 2016; Hazan et al., 2009; Noordenbos et al., 2012). Not all studies, however, have found significant group differences in categorical perception (Hazan et al., 2009; Ramus, Rosen, et al., 2003; Robertson, Joanisse, Desroches, & Ng, 2009; White et al., 2006), suggesting that this deficit may characterize some individuals with dyslexia or only under specific conditions (e.g., with synthetic but not with natural speech; Blomert & Mitterer, 2004). Others have suggested that speech perception deficits could be epiphenomenal to other deficits in dyslexia such as phonological awareness, working memory, and attention (Manis et al., 1997; Rosen, 2003).

#### 1.4.1.2 Auditory processing

Phonological studies have also been inconclusive on whether speech perception deficits are due to downstream impairments in processing of lower-level auditory information. Dyslexia has been associated with atypical auditory processing of rapid auditory transitions and of rise time cues (a rate of amplitude change) in non-linguistic stimuli (Beattie & Manis, 2012; Goswami, Fosker, Huss,

Mead, & Szucs, 2011; Hamalainen, Leppanen, Torppa, Muller, & Lyytinen, 2005). Such acoustic information provides crucial segmental and suprasegmental cues for discriminating among linguistic units and detecting speech rhythm and prosody (Cutler, Dahan, & Van Donselaar, 1997; Fletcher, 2010). Not all studies were able to establish auditory temporal processing impairments in dyslexia (Ahissar, 2007; Breier, Gray, Fletcher, Foorman, & Klaas, 2002; Mody, Studdert-Kennedy, & Brady, 1997; White et al., 2006) however, and other studies attributed the findings of poor auditory processing of rapid acoustic transitions to comorbid deficits with attention (Breier et al., 2002; Waber et al., 2001).

Thus, there is strong support for speech perception deficits in individuals with dyslexia, but evidence for an underlying deficit in basic auditory processing is weak and equivocal. Furthermore, there is still inadequate evidence for causal links between speech processing deficits and reading impairment.

#### 1.4.1.3 Phonological awareness

Studies of phonological awareness deficit in dyslexia have met many of the gold standards of etiological examinations: studies in pre-readers, longitudinal examinations, training studies, neuroimaging evidence, and studies across different orthographies (Goswami, 2015). Poor phonological awareness has been demonstrated in individuals with dyslexia across different ages. Pre-reading deficit in phonological awareness have associated with poor longitudinal reading outcomes across multiple studies (e.g., Catts, McIlraith, Bridges, & Nielsen, 2017; Elbro, 1996; Scarborough, 1989). Persistent phonological awareness deficits have been shown in poor readers and compensated adults with a history of dyslexia

(Bruck, 1992; Wilson & Lesaux, 2001). The findings of such deficits have held consistent even when comparing children with dyslexia with reading level matched controls (Blachman, 2000; Fletcher et al., 1994; Shankweiler et al., 1979; Share & Stanovich, 1995; Stanovich, 1994; Velluntino, 2004). As reviewed in the previous section, neuroimaging studies of dyslexia most consistently report structural and functional alterations in the dorsal phonological system as early as in infancy.

Although most researchers agree that phonological awareness deficits characterize some, if not most, individuals with dyslexia, it has been argued that since phonological awareness is such a high-level skill, impairment on this skill is meaningless for informing about dyslexia etiology (Stein, 2018). Furthermore, there are ongoing debates on whether phonological deficits represent underspecified stored phonological representations or deficits with the retrieval of otherwise typical phonological information (Boets, 2013; Ramus & Szenkovits, 2008). Somewhat related debate is whether the rapid automatized naming (RAN) deficit represents a separate etiological mechanism in dyslexia or whether it is an extension of the phonological retrieval deficit (Ramus & Szenkovits, 2008).

#### **1.4.2 RAN/Phonological retrieval**

Based on hypotheses by Geschwind, Denckla (1972) first reported among impaired readers is an “unusual hesitancy” in rapidly naming a series of colors. Subsequently, in the original *double deficit hypothesis*, Wolf and Bowers (1999) argued that phonological awareness and rapid automatized naming (RAN) deficit, an impairment in rapid retrieval, or naming speed, represent two independent deficits in dyslexia that can converge in the most impaired children. Naming speed

deficits have been demonstrated across multiple languages, with increased salience in regular languages, as compared to phonological awareness deficits (Norton & Wolf, 2008). Accordingly, RAN has been called the “universal” predictor (Tan et al., 2005). Although some older studies showed moderate to high correlations between phonological awareness and RAN (Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997), most recent evidence across languages indicates that RAN and phonological awareness are largely independent systems with some overlap and with each contributing independent variance to specific literacy skills (see a review by Norton & Wolf, 2012). Behaviorally, RAN measures the speed required to name serially presented, visual stimuli (e.g., letters, colors, etc.), with some subtests involving set switching between these categories (RAS). Cognitively, RAN represents the time it takes for the attentional, perceptual, linguistic, and motoric processes to be integrated. Neurologically, RAN was proposed to represent a microcosm of the reading circuit, requiring timing, accuracy, and synchrony within and across components shared by reading and naming (Wolf, Bowers, & Biddle, 2000). Although RAN measures are widely used in reading research, particularly in prediction (Scarborough, 1998), RAN’s exact role in dyslexia continues to be debated.

Due to challenges involved in conducting rapid naming inside MRI scanners, only two studies investigated the neural substrates of RAN, using covert naming (Misra, Goswami, & Pandav, 2004; Misra et al., 2009). In this study activation during naming was observed in a distributed network, including inferior frontal and temporo-parietal regions and posterior fusiform gyrus. Two other

studies have demonstrated distinct associations of RAN and phonological awareness with task-related activation during word reading (Turkeltaub et al., 2003) and visual rhyming (Norton et al., 2014). Studies of neuroanatomy reported associations of grey matter volume (Eckert et al., 2003; He, Xue, Chen, Lu, & Dong, 2013; Jednoróg, Gawron, Marchewka, Heim, & Grabowska, 2014) and of white matter coherence (Deutsch et al., 2005; Saygin et al., 2013) with PA and RAN with multiple clusters in the occipital, temporal, parietal, and frontal cortices.

It has been proposed that RAN indexes a subset of skills associated with phonological awareness. Phonological awareness can be divided into two distinct cognitive functions: phonological retrieval and phoneme representation. Dissociation between neural regions supporting these functions would be expected. The studies reviewed so far considered both functions together, allowing for the possibility that the neuroanatomical independence between RAN and phonological awareness is due to this dissociation. Thus, it is plausible that phonological awareness deficit is a phonemic representation deficit, but RAN deficit reflects compromised phonological retrieval. For example, in the *core phonological deficit model*, Ramus (2003) conceptualizes RAN deficits as a failure to access and retrieve phonologically based information fluently. Recent fMRI findings challenge this conceptualization: A study in adults with and without dyslexia investigated patterns of brain responses to auditory-presented bi-syllabic pseudowords (Boets, 2013). Individuals with dyslexia demonstrated similar neural activation patterns in response to phonological similarity in the dorsal temporo-parietal regions, but atypical functional and structural connectivity patterns among these regions as

compared to typical readers. This suggests normal activation to phonological representations in dyslexia, thereby supporting the hypothesis that phonological deficit in dyslexia is not a result aberrant phonemic representations, but rather a result of impaired access to otherwise typical representations (Ramus & Szenkovits, 2008). Interestingly, individuals in this study showed poor performance on both RAN and phonological measures, demonstrating the double-deficit profile. Thus, it is not clear whether the poor access finding was related to phonological, RAN, or both deficits.

In summary, behavioral and neuroimaging studies provide considerable evidence both for the segregation of phonological awareness and RAN deficits in dyslexia, and for the independence of at least two major profiles or subtypes. They also illumine why there is some overlap. Like any language-based task, RAN involves phonological retrieval of verbal labels, which explains moderate correlations between the two, but does not account for all the variance in slower naming speed in dyslexia. Still missing are the cognitive and neuroimaging investigations of the two profiles in pre-reading children, important for understanding both the distinct and the shared development of the neural components of RAN and phonological awareness.

### **1.4.3 *Visual***

Historically, dyslexia, termed “word blindness,” has been viewed as stemming from deficits in the visual domain. Multiple visual theories of dyslexia have been proposed and subsequently empirically dismissed (Vellutino, 2004). Most prominently, dyslexia was thought to represent a deficit in reversing letters



and words (optical reversibility—Orton, 1925). More recent theories have proposed low-level magnocellular deficits and higher-level visual attentional deficits in dyslexia.

#### 1.4.3.1 Low-level

The visual processing in reading occurs through two anatomically and functionally independent systems: magnocellular and parvocellular. The parvocellular system communicates the “what” information by transmitting the details of visual targets to the ventral route that passes from the primary visual cortex towards the visual word form area (Dehaene & Cohen, 2011), which is involved in letter-form and word-form recognition. The magnocellular system provides “where” information from the primary visual cortex to parietal attention regions and is specialized for timing in visual events. It plays a crucial role in signaling letter order and directing attention and eye movements (Bosse, Tainturier, & Valdois, 2007).

Visual deficits in dyslexia have been reported on tasks that require magnocellular system-related processing of high temporal and low spatial frequency stimuli, but the magnocellular deficit explanation of dyslexia received mixed empirical support (Stein, 2014). Several studies indicate neuroanatomical abnormalities in the magnocellular brain network, including a magnocellular layer of the lateral geniculate nucleus (Livingstone, Rosen, Drislane, & Galaburda, 1991), the MT/V5 region (Demb, Boynton, & Heeger, 1998; Eden et al., 1996; Heim et al., 2010; Olulade et al., 2013), and the posterior parietal lobe that receives

input from the MT/V5 regions (Lobier, Peyrin, Pichat, Le Bas, & Valdois, 2014; Zhang, Whitfield-Gabrieli, Christodoulou, & Gabrieli, 2013).

Causal evidence for the magnocellular deficit in dyslexia, however, is limited. It has been argued that these deficits are epiphenomenal to reading impairment (Ramus, 2004). Indeed studies that disrupted neural activity in the magnocellular pathway through transcranial magnetic stimulation (TMS) showed that single-word reading remained intact (Laycock, Crewther, Fitzgerald, & Crewther, 2009; Rauschecker et al., 2011). Furthermore, decreased activity in the magnocellular regions during a motion perception task in children with dyslexia was present when compared to age-matched, but not to reading-level-matched controls, suggesting that visual magnocellular deficits in dyslexia are a consequence rather than a cause (Olulade et al., 2013). Nevertheless, several studies questioned the links between reading experience and motion processing, failing to find significant relationship between the two using psychophysical and brain imaging measurements (Amitay, Ben-Yehudah, Banai, & Ahissar, 2002; Paulesu, Danelli, & Berlingeri, 2014; Skottun, 2000). Furthermore, 7-12 year old children with dyslexia have failed to demonstrate improvements in motion sensitivity despite improved reading performance after receiving 8 weeks of intensive reading intervention (Joo, Donnelly, & Yeatman, 2017). These findings were interpreted by the authors to suggest that magnocellular deficits are not the result of reduced reading experience. It seems unlikely, however, that an 8-week intervention program can ameliorate the vast gaps in reading experience, estimated at one

thousand more words a week read by typically-reading children as compared to children with dyslexia since 1<sup>st</sup> grade (Stanovich, 1986).

#### 1.4.3.2 Visual attention deficits

Visual attention plays a significant role in reading. Attentional mechanisms highlight the visual stimuli, direct gaze, extract information, bind visual features, and shift to the next stimuli (Pammer, 2014). It is plausible that attentional deficits could impair the on-line, perceptual processes during reading, undermining one or all them. Evidence that some individuals with dyslexia show attentional, but not phonological, deficits gives some support to this modality-specific view for visual processing (Heim et al., 2008). In a study of Italian kindergartners, two visual attention tasks - a visual search task and a spatial cuing task - reliably predicted 2<sup>nd</sup> grade reading outcomes better than phonological and RAN measures. In other studies, children with dyslexia demonstrated visual attention deficits even when compared with reading-level matched controls (Bosse & Valdois, 2009; Zoubrinetzky, Bielle, & Valdois, 2014). Several training studies targeting visual attentional mechanisms demonstrated reading enhancements (Franceschini et al., 2013; Valdois et al., 2014) and increased activation of cortical regions supporting visual attention processing in individuals with dyslexia (Valdois et al., 2014).

Several studies have demonstrated the importance of attentional system for successful reading. A functional resting state connectivity study demonstrated an increase in connectivity between attentional regions and the visual word form area with increased reading skill (Vogel, 2013). A study in children with dyslexia reported reduced resting state correlations among frontal and parietal attention

systems and visual word form region as compared with typical readers (Koyama, 2011). Another study examined functional co-activation between bilateral medial frontal gyrus and left anterior and posterior language regions during a visual rhyming fMRI task (Cao, Bitan, & Booth, 2008). Confirming the role of top-down attention mechanisms in regulating language processing, modulatory effects from attentional to phonologic systems were positively correlated with reading ability. Importantly, however, this pattern was absent in children with dyslexia, suggesting possible compromised top-down attentional modulation in reading-related tasks.

To summarize, attention plays an important role in reading and neural atypicalities within this visual attentional network, and impaired behavioral performance on visual attention tasks suggests a visual attentional deficit as one of several, potential contributors to reading difficulties in at least some individuals with dyslexia. It has been argued, however, that visual attention is a skill significantly enhanced by reading experience and therefore these deficits are experientially driven rather than etiological causes of dyslexia (Goswami, 2015). It has also been suggested that these deficits are specific to shallow orthographies where phonological demands on reading are less severe (Georgiou, Papadopoulos, Zarouna, & Parrila, 2012).

#### **1.4.4 *Temporal/motor coordination***

Temporal deficits in individuals with dyslexia have been reported for pure sensory tasks in the auditory, visual (discussed earlier), and motor domains (Nicolson, Fawcett, & Dean, 2001) and hypotheses have been proposed (Wolf, 1991; Goswami, 2011; Breznitz, 2008). Wolf (1991) have suggested that a failure

in underlying temporal processing mechanisms impedes processing in the cognitive and linguistic subprocesses of reading affecting the development of automaticity for reading. Breznitz (2008) suggested that temporal asynchrony between the varied systems involved in reading underlies slow, inaccurate reading development. Breznitz and colleagues demonstrated an atypically wide temporal gap in the speed of processing between visual and auditory modalities. On linguistic tasks, delays between phonological and orthographic processing strongly predicted reading accuracy (Breznitz & Misra, 2003). This gap is indicative of a neural asynchrony that would impact the temporal precision required in automatic grapheme-phoneme correspondence.

Corriveau and Goswami (2009) described the dyslexic brain as “in tune, but out of time” and proposed an alternative view. In the Temporal Sampling Framework of dyslexia, Goswami (2011) argued that impaired phase-locking of the oscillatory activity in the auditory cortex of individuals with dyslexia results in poor neural entrainment to auditory input at the syllabic processing rate. This, in turn, impacts speech segmentation due to poor coding of the low-frequency envelope information, subsequently undermining phonological development. Goswami further hypothesized that similar mechanisms may be responsible for poor temporal processing in the visual modality. In a similar vein, Kraus and colleagues proposed the Neural Synchrony Hypothesis to account both for delays in cortical and subcortical neural responses to sound for children with language impairments and for less robust neural speech representations in struggling readers (Tierney & Kraus, 2013b). Both these theories provide a framework for the burgeoning

research into the relationship between rhythm and phonology and the promise of rhythm interventions in contributing to remediation (Moritz, Sasha, Papadelis, Thomson, & Wolf, 2012; Thomson & Goswami, 2008).

#### 1.4.4.1 Musical rhythm

A growing number of studies report relations between non-linguistic musical rhythmic processing abilities and reading-related skills in children and adults. Research on dyslexia has suggested that rhythm skills are impaired in this population (Banai et al., 2009; Bishop-Liebler, Welch, Huss, Thomson, & Goswami, 2014; Dellatolas, Watier, Le Normand, Lubart, & Chevrie-Muller, 2009; Flaughnacco, Lopez, Terribili, Zoia, Buda, Tilli, Monasta, Montico, Sila, & Ronfani, 2014; Goswami & Leong, 2013; Hornickel & Kraus, 2013; Huss, Verney, Fosker, Mead, & Goswami, 2011; Lee, Sie, Chen, & Cheng, 2015; Wolff, 2002). Individuals with dyslexia tap more variably to a beat, show poor rhythm discrimination skills, and show delayed neural response to rhythmic cues. Moreover, musical training in individuals with dyslexia, specifically targeting rhythmic perception and production, has been shown to result in improved phonological and reading skills (Bhide, Power, & Goswami, 2013; Habib et al., 2016; Overy, 2003; Schön, Magne, & Besson, 2004). This crossover effect of enhanced reading skills afforded through rhythmic training suggests a shared substrate underlying these two cognitive abilities.

One link between the two domains may lie in the temporal processing of sound (Goswami, 2011; Tallal & Gaab, 2006; Tierney & Kraus, 2013b). As previously discussed, reading acquisition depends heavily on phonological awareness that draws on auditory processes involved in analyzing the temporal structure of sound patterns. For example, distinctions between phonemes can involve very subtle timing cues (e.g., voice onset time differences between /b/ and /p/), and in some languages (such as English) syllable duration patterns are important in cueing stress (Greenberg, 2006), which in turn provides statistical cues to word boundaries (Cutler, 2012).

Rhythm is present in both music and speech. Both involve complex sound sequences with systematic patterns of timing, accent, and grouping (Patel, 2008). One salient difference between musical and speech rhythm, however, is that much of the world's music has an underlying time frame of equal intervals that organizes the timing of musical notes (Savage, Brown, Sakai, & Currie, 2015). Ordinary speech does not have this structure: the temporal patterning of linguistic units (e.g., syllables, or words) cannot be subdivided into equal time intervals (Nolan & Asu, 2009). Povel and Essens (1985) define the former type of rhythmic pattern as *metrical* and the latter as *nonmetrical*. The processing of metrical sequences engages predictive mechanisms which support the perception of an underlying beat (van der Weij, Pearce, & Honing, 2017). That is, beat perception involves processing and adapting to regular temporal structure. In music processing, this ability supports synchronized movement to the beat and also provides cognitive scaffolding for the encoding of rhythmic patterns (London, 2012). Suggesting the

involvement of a domain-general cognitive mechanism, beat-based scaffolding has been demonstrated beyond the domain of music. A number of studies have shown that auditory beat-based rhythms can influence the processing of information in visual (Escoffier, Sheng, & Schirmer, 2010) and linguistic (Schön & Tillmann, 2015) domains.

Investigating whether beat processing is uniquely associated to reading development can illuminate the nature of the rhythm deficits in dyslexia. Specifically, if processing of rhythm regularity is uniquely and positively associated with reading development, this could indicate that children who have difficulty with extracting temporal structure from non-linguistic musical sequences, also struggle with benefiting from structural regularities related to reading (Ahissar, 2007). As previously noted, in metrical processing listeners use temporal structure to make predictions about upcoming musical material. In language processing listeners implicitly use phonological, semantic, and syntactic structure to make predictions about upcoming linguistic material (Kuperberg & Jaeger, 2016). Thus, the neurological capacity to form predictions based on contextual structure in rapidly unfolding sequences is common to metrical processing and language processing, even if the former is based on temporal periodicity and the latter is not. If such commonality can explain differences in reading development beyond auditory processing, it could provide a putative mechanism to explain the variety of learning-related deficits demonstrated in dyslexia (reviewed next). Interestingly, earliest theories on links between musical rhythm and dyslexia had to do with sensorimotor coordination and were developed in response to observations of deficits with



motoric automatization in some individuals with dyslexia (Wolff, Michel, Ovrut, & Drake, 1990). These deficits will be reviewed next.

#### 1.4.4.2 Cerebellum

Some theories of dyslexia posit that slow performance reflects poor coordination between sensory and motor systems. This would suggest that slow reading performance in dyslexia may be only the tip of the iceberg of a general deficit in acquiring automatization. Atypical cerebellar structure has been implicated in dyslexia that compromises procedural motor learning and automatization across different functions, including reading. The *cerebellar deficit theory* has some limited support from findings of reduced performance of individuals with dyslexia on tasks considered under cerebellar control, such as dual balance tasks (Needle, Fawcett, & Nicolson, 2006; Yap & Leij, 1994), and implicit motor learning tasks (Stoodley, Harrison, & Stein, 2006). Ramus (2003) estimated that 30 to 50% percent of individuals with dyslexia exhibit these deficits. Other studies, however, suggest that a third moderator, such as comorbidity with ADHD, accounts for the relationship between dyslexia and deficits in balance (Raberger & Wimmer, 2003; Rochelle & Talcott, 2006). Additionally, a study in pre-readers demonstrated that motor skills were not predictive of reading outcomes (Carroll, Solity, & Shapiro, 2016). Unresolved here is increasing evidence for the importance of the cerebellum for fluent reading.

Although the cerebellum was traditionally associated with motor control and skill automatization, recent research emphasized cerebellar involvement in higher cognitive functions. The extensive connectivity of the cerebellum to the

different cortical regions and its functional diversity is well known, but its specific roles in reading are just unfolding. Neuroimaging studies consistently demonstrate cerebellar activation in multiple, reading-related functions, including silent reading and passive language processing, word/letter generation, word stem completion, semantic processing, phonological processing, and verbal fluency tasks (Stoodley & Stein, 2013). Studies of dyslexia demonstrate reduced grey matter volume in the cerebellum, particularly in the right lobule VI, and cerebellar asymmetries are some of the most consistently reported structural differences in individuals with dyslexia (Stoodley, 2014). Importantly, reduced activation in cerebellar lobule VI, discriminated children with RAN deficits from typically reading and phonological deficit groups (Norton et al., 2014).

Thus, increasing neuroanatomical evidence suggests the cerebellum's importance for reading. For example, lobule VI is functionally connected to prefrontal and premotor cortical areas, and inferior frontal and temporo-parietal regions, areas necessary for varied sensorimotor functions in reading (Stoodley & Stein, 2013). A strong case for the role of this region in timing has been made (Spencer & Ivry, 2013). Citing evidence from ERP, TMS, MRI, and brain lesion studies, the authors described the importance of lobule VI in temporal prediction, rhythm synchronization, and time perception, as well as right lobule VI's role during time discrimination, and left lobule VI's role during timed movement generation (Aso, Hanakawa, Aso, & Fukuyama, 2010). The role of the cerebellum in dyslexia remains controversial. It has been suggested, for example, that "the cerebellum might stand unfairly accused, an innocent bystander in the processes

responsible for disordered motor control in developmental dyslexia... [that] reflect a remote effect of neocortical perisylvian damage on cerebellar function” (Zeffiro & Eden, 2001).

#### **1.4.5 *Statistical learning/perceptual adaptation***

Somewhat related to deficits in automatization subsumed under the cerebellar hypothesis, are theories that attribute dyslexia to deficits in cross-domain learning. These theories emerged to explain deficits in tasks requiring perceptual learning of implicit structure and stimulus regularity. For example, individuals with dyslexia have demonstrated deficits with associative learning (Brewer, 1967; Gascon & Goodglass, 1970), cross-modal transfer (Birch, 1962), serial-order processing (Bakker, 1972), pattern analysis and rule learning (Morrison & Manis, 1982), statistical learning (Gabay, Thiessen, & Holt, 2015), perceptual anchoring (Ahissar, Lubin, Putter-Katz, & Banai, 2006). Most of these studies, however, failed to control for differences in verbal working memory, which could have affected learning on these tasks. These deficits have been interpreted in terms of fundamental deficits in extracting and benefiting from implicit structure in sensory signals (Achal, Hoefft, & Bray, 2016; Ahissar et al., 2006; Howard Jr, Howard, Japikse, & Eden, 2006; Vicari, Marotta, Menghini, Molinari, & Petrosini, 2003). One such theory, the *anchoring hypothesis model* of dyslexia, suggests that deficits in automatic extraction of contextual regularities undermine the establishment of robust linguistic models and weaken predictive processing (Ahissar, 2007). This model was developed to account for findings in the auditory domain, but similar hypotheses have been proposed to account for cross-domain deficits (“noise

exclusion hypothesis”, Sperling, Lu, Manis, & Seidenberg, 2005; “neural noise hypothesis”, Hancock, Pugh, Hoefft, 2017; statistical learning, Gabay, Thiessen, & Holt, 2015; Schmalz, Altoe, & Mulatti, 2016), and for the range of attentional deficits (reviewed in previous section) reported in dyslexia.

On the neural level, there has been evidence for reduced neural adaptation to a variety of linguistic and non-linguistic stimuli in children and adults with dyslexia (Perrachione et al., 2017). Individuals with dyslexia have also demonstrated atypical hyper involvement of the fronto-striatal network, a network involved in implicit learning, across different fMRI tasks (Achal et al., 2016; Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan et al., 2009). In EEG and MEG studies of speech processing, individuals with dyslexia exhibited reduced phase-locking of neural oscillations to regular stimuli (Hamalainen, Rupp, Soltesz, Szucs, & Goswami, 2012; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013) as well as less reliable neural representation of speech (Hornickel & Kraus, 2013). While the theory of learning deficits in dyslexia is gaining much support, its critics argue that such a theory would predict pervasive deficits across domains unrelated to reading (e.g., face recognition). Deficits in dyslexia, however, are often specific to reading-related stimuli and do not spill over into unrelated domains.

#### **1.4.6 *General language***

Language skills such as vocabulary, comprehension, and grammar are closely and reciprocally intertwined with reading development. Studies examining infants and pre-reading children with a hereditary risk for dyslexia have identified atypical language development in these children, as they tend to show delayed onset

of talking, shorter mean length of utterances, lower complexity of syllables produced, and poor receptive or expressive vocabulary (Pennington, 2001; Raschle et al., 2015; Scarborough, 1990). Longitudinal studies have further demonstrated the importance of these early language skills for the development of reading. For example, in one study, oral language skills in 4 year old children were associated with reading comprehension at age 12 (van Viersen et al., 2018). However, these symptoms have also been observed in children with a subsequent diagnosis of specific language impairment or speech sound disorders, and could be a consequence of the high occurrence of comorbidities between dyslexia and language disorders (Bishop & Snowling, 2004; Catts, Adlof, Hogan, & Ellis Weismer, 2005). Indeed, it has been suggested that a language component could increase the probability of reading difficulty in at-risk children or alternatively, serve a protective function in these children (Snowling et al., 2003). Several studies found that unaffected at-risk for dyslexia siblings have better language skills than the affected siblings (Moll, Loff, & Snowling, 2013; Pennington & Bishop, 2009; Snowling, 2003). A neuroimaging study that compared children with a poor comprehension profile (representing general language deficits), children with dyslexia and typical readers demonstrated patterns of gray matter dissociation between the profiles, suggesting that language deficits are neuroanatomically independent from dyslexia (Aboud, Bailey, Petrill, & Cutting, 2016). Therefore, delayed language development in some children with dyslexia risk is likely indicative of a cumulative contribution of etiological factors for each disorder, rather than being a specific marker of dyslexia risk.

To summarize, there has been little consensus on the causal nature of the different dyslexia deficits for reading failure (Elliott, 2014). Researchers have challenged even the most agreed-upon deficit of phonological awareness, arguing that due to the reciprocal relationship of phonology with reading, its causal influences on reading development cannot be established (Gori & Facoetti, 2014). Instead, it has been suggested, as for many of the other deficits, that phonological awareness deficits are largely a consequence of a lack of adequate reading experience (e.g., Huettig et al., 2017). Thus, due to the variability of dyslexia symptoms reported in the literature, forming a cohesive definition of dyslexia has been challenging. To resolve some of the ambiguities, it is important to investigate whether the impairments are already present in pre-reading children who go on to be diagnosed with dyslexia, and whether they affect reading acquisition once they start learning to read. Neuroimaging evidence from pre-readers can inform whether particular deficits are linked to distinct neural systems early in development, suggesting their independent emergence. Finally, it is important to consider that reading development occurs in the context of experiences that can drastically shape its course. There are great disparities in environments in which children learn how to read, and these disparities are important to study in relation to dyslexia.

In my dissertation I will adopt a developmental cognitive neuroscience approach to investigating dyslexia etiology. Specifically, *Study 1* will investigate the heterogeneity of dyslexia-risk profiles in pre-reading children longitudinally. Latent profile analysis, a data-driven approach will be used to investigate whether profiles of deficit emerge independently early in reading development. Such early

independence of deficits would point to their distinct contribution to reading failure. *Study 2* will evaluate, using magnetic resonance imaging (MRI), the neuroanatomical characteristics underlying the distinct risk profiles identified in *Study 1*. Investigating whether different dyslexia risk subtypes are associated with distinct structural differences in brain anatomy, before these differences are confounded with reading failure or experience, has important implications for supporting the heterogeneous neurobiological basis of dyslexia. Processing of metrical structure (the underlying grid of equal time intervals) in musical rhythm involves the extraction of implicit regularity. Increasing evidence links rhythm processing with reading, although the unique role of metricality in this relationship has yet to be investigated. Several studies have documented rhythm deficits in individuals with dyslexia raising two possibilities: first is that the deficits are due to underlying impairments with auditory processing, and second is that these deficits are due to impairments in perceptual adaptation. To better understand the underlying mechanisms of musical rhythm deficits in dyslexia, *Study 3* will dissociate the role of temporal regularity processing—i.e., metricality, in the relationship between rhythmic and reading abilities. Finally, in *Study 4*, I will test the environmental influences on early development of brain circuits for reading and longitudinal reading outcomes of at-risk children. Although traditionally dyslexia has been defined based on its independence from environmental influences such as access to adequate schooling, there is strong evidence that environmental factors exert a significant influence on reading development. Nevertheless, there is no clear understanding of how environmental factors interact with dyslexia risk to shape

reading outcomes. The final chapter will report behavioral and neuroimaging evidence that SES modulates risk-reading outcomes.



## **2 Study 1: Longitudinal Stability of Pre-Reading Skill Profiles of Kindergarten Children: Implications for Early Screening and Theories of Reading<sup>3</sup>**

### **2.1 Abstract**

Research suggests that early identification of developmental dyslexia is important for mitigating the negative effects of dyslexia, including reduced educational attainment and increased socioemotional difficulties. The strongest pre-literacy predictors of dyslexia are rapid automatized naming (RAN), phonological awareness (PA), letter knowledge, and verbal short-term memory. The relationship among these constructs has been debated, and several theories have emerged to explain the unique role of each in reading ability/disability. Furthermore, the stability of identification of risk based on these measures varies widely across studies, due in part to the different cut-offs employed to designate

---

<sup>3</sup> Portions of this chapter were originally published as Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles of kindergarten children: implications for early screening and theories of reading. *Developmental science*, 20(5).

risk. We applied a latent profile analysis technique with a diverse sample of 1215 kindergarten and pre-kindergarten students from 20 schools, to investigate whether PA, RAN, letter knowledge, and verbal short-term memory measures differentiated between homogenous profiles of performance on these measures. Six profiles of performance emerged from the data: average performers, below average performers, high performers, PA risk, RAN risk, and double-deficit risk (both PA and RAN). A latent class regression model was employed to investigate the longitudinal stability of these groups in a representative subset of children ( $n = 95$ ) nearly two years later, at the end of 1<sup>st</sup> grade. Profile membership in the spring semester of pre-kindergarten or fall semester of kindergarten was significantly predictive of later reading performance, with the specific patterns of performance on the different constructs remaining stable across the years. There was a higher frequency of PA and RAN deficits in children from lower socioeconomic status (SES) backgrounds. There was no evidence for the IQ–achievement discrepancy criterion traditionally used to diagnose dyslexia. Our results support the feasibility of early identification of dyslexia risk and point to the heterogeneity of risk profiles. These findings carry important implications for improving outcomes for children with dyslexia, based on more targeted interventions.

## **2.2 Introduction**

Developmental dyslexia (henceforth, dyslexia) affects 5-17% of children, with the prevalence rates varying widely across studies depending on the exact definition and measures used for diagnosis (Elliott & Grigorenko, 2014). Dyslexia is a neurological condition characterized by difficulties with accurate and/or fluent

word recognition, poor spelling, and poor decoding abilities (Lyon, Shaywitz & Shaywitz, 2003). Dyslexia is also often associated with impediments in a range of perceptual and cognitive processes important for reading, such as verbal short-term memory, rapid naming, and phonological awareness, as well as differences in the brain regions supporting these processes (Norton, Beach & Gabrieli, 2015). Due to the variability of dyslexia symptoms reported in the literature, forming a cohesive definition of dyslexia has been challenging, and instead a multi-deficit conceptualization of dyslexia is becoming increasingly accepted (Pennington et al., 2012). Traditionally, there has been an emphasis on the independence of dyslexia from other causes that could explain reading failure (i.e., low intelligence, socioeconomic disadvantage, inadequate schooling, or physical disability) (Lyon, 1995). Due to the complex interactions among environmental, cognitive, and neurological factors, however, the merits of such an approach are under considerable debate (Elliott & Grigorenko, 2014).

Similar to the complexity of dyslexia's definition, and potentially because of it, dyslexia remediation efforts have been challenging, with modest effect sizes for interventions ranging from 0.07 to 0.56, according to a meta-analysis (Wanzek & Vaughn, 2007; Wanzek et al., 2013). Converging research points to the importance of early and individualized interventions for at-risk students for improving the effectiveness of remediation (Denton & Hocker, 2006; Flynn, Zheng, & Swanson, 2012; Morris et al., 2012; Shaywitz, Morris, & Shaywitz, 2008; Torgesen, 2000; Vellutino et al., 1996; Wanzek & Vaughn, 2007). A recent study demonstrated that when an intervention is administered in kindergarten and 1<sup>st</sup>

grade it had an effect size of up to 0.84, the highest effect size for intervention administered in 2<sup>nd</sup> or 3<sup>rd</sup> grade, however, was 0.27 (Lovett et al., 2017). Therefore, an important question is whether distinct patterns of pre- reading performance could serve as reliable predictors of particular profiles of dyslexia. The current study, for the first time, implements latent profile analysis (LPA) methods to investigate the heterogeneity of dyslexia risk profiles in pre-reading and early-reading children longitudinally. LPA is a mixture-modeling technique that aims to classify individuals into distinct groups based on individual response patterns.

In order to characterize the heterogeneity and prevalence of latent early literacy profiles as well as their longitudinal stability and distribution across school SES levels, a large sample of kindergarten and pre- kindergarten children from 20 diverse schools was evaluated on measures of early literacy and cognition. LPA was implemented to reveal homogenous profiles of performance and to examine these profiles in relation to reading status (readers or pre-readers) and school SES. Latent class membership was then used to predict end- of-1st-grade reading abilities of a subsample of children.

### **2.2.1 *Early identification of dyslexia risk***

The cascading effects of early reading ability have been well documented: children who are early readers receive more print exposure and develop superior automaticity, comprehension skills, vocabulary, and cross-domain knowledge (Mol & Bus, 2011; Stanovich, 1986). In contrast, children who lag behind in their early reading abilities receive fewer opportunities to enhance their vocabulary or to develop reading comprehension strategies (Paris & Oka, 1989). In addition, these

children tend to acquire negative attitudes about reading, and often remain poor readers throughout their school years and beyond, never achieving fluent reading (Ferrer et al., 2015; Lyon et al., 2003). Thus, an important aim of reading studies is to determine which pre-reading measures predict dyslexia in order to offer the potential to effectively intervene and prevent reading failure.

Several pre-reading measures, when administered in kindergarten, are predictors of later reading abilities (for a review see Ozernov-Palchik & Gaab, 2016). These measures include letter name and letter sound knowledge (LSK), phonological awareness (PA), verbal short-term memory (VSTM), and rapid automatized naming (RAN) (Pennington & Lefly, 2001; Scarborough, 1989; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004; Wolf, Bally, & Morris, 1986). PA is the meta-understanding of the sound units of oral language, measured by the ability to manipulate linguistic sounds independent of meaning (Stahl & Murray, 1994). Short-term memory is a separate, but related, construct that measures the capacity to maintain and process information (e.g., digits, pseudowords) for a short period of time (Stahl & Murray, 1994; Stanovich, Cunningham, & Feeman, 1984). VSTM, a short-term memory for linguistic (verbal) material (e.g., a string of letters), is sometimes subsumed under PA, since both involve phonological processing, but there is evidence that it represents a distinct construct and accounts for unique variance in reading (Mann, 1984; Scarborough, 1998). RAN is the ability to rapidly retrieve the names of visually presented, familiar items in a serial array (e.g., objects, colors, numbers, or letters,

or a combination of these in rapid alternating stimulus formats) (Denckla & Rudel, 1976; Norton & Wolf, 2012).

Although these measures demonstrate a strong association with later reading performance, studies that used kindergarten performance on these measures to ascertain risk for dyslexia showed limited success in predicting which children truly develop dyslexia, with false positives ranging from 20% to 60% (Jenkins, 2002; Torgesen, Rashotte, & Alexander, 2001) and false negatives from 10% to 50% (Catts, 1991; Scarborough, 1998; Torgesen, 2002). These findings prompted suggestions of delaying identification until at least 1<sup>st</sup> grade, when language-based and literacy-based activities at home are less influential and measures can be more reading-specific (Fletcher et al., 2002). Demonstrating stability in risk-status classification between kindergarten and 1<sup>st</sup> grade, however, has important implications for early diagnosis of dyslexia risk and, subsequently, early remediation.

### ***2.2.2 Theories of dyslexia and implications for diagnosis and treatment***

In the double-deficit view of dyslexia, deficits in PA and RAN represent distinct deficits across different languages; further, the combination of both deficits in some individuals can be additive, creating reading impairment that is more severe than it is in individuals with single deficits (Compton, Defries, & Olson, 2001; Kirby, Parrila, & Pfeiffer, 2003; O'Brien, 2012; Papadopoulos, Georgiou, & Kendeou, 2009; Wimmer, Mayringer, & Landerl, 2000; Wolf & Bowers, 1999). Others contend, however, that rather than representing distinct dyslexia subtypes,

both PA and RAN deficits represent the failure to fluently access and retrieve phonological information (Lervåg & Hulme, 2009; Ramus, 2003).

One of the main challenges to the evidence for the double-deficit hypothesis is the inconsistency in criteria applied to designate dyslexia diagnosis across studies (Vukovic & Siegel, 2006). The manner in which deficit groups are defined can strongly influence the results, and thus the understanding of how these deficits relate to reading development. For example, some studies define dyslexia based on an IQ–achievement discrepancy model that designates dyslexia as a low performance on reading assessments relative to performance on tests of general intelligence (e.g., IQ), while others do not. A similar issue concerns the use of arbitrary cut-off criteria to designate risk. For example, across studies, the threshold used to define risk spans the range of the 10th to 25th percentile, or 1 to 2 standard deviations below the standardized mean performance on reading tests. Due to the lack of consensus on the definition of risk, the cut-off method may impose an artificial structure onto data and bias the interpretation of results (Catts, Compton, Tomblin & Bridges, 2012; Fletcher et al., 2002; Francis, Fletcher, Stuebing, Lyon, Shaywitz et al., 2005; Waesche, Schatschneider, Maner, Ahmed & Wagner, 2011). Another consideration is the moderate correlation between PA and RAN that has been shown to impede the methodological validity of classifying children into predetermined discrete PA and RAN deficit subgroups (Compton et al., 2001; Schatschneider, Carlson, Francis, Foorman & Fletcher, 2002).

As an alternative to predefining risk-group membership, several studies used latent analysis methods to ascertain reading profiles or dimensions within

large samples of already-reading children. In one study, LPA was used to characterize a large sample of 9-year-old Swedish children on reading performance measures (i.e., reading of continuous texts, reading of document texts, word reading, and reading speed) (Wolff, 2010). Eight stable profiles of readers emerged: (1) high performance; (2) average performance; (3) poor document (e.g., tables, graphs) reading; (4) average decoding, average fluency, poor comprehension (hyperlexic); (5) low decoding, poor fluency, low comprehension (garden-variety poor readers) (6) low decoding, low fluency, poor comprehension (garden-variety poor readers); (7) low-average decoding, low-average fluency, low comprehension; and (8) low decoding, poor fluency, average comprehension (dyslexic).

In a longitudinal study, latent class modeling was also used to identify distinct subtypes of reading development in a large sample of children tested two times per year in the 1<sup>st</sup> and 2<sup>nd</sup> grades (Torppa et al., 2007). Several groups of readers emerged based on children's performance on single word identification, reading fluency, and reading comprehension measures: (1) poor readers, (2) slow decoders, (3) poor comprehenders, (4) average readers, and (5) good readers. These studies support the use of data-driven analysis methods for identifying homogenous profiles of reading and suggest that the heterogeneity of reading development is present early in schooling.

In another approach, a taxometric method was applied to identify the latent structure of reading performance in a large sample of 6–8-year-old children who were identified as dyslexic based on IQ–achievement discrepancy or simply low reading achievement (O'Brien, 2012). This method allows for revealing latent



categorical traits, ‘taxons’, rather than dimensional classes of the condition of interest in the data. Results from the analysis confirmed the double-deficit view of dyslexia.

### **2.2.3 *Identification of dyslexia risk***

However, these results depended on how dyslexia was defined, as the association between RAN and reading emerged in the IQ–achievement discrepancy group, but not in the low reading achievement group. These findings demonstrate the challenge of interpreting results based on pre-established definitions of dyslexia.

The above studies using group classification methods investigated older children who were already reading. From a theoretical perspective, in order to argue that a particular subtype is a core deficit, it is important to demonstrate that the deficit is present prior to reading instruction and is not an artifact of differential influences of reading development, reading instruction, or a phonological awareness deficit (Goswami, 2015). From an applied perspective, the application of these studies to early identification is, therefore, limited.

### **2.2.4 *Longitudinal stability of risk classification***

Another important aspect of group classification is the stability of reading profiles across time. Despite the importance of validating classification methods longitudinally, few studies have investigated the stability of specific risk subtypes over time. Instead, most longitudinal investigations have focused on the long-term associations of the early literacy components (i.e., testing whether a particular

measure at time 1 correlates with a reading outcome at time 2 (e.g., Scarborough, 1998), or on retrospective investigations of individuals with an existing dyslexia diagnosis (i.e., evaluating performance at time 1 based on outcomes at time 2, e.g., Catts & Weismer, 2006).

The longitudinal stability of PA, RAN, and double- deficit (DD) risk classifications has been investigated in only two studies of pre-readers to date (Spector, 2005; Steacy, Kirby, Parrila, & Compton, 2014). Neither study used data-driven methods, but instead applied a predefined cut-off to determine risk. In one study, pre-reading 1<sup>st</sup>-grade students were classified (using 1 SD below mean criterion) as typical, PA deficit, RAN deficit, or DD. These groups exhibited low group membership stability from the beginning to the end of 1<sup>st</sup> grade (less than 50% accuracy) (Spector, 2005). In another study, kindergarten students were characterized into the same groups using a different cut-off criterion (25th percentile) and were followed longitudinally until the fall of 2<sup>nd</sup> grade (Steacy et al., 2014). Group membership in this study was highly stable (over 70% accuracy), which might reflect the larger range of scores below the cut-off. Since these two studies applied different criteria to establish risk, the inconsistency in findings could be due to the issues of pre-classifying students into risk groups using cut-off scores. Thus, a data-driven longitudinal approach is necessary to ascertain the stability and utility of a multi-deficit model for dyslexia risk identification.

### **2.2.5 Cognitive and environmental influences**

The complexity of risk identification is underscored by the multiple cognitive and environmental components that interact with reading ability and

disability. Two of these components that have been strongly linked to reading performance are intelligence (IQ) and socioeconomic status (SES). Historically, dyslexia has been diagnosed based on a reading achievement and IQ discrepancy model. Several studies have shown that the core mechanisms of dyslexia are consistent regardless of IQ (Siegel, 1989, 1992; Tanaka, 2011), although other studies have demonstrated different patterns of reading profiles based on IQ (Ferrer, 2010; Morris, Stuebing, et al., 1998; O'Brien, 2012). The relationship between SES and reading achievement is complex, as SES indexes a broad range of environmental factors. Nevertheless, school-level SES (commonly measured by the percentage of children eligible for free/ reduced lunch within each school) (Caldas & Bankston, 1997) correlated strongly with reading performance (e.g., Scarborough, 1998; Sirin, 2005). Therefore, both school- SES and IQ are important to consider when investigating reading development.

#### **2.2.6 *Current study***

The current study aimed to examine: (1) the heterogeneity and prevalence of latent early literacy profiles among kindergarten students; (2) the stability of latent class membership across two time points (i.e., the beginning of kindergarten and the end of 1<sup>st</sup> grade); and (3) the latent profile distribution across school SES levels. In using LPA and latent class regression methods, we are controlling for some of the major issues raised in the research to date on the presence of single or multiple core deficits in children with dyslexia.

## **2.3 Methods**

### **2.3.1 *Participants and data collection procedures***

Participating children were recruited from 20 schools in New England. Schools varied in their urbanicity and socioeconomic status and included public district (30%), public charter (20%), private (10%), and Catholic (40%) schools. Schools were classified into three SES groups based on the percentage of students receiving free or reduced lunch, as reported by the school: high SES (0-5% of students, 8 schools), middle SES (12-30%, 6 schools), and low SES (32-79%, 6 schools). Permission and informed consent letters were sent to the parents of kindergarten and pre-kindergarten children to obtain parental permission for their children to participate. Children whose parents provided written informed consent and who provided verbal assent completed a 30-45 minute assessment battery. Assessments were administered by trained research assistants and speech-language pathology students on a 1:1 basis. In total, 1,433 English-speaking children were tested at the end of pre-k or beginning of kindergarten (Year 1). Testing was completed over three years, and therefore the final sample included three cohorts of students. Only children with valid and complete data were included in the current analysis. The final sample included 1,215 participants (48% males) with diverse racial (69% Caucasian, 24% African-American/Black, 6% Asian, 1% other) and ethnic (12% Hispanic/Latin) backgrounds. A subset of these children ( $n = 95$ , 49% male; 79% Caucasian, 20% African-American/Black, 1% Asian; 5% Hispanic) was followed longitudinally as part of a larger neuroimaging study and was assessed again at the end of 1st grade (Time 2). Children were recruited to maintain a representative composition to that of the larger sample in regards to gender, age,

ethnicity/race, school type, and behavioral scores. Children with kindergarten IQ scores below 80 and/or who did not speak fluent English, and/or who were born pre-term were excluded from longitudinal analysis.

### **2.3.2 Measures**

Group performance on the age-standardized scores of the measures below is summarized in Table 1.

#### **2.3.2.1 Classification variables (the pre-kindergarten and kindergarten screening battery)**

**Phonological awareness (PA) and phonological working memory (PWM).** Three subtests from the Comprehensive Test of Phonological Processing (CTOPP, Wagner, Torgesen, & Rashotte, 1999) were administered, (1) Elision: the child repeats a word after removing a given syllable or sound; 2) Blending Words: the child blends sounds together to make a real word; (3) Nonword Repetition (NWR): the child repeats a nonsense word. The mean of Elision and Blending scores were used to calculate the PA composite score.

**Rapid automatized naming (RAN).** The Colors and Objects subtests of the Rapid Automatized Naming/Rapid Alternating Stimulus (RAN/RAS) tests (Wolf & Denckla, 2005) were administered. The child names an array of familiar items (colors or objects) on the page as quickly and accurately as possible. The raw score is the time to name all items.

**Word ID.** The Woodcock Reading Mastery Tests - Revised-Normative Update (WRMT-R/NU, Woodcock) was administered to some of the children (65%) and the Woodcock Reading Mastery Tests, Third Edition (WRMT-III,

Woodcock, 2011) was administered to other children (35%) based on their cohort of participation. For both tests, the Word ID subtest assesses single word reading skills. The child reads aloud single words of increasing difficulty within 5 seconds. Scores from the two editions were used interchangeably in the analysis as items early in the test were similar. Children were considered readers based on a raw score of 3 or higher, and non-readers based on a raw score lower than 3. This criterion was chosen based on the median score of 3 for the sample.

**Letter sound knowledge (LSK).** The Letter Sound Knowledge subtest from the York Assessment of Reading for Comprehension (YARC, Stothard, Hulme, Clarke, Barmby, & Snowling, 2010) assesses knowledge of letter sounds. The scores were normed based on the sample distribution in the current study.

**Non-verbal IQ (IQ).** The Kaufman Brief Intelligence Test, Second edition (Kaufman & Kaufman, 2004). Matrices subtest assesses nonverbal matrix reasoning skills, specifically, the understanding of relations between either concrete stimuli (pictures of objects) or abstract stimuli (e.g., designs or symbols).

2.3.2.2 1<sup>st</sup> Grade (Year 2) assessment included the same measures as in kindergarten and additional measures:

**Phonemic decoding (Word Attack).** The Word Attack subtest of the WRMT-III was administered to measure the ability to apply phonic and structural analysis skills to pronounce unfamiliar words (Woodcock, 2011). The child reads non-words of increasing complexity.

**Sight word recognition (SWE) and phonemic decoding (PDE).** Test of Word Reading Efficiency (TOWRE-2, Torgesen, Wagner, & Rashotte, 1999). Sight Word Efficiency and Phonemic Decoding subtests were administered to

measure word reading accuracy and fluency. A child reads real words or non-words as accurately and as quickly as possible within 45 seconds.

**Reading comprehension and fluency.** For the Gray Oral Reading Test-5th Edition (GORT-5, Wiederholt & Bryant, 2012), the child reads several stories aloud and then answers questions based on these passages. Scores are determined for comprehension (number of correct comprehension responses), reading accuracy (number of oral reading errors only for the oral reading paragraph), and reading fluency (combination of the rate and accuracy score).

**Processing speed (PS).** The Wechsler Intelligence Scale for Children – IV Coding subtest (Wechsler, 2003). A child copies symbols that are matched with simple shapes within a two-minute time limit.

**Spelling (Spell).** The test of Written Spelling (TWS-4, Larsen, Hammill, & Moats, 1999) assesses expressive spelling skills. The experimenter reads a word and a child is asked to write it on paper.

### **2.3.3 *Latent profile analysis***

A Latent Profile Analysis (LPA) approach was employed to identify homogenous subgroups (i.e. profiles) of children based on reading-related variables. Unlike variable-centered approaches (e.g., exploratory factor analysis) that seek to identify correlations between variables of interest, LPA is a person-centered approach that groups individuals by the probability of their response patterns on each of the latent profile indicators. The interpretation of each profile is derived from those probabilities. Specifically, the model-fitting process begins with a one-profile model (i.e. a model in which all readers are hypothesized to

demonstrate a single, homogeneous profile) to which additional profiles are added one at a time, and statistical tests are conducted at each step to determine if the additional profile significantly improves the goodness of fit of the model. Simulation studies in the statistical literature have found that these tests are robust and specific in determining when latent profiles can and cannot be differentiated in the population (e.g., Enders & Tofighi, 2007; Lo, Mendell, & Rubin, 2001). Several criteria are employed for testing optimal fit. These include: (a) magnitude of  $R^2$  values; (b) correct classifications versus misclassifications, e.g., in the longitudinal prediction between original class cases and their representation in the predicted classes; (c) significance of predictors; (d) significant reduction in the likelihood ratio test  $L^2$  when comparing nested models; and, (e) acceptable entropy values. For continuous indicators as those involved in the present study, each latent profile was assumed to have its own mean and variance estimates as shown below:

$$f(y|\theta) = \sum_{\tau=1}^T \pi_{\tau}^x f(y|\mu_{\tau}, \Sigma_{\tau}) \quad (\text{Equation 1})$$

The distribution of a dependent variable  $y$  is a function of a set of unknown parameters  $\theta$ . In the right side of the equation  $\pi$  defines the probability of person  $x$  belong in latent profile  $\tau$ . Each latent profile has its own mean ( $\mu_{\tau}$ ) and variance and covariance estimates of the latent profiles ( $\Sigma_{\tau}$ ).

In the present study 1-7 profile models were fit to the data and the superior model fit was judged as a function of differences in the likelihood ratio between nested models, using the unbiased bootstrap distribution (Magidson & Vermunt, 2002). Furthermore, parsimony was taken into consideration by selecting the model with the smallest AIC or BIC values, due to the fact that the likelihood ratio (LR)



test will likely be influenced by the large sample size. The level of significance was set to 5% (Tofighi & Enders, 2008).

Sample size estimation in latent profile models involves comparing models and thus relates strongly to the power of the LR test. Recommendations from previous simulation studies have suggested that 500 participants would suffice for highly demanding models (those with minimal between-profile membership differences and small numbers of indicators, Nylund, Asparouhov, & Muthén, 2007). Our sample size of 1,215 participants was more than adequate for estimating our 6-profile model. As a secondary precaution and because the chi-square test is sensitive to sparse data, it is recommended to bootstrap *p*-value estimates in order to test the improvement in fit between two models using population-based estimates. This approach was followed in the present study using 1,000 replicated data sets, thus simulating population parameters based on our large sample. All models were run using Latent Gold 5.0 (Vermunt & Magidson, 2013).

Parameter estimates were presented in effect size metric, specifically Cohen's *d* statistic (Cohen, 1992). Cohen's *d* is the metric of standard deviations and therefore differences in the latent class membership (figures 1 and 2) are expressed as standard deviations from the mean of zero (i.e. *z*-scores). As is customary, effect sizes of .5 are considered medium size and significant (as would be derived from inferential analyses), effect greater than .8 standard deviations as large, and effects between 0 SD and .49 as small to medium and non-significant.

## 2.4 Results

### 2.4.1 *Subtypes of early reading profiles*

The baseline model estimated a 1-profile solution which formed the basis for subsequent comparisons. When comparing a 2-profile model to the baseline model, the fit of the 2-profile model was superior, but still not acceptable because the classification errors were at the level of significance and the residual values<sup>4</sup> exceeded the 1.0 recommended value (Magidson & Vermunt, 2002; Magidson & Vermunt, 2001). Subsequently, a 3-profile solution was estimated with the purpose of minimizing those residual co-variations and improving model fit (Table 2). Results suggested that the classification errors were still unacceptably high ( $p < 0.05$ ). Thus, the 3-profile model was discarded in favor of a 4-profile model. When this model was tested using the log-likelihood  $-2LL$  statistic based on the bootstrap distribution and 1,000 replications [ $-2LL_{Diff} = 357.572$ ,  $p < 0.001$ ], it was superior to the 3-profile model. The process was repeated until the superiority of any subsequent model would not be evident using the BIC and/or Corrected AIC index (Tofighi & Enders, 2008). A 5-profile solution was tested and was statistically superior to the 4-profile solution [ $-2LL_{Diff} = 278.551$ ,  $p < 0.001$ ], with a significant improvement in fit by also employing the BIC and corrected AIC (Table 2). The superiority of a 6-profile model was tested against the 5-profile model, which was also supported [ $-2LL_{Diff} = 158.084$ ,  $p < 0.001$ ]. However, when moving to a 7-profile model, its superiority was not substantiated. First, the BIC and corrected

---

<sup>4</sup> Reflecting chi-square statistics regarding the conditional independence assumption. They are bivariate correlations of error between pairs of independent variables. Their expected value is 1.0 when no significant correlation is present.

AIC values were not improved, suggesting a return to the 6-profile model for reasons of parsimony (i.e.,  $BIC_{7\text{-class}} = 37677.9844$ ,  $BIC_{6\text{-class}} = 37656.0927$ ;  $CAIC_{7\text{-class}} = 37777.9844$ ,  $CAIC_{6\text{-class}} = 37741.0927$ ). Thus, the preferred model involved a 6-class solution, which is discussed in detail below.

	Mean	SD	Range
<b>Year 1</b>			
Age (months)	65.7	4.2	57-78
KBIT-2 Matrices (IQ)	98.9	10.5	80-154
YARC Letter-Sound ID (LSK)	99.6	14.8	67-138
CTOPP Composite (PA)	9.9	2.0	5-18
CTOPP Nonword Repetition (VSTM)	8.6	2.6	3-18
RAN Composite (Objects and Colors) (RAN)	97.7	14.7	54-144
WRMT Word ID Raw Score	5.1	11.4	0-71
<b>Year 2</b>			
Age (months)	86.7	4.2	79-100
WISC-IV Processing Speed (PS)	10.4	2.7	4-18
TOWRE-2 Sight Word Efficiency (SWE)	106.2	14.3	71-138
TOWRE-2 Phonemic Decoding Efficiency (PDE)	101.4	14	62-134
CTOPP Composite (PA)	11.7	2.7	7-18
CTOPP Nonword Repetition (VSTM)	11.1	1.9	8-17
WRMT-3 Word ID (WID)	9.1	2.2	4-16
WRMT-3 Word Attack	108.3	14.6	75-145
RAN Composite	105.6	13.6	75-135
TWS-5 (SPELL)	106.8	14.7	71-138
GORT-5 Fluency	10.1	2.8	6-17
GORT-5 Comprehension	10.1	2.2	6-17

**Table 2.1:** Screening and follow-up sample characteristics. Values represent standard scores.

#### **2.4.2 Profile descriptions based on kindergarten data**

The six latent profiles/classes<sup>5</sup> were each defined with ample participants (see *n* for each profile in Table 4). The point estimates of each of the 6 latent classes across IQ, PA, NWR, RAN and LSK predictors (Table 3) demonstrated that (a) each measure was associated with differential effects (levels) across classes (as shown by the significant Wald tests) and (b) the amount of variance of each

<sup>5</sup> The terms profiles and classes have been used interchangeably.

predictor explained by the latent class membership was both significant and large as shown by the R-square values, ranging from 13.3% for IQ to 71.1% for LSK.

The profiles were further characterized in terms of the reading performance (i.e. non-readers versus readers) of their members (Table 4). The order and numbering of the profiles was determined by the number of group members, from largest to smallest.

**Profile 1:** The “Average Performers” group was the largest group and included 378 children (31.1%). This profile was associated with performance near the mean score across all measures. Most of the members in this group were non-readers (78.3%), and the group average performance ranged between 0.0 and 0.5 standard deviations<sup>6</sup> from the age-normed test mean across all measures.

**Profile 2:** The “Low-Average Performers” group was the second largest and included 249 children (20.5%). The scores in this group fell slightly below the test mean on all measures except PA and LSK for which they were .5 SDs and .8 SDs below the mean, thus representing medium and large effect sizes, respectively. Most children in this group were non-readers (85.5%).

**Profile 3:** The “High Performers” group included 235 children (19.3%). This group had a similar, but higher pattern of performance as compared to Profile 1 and was associated with .5 to 1 SD above average performance across all measures. 30.4% of the members in this group were readers. This high performing

---

<sup>6</sup> A standard deviation of 0.5 was selected to represent a medium effect size based on Cohen (1992).

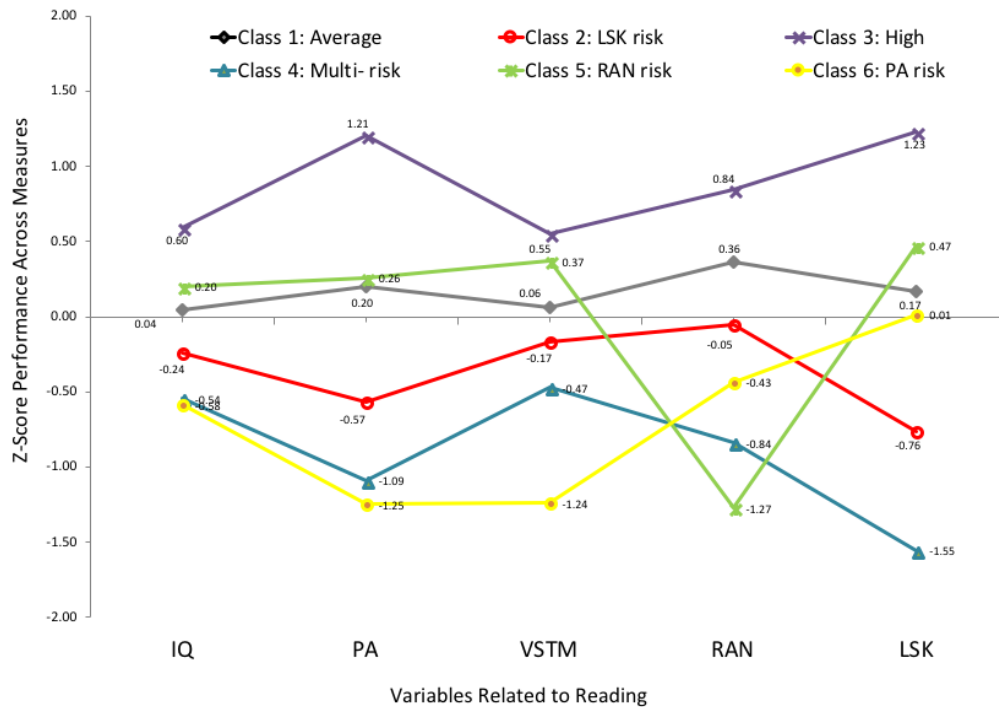
group had achievement levels exceeding a medium effect size (i.e., .5) across all measures. 30.4% of the members in this group were readers.

**Profile 4:** The “Double Deficit (DD) Risk” group included 147 children (12.1%) and was one of the poorest performing classes. This class was associated with -.5 to -1.6 SD below mean performance across all measures. This was the only group in which all members in the group were non-readers (100%).

**Profile 5:** The “RAN Risk” group included 132 children (10.86%). This class was associated with at mean or slightly higher than the mean performance on all measures except the RAN (with effect sizes ranging between .2 and .5). On the RAN, the group performed 1.27 SD below the mean. Over half, 58.3%, of the members of this group were non-readers.

**Profile 6:** The “PA Risk” group was the smallest group including 74 children (6.09%). This group performed .5 SD below the mean on IQ and more than 1 SD below the mean on PA and NWR. Also their RAN performance was close to a medium effect size (i.e., -.43) below the mean and their LSK performance was average. Most members of this class were non-readers (89.2%).

Differences between classes were evaluated by means of the omnibus Wald test and in case of significance, a series of post-hoc tests. However, because those post-hoc tests were run under excessive power levels, due to the large sample size at Year 1, almost all estimates were significant. In order to avoid inflated statistics the comparison between classes was based on effect size metrics, specifically, Cohen's d statistic as discussed above.



**Figure 2.1:** Latent profile analysis model for the identification of reading subgroups: Optimal solution. Raw scores were transformed to z-scores on all variables. PA: phonological awareness, VSTM: verbal short-term memory, RAN: rapid automatized naming, LSK: letter sound knowledge.

#### 2.4.3 Longitudinal prediction based on latent class membership

A latent class regression model (Magidson & Vermunt, 2001) was employed to test the hypothesis that profile membership in kindergarten would be predictive of end-of-1<sup>st</sup> grade reading performance for the subset of children who participated in the follow-up portion of the study ( $n = 95$ ). A Monte Carlo simulation was conducted to estimate the power levels of the mixture model using an  $n$ -size of 95 participants. The simulation involved a latent profile variable with 11 continuous indicators and 6 latent classes for which a standardized mean estimate of .80 (suggesting a large effect) was tested for significance (through fixing the latent class variances to 1). Results indicated that coverage (i.e. the confidence intervals that contained the true population mean of .8) ranged between

80% and 85.8% and power (proportion of correct rejections) at a 5% level of significance ranged between 70% and 74%. Both estimates of coverage and power were acceptable using our proposed sample size of 95 participants.

A bias analysis was additionally conducted to ensure that the Year 2 cases were allocated among the classes the same way as in the original (kindergarten) sample. This analysis ruled out the possibility that the findings at Year 1 are due to the different composition of the sample at Year 2 compared to that of the kindergarten sample. Specifically, a cross-tabulation analysis using Pearson's chi-square statistic was conducted to evaluate the similarity in percentages. The omnibus Pearson chi-square test was non-significant [ $\chi^2(5) = 7.36, p = .92$ ] indicating a similar representation of cases in the 6 profiles for the longitudinal participants, as compared to the full sample. Specifically, the percentages per class were as follows (Year 1 and Year 2): for Class 1, 31%/27%; for Class 2, 20%/23%, for Class 3, 20%/14%, for Class 4, 12%/13%, for Class 5, 10%/18%, and for Class 6, 6%/5%. Consequently, the Year 1 6-class category classification was used as an independent variable and the following Year 2 measures were entered as dependent variables: a) WISC-IV coding (WISC PS), b) TOWRE-2 sight word efficiency, c) TOWRE-2 phonemic decoding efficiency, d) CTOPP elision and blending (PA), e) CTOPP nonword repetition (WM), f) WRMT-III Word ID, g) WRMT-III word attack, h) RAN objects and colors composite score, i) GORT-5 reading fluency, and, j) GORT-5 reading comprehension, k) TWS-4 spelling. The tested means and significance levels shown in Table 3 suggested that the classes were adequately differentiated based on the classification variables. Table 5 shows the means on



each of the Year 1 measure for each class and the Wald statistics, which indicate differences between classes on, mean point estimates. Latent class formation was distinct across all measures, such that no measures were associated with identical point estimates across the latent classes.

Model	LL	BIC(LL)	AIC(LL)	AIC3(LL)	CAIC(LL)	Npar	Class. Err.
1-Class	-20054.9921	40181.0093	40129.9843	40139.9843	40191.0093	10	0.0000
2-Class	-19204.0640	38585.6904	38458.1279	38483.1279	38610.6904	25	0.0574
3-Class	-18923.2936	38130.6872	37926.5872	37966.5872	38170.6872	40	0.0465
4-Class	-18744.5077	37879.6529	37599.0155	37654.0155	37934.6529	55	0.0662
5-Class	-18605.2323	37707.6396	37350.4646	37420.4646	37777.6396	70	0.0876
6-Class	-18526.1901	37656.0927	37222.3803	37307.3803	37741.0927	85	0.0854
7-Class	-18483.8672	37677.9844	37167.7345	37267.7345	37777.9844	100	0.1011

**Table 2.2:** Model comparison for optimum latent class solution.

Note: \* $p < 0.05$ ; \*\* $p < 0.01$ . Optimum solution is in italics and reflects a 6-class latent variable model. LL = log likelihood; BIC = Bayesian Information Criterion; AIC = Akaike Information Criterion; AIC3 = Corrected AIC with a penalty factor of three; CAIC = Consistent AIC; Npar = Number of estimated parameters; Class. Err. = Classification error. Preferred models should have non-significant amounts of classification errors.

†The Bootstrapped Likelihood Ratio Test (BLRT) was employed in order to compare adjacent models using 500 replications. Thus, the 2-class model provided a significant improvement over the 1-class model [ 2LL Diff = 992.244,  $p < .001$ ]; similarly the 3-class model was superior to the 2-class model [ 2LL Diff = 195.254,  $p < .001$ ] and the 4-class model significantly improved on the 3-class model [ 2LL Diff = 539.211,  $p < .001$ ]. The 5-class model improved on the 4-class model [ 2LL Diff = 343.624,  $p < .001$ ] and the 6-class model on the 5-class model [ 2LL Diff = 48.145,  $p < .01$ ]. The 7- class model was statistically a superior model to the 6-class model but the parsimoniousness indices (BIC and CAIC) suggested that it was over- parameterized in relation to the amount of information it provided.

Predictors	Average	Low-average	High	DD risk	RAN risk	PA risk	Wald	R <sup>2</sup>
IQ	98.941	96.179	104.112	92.519	100.049	92.796	106.18***	0.133
PA	10.295	8.722	12.316	7.58	10.338	7.249	1457.53***	0.593
NWR	8.758	8.255	10.026	7.382	9.445	5.474	367.83***	0.191
RAN	103.133	97.223	110.054	85.354	79.253	90.798	1037.15***	0.451
LSK	103.523	90.159	118.537	77.714	107.426	100.213	2365.5***	0.711

**Table 2.3:** Point estimates of each of the six latent classes across IQ, PA, NWR, RAN and LSK predictors.

Note: Values in the table are means in the original score metric, for clarity.\*\*\* $p < 0.001$ .

	Average		Low-average		High		DD risk		RAN risk		PA risk		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Non-readers	287	75.9	242	97.2	24	10.2	147	100.0	77	58.3	66	89.2	843	69.4
Readers	91	24.1	7	2.8	211	89.8	0	0.0	55	41.7	8	10.8	372	30.6
Total*	378	34.0	249	28.7	235	2.8	147	17.4	132	9.1	74	7.8	1215	100.0

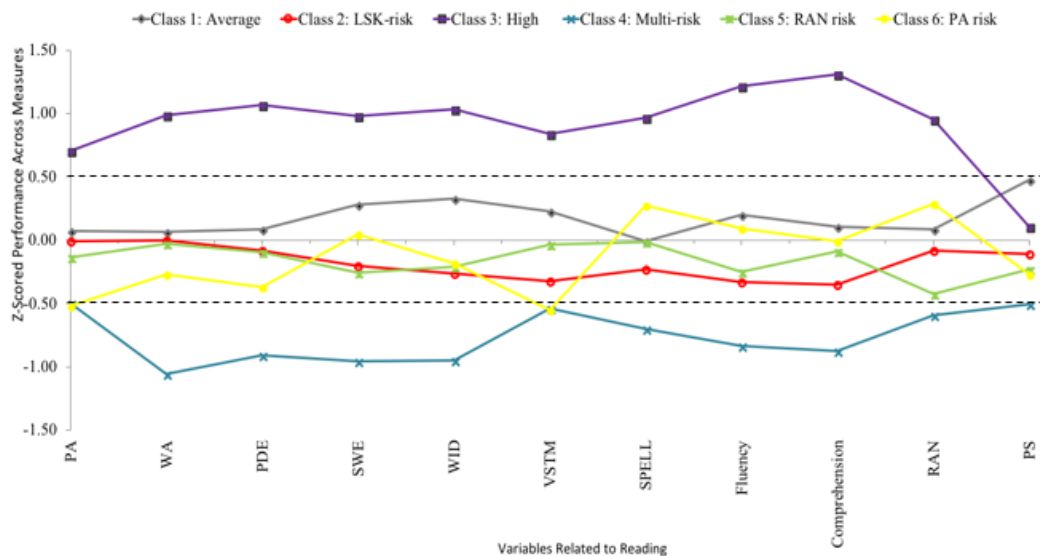
**Table 2.4:** Reading performance of the members of each profile.

Note: Column percentage totals represent the percentage of non-readers in each class as compared to the total number of non-readers

Predictors	Average	Low-average	High	DD risk	RAN risk	PA risk	Wald	R <sup>2</sup>
PS	11.779	9.935	10.316	8.916	9.706	9.518	14.264***	0.148
SWE	109.674	103.157	122.754	92.135	101.76	102.895	86.013***	0.327
PDE	101.909	99.729	117.185	87.858	98.763	91.9	56.200***	0.304
PA	11.724	11.724	13.836	10.447	11.099	9.599	18.511***	0.157
NWR	9.425	8.33	10.86	7.895	8.853	7.519	28.588***	0.216
Word ID	112.342	104.168	123.68	94.616	104.081	101.159	80.041***	0.348
Word Attack	105.259	105.368	118.634	90.279	103.318	97.798	48.890***	0.292
RAN	108.27	106.309	121.512	97.122	99.808	109.55	19.849***	0.224
SPELL	104.762	101.785	124.111	92.906	104.451	106.335	34.349***	0.300
Fluency	10.837	9.504	13.885	8.132	9.663	10.125	50.975***	0.383
Comprehension	10.142	9.328	13.078	8.2	9.751	9.549	50.386***	0.413

**Table 2.5:** Mean point estimates and significance tests for longitudinal prediction of reading outcomes from latent class formation at kindergarten. Note: \* $p < .05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

A latent class analysis on all Year 2 measures revealed 6 distinct profiles of performance. The patterns of performance across profiles on variables that overlapped between the two years, as well as on additional variables, closely resembled the pattern of performance on Year 1 measures. A predictive model was developed to test the likelihood that a particular child who belonged to one class in Year 1 will remain in the same class in Year 2. Results (Figure 2.2 and Table 2.5) indicate that all of the children were classified to the same latent class in Year 2 as in Year 1, reflecting 100% stability in class membership.



**Figure 2.2:** Latent class regression predicting reading outcomes from the earlier-formed kindergarten latent profiles. The values in the figure are z-scores. The Wald z-statistic suggested that all variables were significant in differentiating between latent classes. That is, for all variables there were differential levels of performance per class. PA: phonological awareness, VSTM: verbal short-term memory, WA: word attack, PDE: phonemic decoding efficiency, SWE: sight word efficiency, WID: word identification, SPELL: spelling, RAN: rapid automatized naming, PS: processing speed.

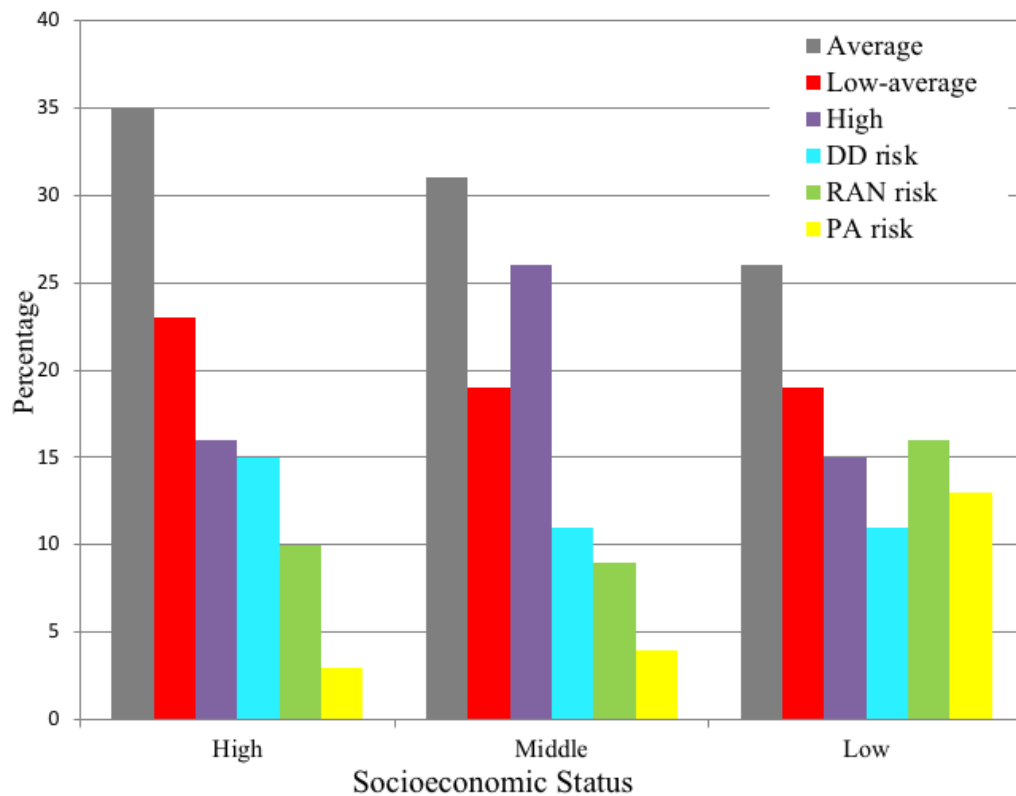
In terms of general performance on Year 2 measures, children in the High-Performers profile (Class 3) had the highest means across all measures except on the WISC PS at Year 2, with effect sizes ranging from medium to large. The Double Deficit profile (Class 4) maintained low performance in 1<sup>st</sup> grade across measures with medium to large effect sizes. Specifically, this profile performed worse, than all other profiles, on all measures except PA, whereas the PA deficit profile scored the lowest but very close to the DD group's estimates (with effect sizes of -.52 and -.50, respectively). The PA-Deficit Group (Class 6) maintained low performance on all phonological measures (PA, phonemic decoding efficiency, and word attack)

as well as WM with small and medium effect sizes, and maintained above average performance on RAN and other speeded measures (reading fluency and sight word efficiency). Furthermore, the Average performing profile (Class 1) demonstrated a slight advantage (small effect size) in performing on the sight word efficiency task as compared to the phonemic decoding efficiency task. This advantage was significantly more pronounced (medium effect size) in the PA deficit group and there was no advantage for the RAN deficit (Class 5) and the Low-Average (Class 2) groups. The low-average group demonstrated a below average (small to medium effect size) performance on NWR, Word ID, spelling, and comprehension and fluency measures, but not on any of the phonological decoding and awareness measures, for which performance was at average levels. The RAN deficit group (Class 5) remained average performing on all measures except RAN, for which performance was below average (small to medium effect size). The PA deficit group outperformed the RAN group on 1<sup>st</sup> grade speeded reading measures (i.e. sight word efficiency, fluency) with small effect size. Children in the RAN deficit and DD groups were the only ones who demonstrated higher reading comprehension than reading fluency skills.

#### **2.4.4 *Cognitive and environmental factors***

To evaluate the relationship between SES and class membership, the distribution of profiles across the three school-level SES groups was tested. Chi-Square tests revealed a significant ( $p < 0.001$ ) difference in profile distribution across the three SES groups. Whereas the majority of high-performing and average-performing students (Profiles 1 and 3) belonged to the high (41.8% and 31.1%

respectively) and medium (36.2% and 49.4% respectively) SES groups (versus 22% and 19.6% in low the SES group), the majority of the PA and RAN deficit students (Profiles 5 and 6) belonged to the low-SES group (37.1% and 56.8% versus 32.6% and 17.6 in high SES and 30.3% and 25.7% in middle). These results are especially striking considering that the low SES class had fewer students overall ( $n = 314$  versus  $n = 457$  for high SES and  $n = 444$  for middle SES). The double-deficit and low-average class distribution was proportional to the SES group size (Figure 3). Pearson correlation revealed that performance on the IQ measure was significantly positively correlated with all Year 1 measures (Pearson  $r$  estimates with PA = .344,  $p < 0.001$ , with LSK = .140,  $p < 0.001$ , with RAN = .226,  $p < 0.001$ , and with NWR = .276,  $p < 0.001$ ).



**Figure 2.3:** Latent class distribution across SES groups.

## **2.5 Discussion**

This study was the first to apply latent profile analysis (LPA) and longitudinal regression approaches to characterize the heterogeneous profiles of early reading performance of a large sample of kindergarten and pre-kindergarten students and to evaluate the predictive capacities of these profiles longitudinally in the context of socioeconomic and cognitive factors (i.e., IQ). Six distinct profiles of reading emerged and were characterized as follows: average performers, high performers, low-average performers, RAN risk, PA risk, and double- deficit risk. Importantly, these patterns of performance were in accordance with previous risk classification studies and significantly predicted performance on end- of-1st-grade reading and language measures, revealing a longitudinal stability of class membership of 100%.

### **2.5.1 Implications for dyslexia risk subtypes**

Similar to previous studies that did not use a predetermined cut-off to characterize risk (Boscardin, Muthén, Francis, & Baker, 2008; Torppa et al., 2007; Wolff, 2010), multiple reading profiles emerged in our sample. Three distinct profiles of deficits that differed in performance level and pattern were identified: PA deficit, RAN deficit, and double-deficit (DD). In terms of general performance on all measures, the DD group performed more poorly than the PA risk group, which in turn had lower scores than the RAN risk group. These results are in line with previous double-deficit studies that found similar relative performance among

the PA, RAN, and DD groups (Katzir, Kim, Wolf, Morris, & Lovett, 2008; Lovett, Steinbach, & Jan; Vaessen, Gerretsen, & Blomert, 2009; Wolf & Bowers, 1999).

Whereas previous findings of lower PA scores in the DD group as compared to the PA deficit group led some authors to question the validity of the double-deficit distinction (Compton et al., 2001; Schatschneider et al., 2002), the DD group in our sample had comparable PA scores to the PA risk group. In fact, on the verbal short-term memory measure, the DD group scored significantly higher than the PA group, albeit with a small effect size. This suggests that the reduced performance of the DD group in both years is due to the cascading effects of both phonological and RAN deficits that impair reading acquisition across several levels of processing, e.g., phonological, visual, attentional, and retrieval (Wolf & Bowers, 1999, 2000).

In contrast, the RAN risk group had intact performance on all other kindergarten measures (including PA), further supporting the independence of the RAN construct from PA. RAN is thought to index the automaticity with which cognitive processes important for reading are executed and integrated (Norton & Wolf, 2012). Consequently, RAN has been strongly linked to timed word identification measures and reading fluency. Indeed, the RAN risk group performed below the other profiles (except DD risk) on 1st-grade rate-related skills (i.e., Sight Word Efficiency, Fluency). In addition, the RAN risk group's pattern of low fluency performance as compared to comprehension is in contrast to that of the other groups that demonstrated similar performance on comprehension and fluency.

Thus, the current results bolster the specificity of RAN's association to speeded and fluency-related measures.

The low performance of the PA risk group on phonological measures, but not reading or spelling measures, both supports the stability of the PA construct and suggests that the PA deficit on its own is insufficient to cause reading impairment. PA indexes the ability to decode (i.e., sound out) words that are not yet automatic as well as non-words (Stahl & Murray, 1994). Accordingly, the PA risk group had lower phonemic decoding skills (Phonemic Decoding Efficiency, Word Attack) as compared to Sight Word Efficiency, showing a different pattern from the DD and RAN risk groups. The PA risk group's impairment on phonological measures was specific, as they did not show reduced LSK in kindergarten or impaired 1<sup>st</sup> grade reading comprehension and fluency performance. This is in line with studies demonstrating a dissociation between phonological deficits and reading performance in the absence of other exacerbating risk factors (Moll et al., 2013). The small size of the PA risk profile further suggests the rarity of pure phonological deficits early in reading development.

The low-average profile comprised the largest group in the sample and was characterized by below average performance on all kindergarten measures. This group also was characterized by low performance on LSK in kindergarten (as compared to the other measures) and by the non-reading status of the majority of the group. In Year 2, this group demonstrated below-average performance on single word identification measures (Word ID, Sight Word Efficiency), as well as on spelling, reading comprehension, and reading fluency. In both years, the low-



average group demonstrated low performance on the verbal short-term memory measure. This unique pattern of poor performance on orthographic measures in kindergarten and typical performance on phonological as compared to orthographic and lexical reading measures in 1st grade, is reminiscent of another conceptualization of dyslexia reported in literature: the surface deficit of dyslexia (Castles & Coltheart, 1993). Surface dyslexia has been characterized by intact phonological abilities and intact regular word reading, but poor exception word reading (Coltheart, Masterson, Byng, Prior, & Riddoch, 1983). Exception words are words that have irregular spelling and, therefore, cannot be read by applying phonological grapheme-to-phoneme conversion rules. Instead, these words are read holistically through direct access to the lexical information underlying a specific orthographic pattern. Children with a surface-dyslexia-like deficit have a problem in developing direct visual representations of words and are thus differentially impaired at tasks emphasizing orthographic knowledge (Jiménez, Rodríguez, & Ramírez, 2009; Manis, Seidenberg, & Doi, 1999; Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996; Stanovich, Siegel, & Gottardo, 1997). Accordingly, the Sight Word Efficiency measure, on which the low-average group showed lower performance as compared to Phonemic Decoding Efficiency, included many irregular words.

There has been mixed evidence for the validity of the surface dyslexia subtype. Some studies with reading-level controls suggested that it represents a developmental delay rather than a distinct deficit (Manis et al., 1996; but see Peterson, Pennington & Olson, 2013; Stanovich et al., 1997). These developmental

delays have been attributed to poor home literacy or language environments (Castles, Datta, Gayan & Olson, 1999; Sprenger-Charolles, Siegel, Jimenez & Ziegler, 2011). This profile was not over represented in the low SES schools, suggesting that in the current study low-risk deficits in LSK are not environmentally driven. In addition, studies suggest that phonological and surface dyslexia differ only in the degree of severity of phonological deficits and in cognitive resources available to compensate for these deficits (Snowling, 1998). Future studies will determine whether the initial orthographic deficits demonstrated in the current study for the low-average group will be ameliorated with additional reading instruction or become more pronounced in later grades.

### **2.5.2 *Longitudinal stability of risk classifications***

The longitudinal stability of early pre-reading literacy profiles has important implications for dyslexia risk identification and intervention. Our results demonstrated perfect stability in classification from the beginning of kindergarten to the end of 1st grade. Importantly, the patterns of performance on pre-reading measures across the groups correlated with performance on more advanced reading measures in a manner that is consistent with the theoretical expectations of the double-deficit and the surface-phonology deficit approaches. Since children who are poor readers in 1st grade tend to remain poor readers by the end of elementary school (Boscardin et al., 2008; Francis, 1996; Juel, 1988; Torgesen & Burgess, 1998) and on through 12<sup>th</sup> grade (Ferrer et al., 2015), these findings point to the validity of our kindergarten battery of tests in identifying dyslexia risk and its sensitivity to individual differences in performance.

Indeed, letter knowledge, phonological awareness, verbal short-term memory, and rapid automatized naming have been identified across several studies as the most robust early predictors of reading abilities (Kirby, Desrochers, Roth, & Lai, 2008; Scarborough, 1991; Schatschneider, Carlson, Francis, Foorman, & Fletcher, 2002; Warmington & Hulme, 2012). Our study demonstrated the stochastic independence among these measures and their robustness in distinguishing between and among various profiles of reading development. Importantly, the differences in performance between typical groups and risk groups on pre-literacy measures extended to differences in actual reading performance on word recognition, fluency, and comprehension measures. These findings suggest that early identification of dyslexia risk is possible and that one-size-fits-all interventions will likely be less effective in accommodating the specific deficits and strengths of the various risk profiles (Allor, Mathes, Jones, Champlin, & Cheatham, 2010; Vaughn et al., 2012).

It is important to note that the theoretical interpretation of our findings could be affected by the selection of measures. The inclusion of other measures in the kindergarten battery could have resulted in different profiles of performance in accordance with other theories of dyslexia (e.g., visual attention). Yet, the selection of measures for the current study was motivated by the robust empirical support for their strong predictive value for reading outcomes across languages, supporting the significance of the current findings. In addition, since the LSK measure was administered in Year 1 only, it was not possible to evaluate the longitudinal stability in performance on this measure. Due to the well-documented limited power of LSK

to differentiate between reading abilities beyond kindergarten (due to ceiling effects), the measure was excluded from the Year 2 battery (McBride-Chang, 1999; Wagner & Barker, 1994). Single word measures, however, were administered in Year 2 and are considered a good proxy of early letter knowledge as there is a high concurrent and predictive correlation between the two constructs (Scarborough, 1998; Schatschneider et al., 2004). Indeed, in the current study, children demonstrated similar performance on LSK in Year 1 and Word ID in Year 2. Furthermore, the small size of the longitudinal sample ( $n = 95$ ) raises the possibility of Type-II error. A Monte Carlo simulation was conducted to test this possibility and showed that such an error was unlikely. By employing a bias analysis, we further demonstrated that the longitudinal sample was representative of the kindergarten sample both in terms of the pattern of distribution across profiles and in demographic characteristics.

### **2.5.3 *Cognitive and environmental factors***

Reading development occurs in the context of cognitive and environmental influences. We observed that the frequency of PA and RAN risk was significantly higher in the low-SES schools than in the middle-SES or high-SES schools. This was not the case, however, for DD risk. It is possible that social factors have a higher impact on the single-deficit groups, whereas the double-deficit is influenced more by hereditary factors. Indeed, previous studies reported a higher frequency of a family history of dyslexia in the DD group as compared to other reading profiles (Morris, Steubing, et al., 1998) and studies have demonstrated more severe reading

deficits in children with higher genetic liability for dyslexia (van Bergen, de Jong, Maassen, & van der Leij, 2014).

Since the majority of schools in this study were charter or private schools, however, students in these schools, even with free/reduced lunch qualification, may not be representative of low-income children who attend non-charter public schools (Lubienski & Lubienski, 2006). For example, in many cases, parents must put forth substantial effort to gain admission and scholarships to a private school or to secure a spot for their child in an oversubscribed charter school. These parents may be more invested in their children's early literacy development. Therefore, different home and school environments may underscore differences in literacy development in children in charter versus public schools. Thus, the current SES results should be interpreted with caution, until future investigations can focus on both school-level and family-level socioeconomic and environmental factors, using a higher proportion of low-SES public schools and a family-level measure of SES.

Our results did not provide support for the IQ– discrepancy model of dyslexia. The PA and DD groups had low-average performance on the non-verbal IQ measure, while the RAN group had average performance. Thus, across the deficit subgroups, the pattern of low language skills despite average IQ did not emerge. In fact, the non-verbal IQ scores were significantly correlated with all Year 1 language measures, indicating a strong coupling between general cognitive abilities and reading. This is in line with previous reports of a strong relationship between cognitive and reading abilities in early grades and the gradual weakening of this relationship across development and into adulthood (Ferrer, 2010). Taken

together, the current results join an increasing body of evidence against using an IQ-based discrepancy criterion to classify dyslexia risk (Fletcher et al., 1994; Pennington, Gilger, Olson, & DeFries, 1992; Stanovich, 1994; Tanaka et al., 2011). Future investigations should examine how the interaction between general cognitive abilities and reading achievement is mediated by social and hereditary factors in order to best determine particular profiles of reading and dyslexia.

#### **2.5.4 *Summary***

These findings are novel in applying a data-driven analysis approach to demonstrate the robustness of RAN, PA, VSTM, and LSK tasks administered in early kindergarten in differentiating the discrete subtypes of dyslexia and predicting later reading performance with high accuracy. Current results carry important implications for improved early identification, differentiated remediation, and an evolving understanding of dyslexia. The high stability of group membership supports the feasibility of early identification of risk, prior to reading failure. This is important for optimizing the educational and psychosocial outcomes of children with dyslexia. Performance on the non-verbal IQ measure of the different groups was proportional to the general level of performance across measures, showing no supporting evidence for the IQ–discrepancy model of dyslexia. Finally, the over-representation in low-SES schools of PA and RAN deficit profiles, but not double-deficit or surface-deficit profiles, provides insight both into the environmental factors influencing dyslexia risk, and also possible hereditary factors.

### **3 Study 2: Distinct Neural Alterations of Heterogeneous Dyslexia Risk Profiles**

#### **3.1 Abstract**

Studies in adults and school-age children demonstrated distinct cognitive and neural profiles of dyslexia subtypes. It remains unclear whether these subtypes emerge prior to reading instruction and whether they are associated with subtype-specific structural brain differences. A latent profile analysis was used to identify distinct patterns of performance on rapid automatized naming (RAN), phonological awareness (PA), letter-sound knowledge (LSK), and working memory (WM) in 1,215 kindergarten and pre-kindergarten students from diverse backgrounds in New England. Four distinct profiles of risk emerged: LSK risk, PA risk, RAN risk, and multi-risk. Alterations in whole-brain gray matter volume and fractional anisotropy in three left-hemispheric tracts important for reading: inferior longitudinal fasciculus (ILF), arcuate fasciculus (AF) and superior longitudinal fasciculus (SLF) were investigated in the four risk groups as compared to the average performers in a subset of participants ( $n = 100$ ). Compared to the average group, risk was associated with significantly reduced grey matter volume indices in the LSK group

in the bilateral temporal-parietal regions, in the multi-risk group in the right inferior temporal/hippocampal regions and precuneus, in the RAN group in the right middle-temporal regions, and in the PA group in the right inferior temporal/fusiform region and posterior parietal regions. Reduced FA was demonstrated in AF in the LSK group, SLF in the multi-risk group, and in both tracts in the RAN risk group. These findings suggest that different subtypes of dyslexia risk profiles are associated with distinct neuroanatomical regions of the reading network in early kindergarten and highlight the heterogeneity of dyslexia etiology.

### **3.2 Introduction**

Research suggests that developmental dyslexia lies on a multidimensional continuum of reading-related abilities with different possible subtypes (Morris, Stuebing, et al., 1998; Peterson & Pennington, 2015). Indeed, a range of perceptual and cognitive deficits have been reported for dyslexia, including rapid automatized naming (RAN), phonological awareness (PA), and orthographic processing (Ozernov-Palchik, Yu, et al., 2016; Pennington et al., 2012). There is not yet consensus, however, on the distinct subtypes of dyslexia as the evidence for the cognitive dissociation of the different subtypes has been mixed (Elliott, 2014).

Whereas behavioral studies have provided some understanding of dyslexia, measures of brain structure have provided additional insights. The advent of MRI technology allowed for in vivo investigations of the neuroanatomical correlates of dyslexia to advance the understanding of the neurobiology of the disorder. These investigations revealed that dyslexia is associated with reduced grey matter volume



in brain structures important for reading including bilateral temporo-parietal regions and left occipito-temporal regions (Linkersdörfer, Lonnemann, Lindberg, Hasselhorn, & Fiebach, 2012; Richlan et al., 2013). Further, dyslexia is associated with altered volume and diffusion properties in three left-hemisphere tracts: the arcuate fasciculus (AF), connecting the superior temporal lobe with the inferior frontal gyrus (IFG), the left superior longitudinal fasciculus (SLF), connecting the inferior parietal with the inferior frontal/premotor regions, and the left inferior longitudinal fasciculus (ILF), connecting the posterior inferior temporal gyrus with the ventral anterior and medial temporal lobe (Hoeft, 2011a; Myers, 2014; Saygin et al., 2013; Vandermosten, Boets, Wouters, & Ghesquiere, 2012; Yeatman, 2012a).

Despite the heterogeneous nature of dyslexia, most of the MRI studies of dyslexia to date have adopted a dichotomous group approach, comparing individuals with and without dyslexia. Studies that examined each deficit separately and the few studies that examined multiple skills and/or deficits in the same individuals produced some positive evidence for the neural dissociation of the different skills associated with dyslexia subtypes. The neural underpinnings of PA skills and deficits have been associated with the left temporo-parietal and inferior frontal regions (Bookheimer, Zeffiro, Blaxton, Gaillard, & Theodore, 1995; Conant, Liebenthal, Desai, & Binder, 2014; Frost et al., 2005; Pugh et al., 2013; Shaywitz et al., 2002). Orthographic processing has been associated with occipito-temporal regions, most notably, the putative visual word form area (VWFA-Cohen et al., 2002). The neural underpinnings of RAN are less well characterized due to

the paucity of studies, but findings to date implicate distributed brain regions including bilateral inferior/middle frontal, inferior parietal, right middle temporal, and ventral visual regions, as well as the right anterior region of the cerebellum (Cummine, Chouinard, Szepesvari, & Georgiou, 2015; Eckert et al., 2005; He, Xue, Chen, Chen, et al., 2013; Misra, Katzir, Wolf, & Poldrack, 2004; Norton et al., 2014; Turkeltaub, 2003).

The main limitation of neuroimaging studies of children who are already reading is that they represent both the innate neurobiological differences in the brains of readers with dyslexia, and also the consequences of environmentally-driven neuroplasticity; that is, the conflated effects of reading acquisition and reading deficits (Goswami, 2014). Furthermore, because several different etiological mechanisms can lead to the same outcome of reading failure, in order to understand the underlying causes of dyslexia, it is important to study it early in development when environmental and experience differences are more minimal. Results from recent MRI studies of pre-reading children at risk for dyslexia have demonstrated brain alterations associated with dyslexia risk in preschoolers and infants (e.g., Black, Myers, & Hoeft, 2015; Im, Raschle, Smith, Grant, & Gaab, 2015; Langer et al., 2017; Maurer et al., 2009; Myers, 2014; Ozernov-Palchik & Gaab, 2016; Raschle, Zuk, Ortiz-Mantilla, et al., 2012; Raschle et al., 2011; Vandermosten et al., 2015; Wang et al., 2016). The locations of these alterations generally reflect those reported in studies of already-reading individuals (See Gaab & Ozernov-Palchik, 2016; Vandermosten, Hoeft & Norton, 2016, for reviews).

No MRI studies to date have investigated the segregation of the RAN, PA, and orthographic skills in pre-reading children. The earliest dissociation of RAN and PA in developing readers has been demonstrated in an fMRI study of children aged 6 and older (Turkeltaub et al., 2003). In that study, RAN and PA correlated with distinct patterns of brain activity during an implicit in-scanner word-reading task. The PA measure correlated with activation in the left hemispheric language regions, including the superior temporal sulcus and ventral inferior frontal gyrus. In contrast, RAN skills correlated with activation in the right posterior temporal gyrus, right middle temporal gyrus, bilateral middle frontal gyrus, and left ventral inferior frontal gyrus during the reading task. Studies in younger children have demonstrated bilateral involvement of the temporo-parietal regions in letter-ID, the earliest measure of orthographic skills in pre-readers (Specht et al., 2009; Yamada, 2011). For example, in one study, kindergarten children who were classified as at-risk based on letter-sound knowledge (LSK) measures demonstrated reduced bilateral activation within temporo-parietal regions to letters, as compared to no-risk children (Yamada, 2011). No studies to date, however, investigated the neuroanatomical segregation of the different deficits in pre-reading children.

Neuroimaging studies of reading, therefore, provide some evidence for the multi-deficit view of dyslexia and research in pre-readers has shown that dyslexia risk is associated with brain alterations early in reading development. Demonstrating that heterogeneous profiles of reading-related skills are associated with unique brain alterations in pre-reading children would point to the distinct etiology of these profiles and provide further support for the multi-deficit approach

to dyslexia. Identifying the structural correlates of dyslexia profiles is particularly important, since these are independent of task performance and can be assessed in pediatric populations as young as infancy, allowing for future replication of findings and clinical application (Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015).

In a recent study, we conducted a latent profile analysis with a diverse sample of 1,215 kindergarten and pre-kindergarten students from 20 schools (see Study 1). Six early literacy ability profiles of performance emerged from the data: Average performers, LSK risk<sup>7</sup>, High performers, PA risk, RAN risk, and Multi-risk<sup>8</sup>.

In order to better characterize the etiological basis of different dyslexia subtypes, in the current study, we sought to extend previous research by investigating whether distinct dyslexia risk subtypes are associated with independent neuroanatomical differences, before these differences are confounded with reading experience. To this end, we investigated grey matter volume (GMV) using Voxel Based Morphometry (VBM) and white matter integrity, as measured by fractional anisotropy (FA) measures from diffusion-weighted imaging (DWI) in a subset of the 1,215 children ( $n = 100$ ). Each of the risk groups' whole-brain GMV and FA values in three left hemisphere tracts important for reading (i.e., left ILF,

---

<sup>7</sup> The group was initially named 'Low-average' based on its kindergarten performance that was similar in pattern to but with lower scores than that of the average group, but it was renamed to LSK risk because it demonstrated distinct patterns of orthographic deficit on the 1<sup>st</sup> grade measures (for a more detailed description, see Ozernov-Palchik et al., 2017) and of semantic/lexical deficit on 2<sup>nd</sup> grade measures.

<sup>8</sup> The group was previously named "double-deficit," but was renamed here to multi-risk to be theoretically neutral

AF, and SLF) were compared with those of the Average group. In order to establish the long-term significance of early risk identification, we examined longitudinal behavioral outcomes of these groups three years later. If subtypes of dyslexia are associated with independent etiological mechanisms, distinct patterns of grey matter reductions across the four risk groups would be evident. Based on previous MRI studies of dyslexia risk, we expected these reductions in the temporo-parietal regions and occipito-temporal regions bilaterally. If dyslexia risk is associated with atypical development of connections among areas of the reading network, reduced mean FA in the three tracts that have been most consistently associated with reading across development (SLF, AF, and ILF) would be expected. Finally, based on our previous behavioral findings, we expected robust stability between kindergarten risk designation and second grade reading and language outcomes.

### **3.3 Methods**

#### **3.3.1 *Participants***

The study was conducted in four stages: 1) Kindergarten (KG): children from 20 diverse schools in New England completed a short battery of pre-reading assessments administered by trained researchers in their schools (for details see Study 1); In-school testing took place in the spring of pre-kindergarten or fall of kindergarten. 2) At the follow-up session, parents were contacted and screened for eligibility, and eligible children participated in a follow-up neuroimaging study and additional behavioral assessments. While the children were being assessed, their parents completed background questionnaires. Children who participated in the follow-up session in kindergarten came in for two behavioral follow-up sessions at

the end of 1st and 2nd grades. Based on parental report, all children had normal hearing, no neurological or psychiatric disorders other than ADHD ( $n = 9$ ), American English as a native language, full term birth ( $>36$  weeks). All children also had KBIT Matrices (nonverbal IQ) standard score  $>80$ . A subset of children ( $n = 100$ ) was included in the current analysis (54.5% male, mean age at first session 66.9 months,  $SD = 4.0$  months). Data from 64 children was available for inclusion in the longitudinal behavioral analysis (56.3% male, mean age 98.9,  $SD = 4.4$ ). Most children were right-handed ( $n = 88$ ), but there were a few left-handed ( $n = 8$ ) and ambidexterous ( $n = 4$ ) children in the sample. Children were recruited to participate in the neuroimaging portion of the study to maintain a representative composition to that of the larger sample tested at schools in regards to gender, age, ethnicity/race, school type, and behavioral scores. This study was approved by the institutional review boards at the Massachusetts Institute of Technology and Boston Children's Hospital. Parents gave written consent and children gave verbal assent to participate.

### **3.3.2 *KG behavioral measures***

All participants were administered a comprehensive psychometric testing battery measuring cognitive and language skills. The following measures (described in Chapter 1) were used to classify children into latent profiles: phonological awareness (PA), verbal Short Term Memory (VSTM), rapid automatized naming (RAN), letter sound knowledge (LSK), nonverbal IQ (IQ).

### **3.3.3 Background measures**

**Home Literacy and SES Information.** Home literacy and SES information were collected via an extensive background questionnaire completed by one or both parents. SES questions were adapted from the MacArthur Research Network (<http://www.macses.ucsf.edu/Default.htm>), and home literacy questions were adapted from Denney et al., 2001 (as cited in Katzir et al., 2009). Groups were compared on the following criteria: Parental education (highest degree obtained, categorical score: 1=8<sup>th</sup> grade or less; 2= Some high school (HS); 3=HS/GED degree; 4=Associate degree or some college; 5=Bachelor's degree; 6=Master's degree; 7= Doctorate degree or equivalent); 2) Total household income (categorical value: 1=<\$5,000 to 9=\$501,000+); and 4) Hours spent reading with the child per week.

**Familial History of Dyslexia Measure.** Parental history of reading difficulties was evaluated using the Adult Reading History Questionnaire (Lefly & Pennington, 2000). The questionnaire is scored by summing the responses to all questions and dividing by the total number of questions. Greater scores on the ARHQ indicate more reading impairment during childhood.

### **3.3.4 Imaging procedures**

Children practiced in a mock scanner area in the beginning of each MRI session, and child-friendly equipment (e.g., pediatric headphones, head padding, etc.) and procedures were used during data acquisition as described by Raschle and colleagues (2009; 2011). MRI sequences were acquired on a Siemens 3T Trio whole body MRI scanner at the Athinoula A. Martinos Imaging Center at the

McGovern Institute for Brain Research at MIT using a standard 32-channel head coil. The T1-weighted MPRAGE scan used the following specifications: 176 slices, TR= 2350 ms; TE= 1.64 ms; flip angle=9°; FOV= 256 mm; voxel size 1.0 x 1.0 x 1.0 mm. As in Saygin et al., (2013), an online prospective motion correction algorithm was implemented to reduce the effect of motion artifacts during the structural scan, and 10 selective reacquisition time points were acquired and included to replace time points that were affected by head motion (Tisdall, 2012). The diffusion-weighted (DW) MRI scan included 10 non-diffusion-weighted volumes (b=0) and 30 diffusion-weighted volumes acquired with non-colinear gradient directions (b=700s/mm<sup>2</sup>), all at 128x128 base resolution and isotropic voxel resolution of 2.0mm<sup>3</sup>.

### **3.3.5 2<sup>nd</sup> Grade behavioral measures**

An extensive battery of assessments including phonological, RAN, cognitive, language comprehension, vocabulary, and reading measures was administered at the end of 2<sup>nd</sup> grade. For the current study, the following measures were used to characterize children's 2<sup>nd</sup> grade language and reading outcomes:

**Vocabulary.** The Peabody Picture Vocabulary Test (PPVT-4) assesses vocabulary knowledge.

**Language Comprehension.** The Understanding Spoken Paragraphs subtest of the Clinical Evaluation of Language Fundamentals (CELF-5).

**Word ID.** The Word Identification subtest of the Woodcock Reading Mastery Tests (WRMT-III, Woodcock, 2011) assesses single word reading skills.



**Reading Fluency and Comprehension.** The Gray Oral Reading Test (GORT-5) assesses oral reading rate, accuracy, fluency, and comprehension.

**Spelling.** Two measures were administered to evaluate spelling accuracy, the Test of Written Spelling (TWS-5) and the Test of Orthographic Competence-Homophone Choice subtest (TOC-HC).

**Verbal Short Term Memory.** The Memory for Digits subtest from the Comprehensive Test of Phonological Processing was administered (CTOPP, Wagner et al., 1999). Participants hear and then are required to correctly repeat strings of digits.

### **3.3.6 Behavioral analysis and statistics**

Structural MRI images were available for a subset of the participants who were selected to take part in the neuroimaging study ( $n = 100$ , who fell into the following groups: Average Performers=29, LSK risk=27, Multi-risk=12, RAN risk=23, PA risk=9). The four risk profiles (i.e. LSK risk, Multi-risk, RAN risk, PA risk) were then compared to the control group (Average performers) and to each other with respect to KG and 2nd Grade behavioral measures using one-way ANOVA. Planned pairwise comparisons were carried out with Games-Howell correction applied due to unequal sample sizes among the five profiles. PA risk was excluded from the 2<sup>nd</sup> Grade analysis due to the small group size of 4 (for the longitudinal sample).

### **3.3.7 VBM analysis**

Voxel-based morphometry (Ashburner et al., 2014) a whole-brain analysis technique, was used to identify significant grey matter volume reductions in each

of the risk groups as compared to the Average group. The VBM8 toolbox ([www.dbm.neuro.unijena.de/vbm](http://www.dbm.neuro.unijena.de/vbm)) was employed using SPM8 software ([www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)) executed in MATLAB (Mathworks, Natick, MA). All images were visually inspected for image quality. Two observers who were blind to behavioral scores and LPA group membership rated each image on a scale of 1 (unusable, very low quality) to 4 (excellent quality) based on a visual rubric of structural brain images artifacts associated with motion. If ratings differed, a third blind observer made the final decision. Ten participants were excluded because of poor image quality (i.e. image quality=1). Crucially, structural brain image quality did not differ among the groups, Average ( $M=3.31$ ,  $SD=0.6$ ), LSK-risk ( $M=2.96$ ,  $SD=0.9$ ), Multi-risk ( $M=3.23$ ,  $SD=1.17$ ), RAN-risk ( $M=3.39$ ,  $SD=0.6$ ), or PA-risk ( $M=2.78$ ,  $SD=0.8$ ),  $F(4,100) = 1.7$ ,  $p = 0.16$ ) nor was image quality significantly correlated with any literacy scores across all participants (all  $p$ 's > 0.18).

A customized tissue probability map was created in the Template-O-Matic toolbox (Wilke et al., 2008) using data from the Cincinnati Children's Hospital Medical Center sample (<https://irc.cchmc.org/software/tom.php>) with specification for the mean age of our pediatric population (5 years). Structural images were preprocessed using the following stream: segmentation to separate grey matter, white matter, and CSF; bias-correction to correct field inhomogeneity; and spatial normalization to a high-dimensional diffeomorphic anatomical registration using exponentiated linear algebra (DARTEL) template (Ashburner & Friston, 2000). The template was created based on 149 structural images of children (mean age=67.9 months, 50% female) from the larger sample of this study. Segmentation

quality was then assured with visual inspection and with a sample homogeneity test by plotting the standard deviation of the normalized, grey matter segmented brain volumes across all subjects. Analyses were performed with image outliers included and excluded, and since the results did not change, all images were included in the final analysis. Next, bias-corrected, whole brain Jacobian modulated images (preserving total grey matter volume) were smoothed with a 12-mm full width at half maximum isotropic Gaussian kernel (Ashburner & Friston, 2000).

To account for differences in sample sizes between and the Average group ( $n = 29$ ) and the LSK risk ( $n = 27$ ), PA risk ( $n = 9$ ), RAN risk ( $n = 23$ ), and Muti-risk ( $n = 12$ ) groups, subgroups were created which were matched for sample size, gender, and age. A series of whole brain t-tests was performed to compare grey-matter volume indices (GMVI) of each risk profile group to the average performance group. Explicit masking was employed in the general linear models using a study-specific analysis mask. A mean image was created using 148 segmented T1 grey-matter images that from the larger sample in this study. A threshold of 0.1 for the probability of a random field of noise was applied to the mean image to create the grey-matter mask used in analysis. Monte-Carlo cluster-level correction for multiple comparisons was conducted using the Alphasim tool in the REST toolbox (Song et al., 2011); the following parameters were used for the simulation: Gaussian filter width FWHM = 19.8 mm, 20 mm, 20 mm, rmm=5, voxel-wise  $p$ -threshold = .001, iterations = 5000. A cluster-level  $p$ -threshold of .05 was identified for a minimum of 661 voxels. Following the whole-brain analysis, a region of interest (ROI) analysis using the MarsBar toolbox (Brett et al., 2002) was

conducted for the ventral occipito-temporal cluster, representing the putative visual word form region, from the LPA1>LPA4 comparisons. Mean GMV from the cluster was extracted. To test whether the GMV and FA in the regions that showed reductions in the Multi-risk profile significantly explained additional variance in the reading score in 2<sup>nd</sup> grade over behavioral measures alone, a stepwise regression procedure was implemented in SPSS software package, version 22.0 (SPSS Inc., 1999).

### **3.3.8 *DWI analysis***

A subset of participants ( $n = 84$ ) who had usable DWI data and belonged to one of the five LPA profiles was included in the analysis. Diffusion-weighted images were visually checked for motion artifact and processed using FreeSurfer's TRACULA and FSL's FDT software (<http://www.fmrib.ox.ac.uk/fsl/fdt/index.html>). Data was processed and tracts of interest were defined using the same method described previously in Saygin et al. (2013). Briefly, diffusion-weighted images were registered to b0 image to correct for motion and eddy-current distortions. Tracts were defined using a probabilistic framework that incorporated information from an atlas of manually labeled tracts and Free Surfer anatomical segmentation to estimate the likelihood of each tract passing through a specific anatomical region. These anatomical priors were incorporated to guide tract reconstruction of each individual subject allowing for inter-subject variations. FSL's DTIFIT estimated the tensor fits at each voxel, which produced FA images. After visual inspection, FA values were averaged over the tract per individual. Tensor fits were only used to calculate the FA measures

and not in the tractography algorithm itself. Mean signal-to-noise ratio (SNR) values for all DWI were extracted for each subject as a proxy for head motion during DWI scan acquisition.

As in the GMV analyses, average FA values of the LSK risk ( $n = 16$ ), RAN risk ( $n = 22$ ), and Multi-risk ( $n = 8$ ) groups were compared to those of the Average group ( $n = 29$ ). The PA risk group was excluded from the analysis due to a small sample size ( $n = 6$ ). Three white matter tracts of interest which connect critical components of language and reading networks, and that have been most often associated with variation in reading ability were examined: left ILF, SLF, and AF (Saygin et al., 2013; Wang et al., 2016; Yeatman, Dougherty, Ben-Shachar & Wandell, 2012).

### **3.3.9 Power analysis**

A power analysis was conducted to evaluate whether differences between groups defined by a large effect size (based on Cohen, 1992) could be detected. An ANOVA with 4 groups and a semipartial correlation equal to .45, with power levels equal to 80% suggested that 60 participants would suffice to identify between groups differences that were equal to greater than a large effect size a number that was in accord with the longitudinal sample involved in the present study.

## **3.4 Results**

### **3.4.1 KG behavioral results**

The six latent profiles were characterized based on the patterns of performance across the five kindergarten measures (*Chapter 1* for detailed description of the profiles). A one-way ANOVA examining differences among the

groups on Word ID performance was significant ( $F(4) = 6.46, p < 0.001$ ). Post-hoc Games-Howell comparison revealed significantly lower performance of the LSK risk ( $p = 0.002$ ) and the Multi-risk ( $p < 0.001$ ) groups as compared the Average group ( $p < 0.001$ ). Because this paper is focused on risk, the sixth, “high-achieving” profile was excluded from further analysis.

### **3.4.2 Background measures results**

Differences in home literacy, SES, and family history of dyslexia among LPA groups were examined. The five included LPA groups did not significantly differ in any of the home literacy or SES variables (Table S1) except on paternal education, where the PA risk group demonstrated significantly lower levels of education than the average group ( $p = 0.026$ ) and lower number of parent books at home ( $p < 0.002$ ). Additionally, Pearson Chi Square analysis revealed no differences in distribution of gender ( $X^2 = 1.81, p = 0.77$ ), paternal ( $X^2 = 10.32, p = 0.24$ ) or maternal ( $X^2 = 13.53, p = 0.9$ ) familial dyslexia risk across the profiles.

### **3.4.3 2<sup>nd</sup> Grade behavioral results**

To test differences in reading outcomes across the four risk groups, a one-way ANOVA of 2<sup>nd</sup> Grade measures (Table 1) revealed significant effects of group on WRMT Word ID [ $F(4,86) = 5.57; p < 0.0001$ ], GORT Reading Fluency ( $F(4,87) = 4.32, p < 0.003$ ), GORT Reading Comprehension ( $F(4,87) = 5.15, p < 0.001$ ), CELF Understanding Spoken Paragraphs ( $F(4,85) = 3.96, p < 0.005$ ), PPVT Vocabulary ( $F(4, 84) = 5.31, p < 0.001$ ), CTOPP Memory for Digits ( $F(4,88) = 3.02, p < 0.022$ ) and TWS Spelling performance ( $F(3, 54) = 2.88, p < 0.027$ ). Groups did not differ significantly in age ( $F(4, 57) = 1.94, p = 0.941$ ). Post

hoc analysis using Games-Howell tests (robust to unequal variances between groups) revealed differences between the average versus other groups as indicated in Table 3.1.

	<b>Average Performers</b>	<b>LSK Risk</b>	<b>Multiple Risk</b>	<b>RAN Risk</b>	<b>PA Risk</b>
	Mean $\pm$ SD [n]	Mean $\pm$ SD [n]	Mean $\pm$ SD [n]	Mean $\pm$ SD [n]	Mean $\pm$ SD [n]
<b>Kindergarten</b>					
Age (in months)	66.34 $\pm$ 4.47 [29]	67.59 $\pm$ 3.63 [27]	69.54 $\pm$ 3.07 [13]	66.83 $\pm$ 4.89 [23]	65.11 $\pm$ 3.41 [9]
IQ	102.07 $\pm$ 9.98 [29]	98.41 $\pm$ 7.79 [27]	97.15 $\pm$ 7.02 [13]	100.70 $\pm$ 10.62 [23]	94.56 $\pm$ 7.32 [9]
WID	109.72 $\pm$ 18.23 [29]	91.89 $\pm$ 15.44 [27]*	88.85 $\pm$ 9.93 [13]*	107.78 $\pm$ 25.82 [23]	112.78 $\pm$ 10.18 [9]
RAN	98.79 $\pm$ 20.90 [29]	95.93 $\pm$ 11.16 [27]	86.27 $\pm$ 16.52 [13]	70.96 $\pm$ 23.53 [23]**	94.56 $\pm$ 14.88 [9]
PA	10.60 $\pm$ 1.23 [29]	8.89 $\pm$ 1.07 [27]**	8.08 $\pm$ 1.67 [13]**	11.39 $\pm$ 1.43 [23]	7.22 $\pm$ 0.57 [9]**
VSTM	8.28 $\pm$ 2.07 [29]	8.59 $\pm$ 2.49 [27]	8.62 $\pm$ 2.33 [13]	10.22 $\pm$ 2.35 [23]*	5.22 $\pm$ 0.97 [9]**
LSK	100.93 $\pm$ 1.39 [29]	89.00 $\pm$ 9.92 [27]**	79.38 $\pm$ 5.71 [13]**	106.43 $\pm$ 9.53 [23]	101.56 $\pm$ 7.86 [9]
<b>Grade 2</b>					
Age (in months)	98.67 $\pm$ 4.43 [28]	99.32 $\pm$ 4.15 [25]	101.2 $\pm$ 6.18 [10]	98.24 $\pm$ 4.58 [21]	97.67 $\pm$ 2.78 [9]
WID	108.48 $\pm$ 10.39 [27]	101.76 $\pm$ 13.41 [25]	87.22 $\pm$ 11 [9]**	104.81 $\pm$ 11.57 [19]	100.67 $\pm$ 14.25 [9]
Comprehension	10.18 $\pm$ 1.79 [28]	8.68 $\pm$ 2.1 [25]*	7.4 $\pm$ 1.51 [10]**	9.53 $\pm$ 1.84 [19]	8.89 $\pm$ 1.45 [9]
Fluency	10.54 $\pm$ 2.13 [28]	9.20 $\pm$ 2.6 [25]	7.30 $\pm$ 1.83 [10]*	9.55 $\pm$ 2.01 [20]	9.67 $\pm$ 1.41 [9]
Language Comprehension	12.54 $\pm$ 1.75 [28]	11.08 $\pm$ 1.63 [25]*	10.78 $\pm$ 1.92 [9]	11.68 $\pm$ 1.11 [19]	11 $\pm$ 1.58 [9]
Spelling	108.74 $\pm$ 10.16 [28]	100.48 $\pm$ 16.24 [25]	92.78 $\pm$ 10.02 [9]**	107.53 $\pm$ 15.34 [19]	104.11 $\pm$ 17.55 [9]
Memory for Digits	10.22 $\pm$ 2.59 [28]	9.08 $\pm$ 2.02 [24]	9.20 $\pm$ 3.08 [10]	10.43 $\pm$ 1.91 [20]	7.67 $\pm$ 2.24 [9]
Vocabulary	118 $\pm$ 13.6 [28]	110.54 $\pm$ 10.72 [24]	104.33 $\pm$ 14.19 [9]	120.53 $\pm$ 11.88 [19]	105.44 $\pm$ 8.5 [9]*



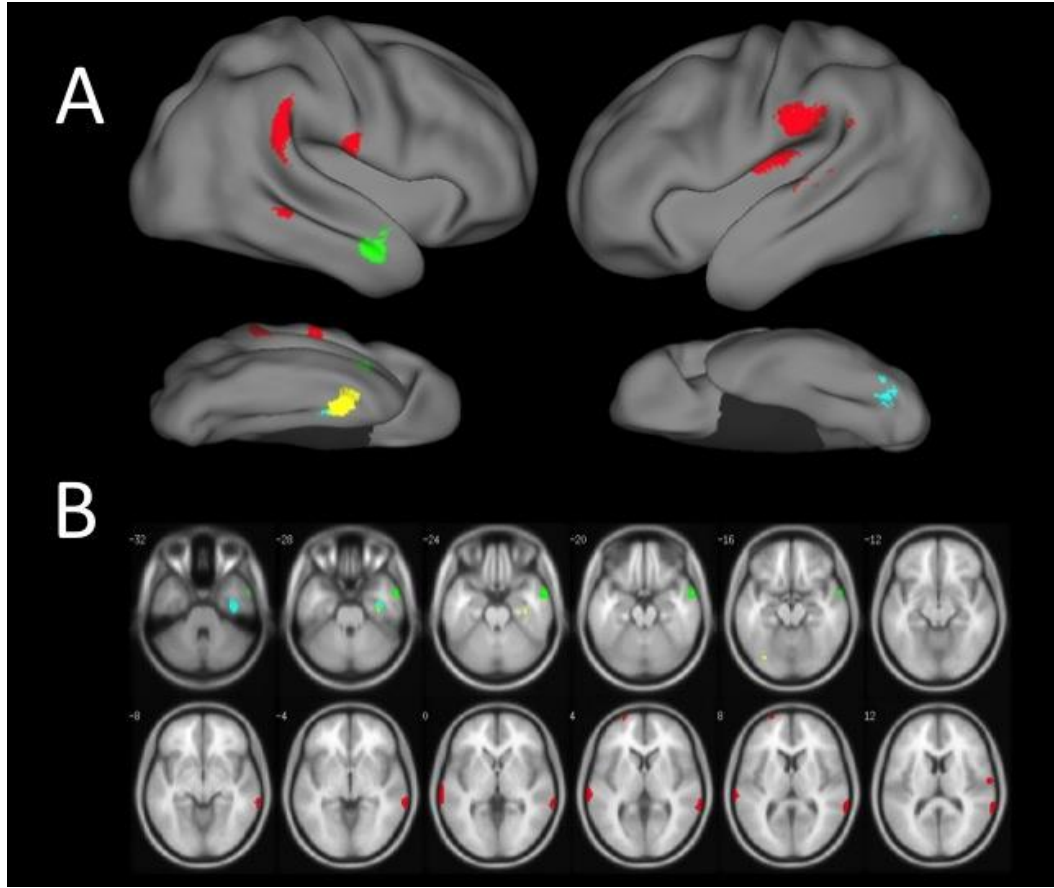
**Table 3.1:** Comparison of behavioral scores across the LPA groups. Group mean significantly below Average Performers with Games-Howell post-hoc comparison: \*  $p \leq .05$ ; \*\*  $p \leq .01$ ; Standard scores are reported for all measures.

#### **3.4.4 KG VBM results**

To investigate differences between the risk groups and the Average group in GMV, a whole-brain ANCOVA with group membership as a factor, controlling for handedness and sex was conducted. Results at the uncorrected threshold of  $p=0.001$ ,  $k=20$  voxels, revealed a main effect of group in the gray matter volume of the right supramarginal gyrus, right superior temporal gyrus, and medial prefrontal cortex. To investigate whether each risk profile was associated with distinct grey matter alterations as compared to the average group, whole-brain two-sample t-test comparisons were then conducted to investigate the specific differences in grey matter volume indices (GMVI) ( $p = 0.001$ ,  $k = 20$  threshold). Results revealed regional reductions in GMVI associated with each of the risk groups. Several such differences were observed (Table 3.2). Significant GMVI reductions were observed in the bilateral temporo-parietal and prefrontal regions for the LSK risk group as compared to the Average group; a right hippocampal cluster extending laterally, a left occipito-temporal region, and a left paracentral lobule for the Multi-risk group; right anterior middle temporal regions in the RAN risk group; and in medial temporal regions and right inferior parietal regions for the PA risk group. Only the right and left superior temporal regions in the Average>LSK risk comparison survived the cluster-extent corrected for multiple comparisons threshold ( $p < 0.001$ )

Cluster size	Region	Hemisphere	Coordinates			Z value
			x	y	z	
Average > LSK Risk						
1636	Angular Gyrus	R	69	-33	34	4.21 *
426	Insula	L	-36	-13	25	3.91
430	Superior Temporal/Middle Temporal	L	-70	-28	3	3.87
755	Supramarginal/Superior Temporal	L	-58	-25	27	3.86 *
304	Superior Temporal/Postcentral	R	63	-12	16	3.75
253	Superior Occipital/Cuneus	R	18	-93	30	3.50
72	Superior Frontal/Middle Frontal	L	-22	71	7	3.30
52	Inferior Parietal	L	-30	-58	52	3.26
Average > Multiple Risk						
68	Fusiform Gyrus/Parahippocampus/Hippocampus	R	34	-16	-27	3.58
85	Paracentral Lobule	R	6	-36	73	3.37
27	Occipito-temporal	L	-32	-75	-15	3.28
Average > RAN Risk						
443	Middle Temporal Pole/Inferior Temporal	R	60	4	-26	4.1
Average > PA Risk						
608	Fusiform/Inferior Temporal/Parahippocampal	R	44	-15	-41	3.82
125	Postcentral Gyrus/Inferior Parietal	R	22	-52	46	3.57

**Table 3.2:** Brain regions showing significant differences in gray matter volume index (GMVI) between LPA groups. Clusters showing reduced GMVI for risk groups compared to the average group are reported, threshold  $p = 0.001$ ,  $k=20$ , uncorrected. \*Survives cluster-level threshold,  $p < 0.05$ , corrected for multiple comparisons using Monte Carlo simulation.

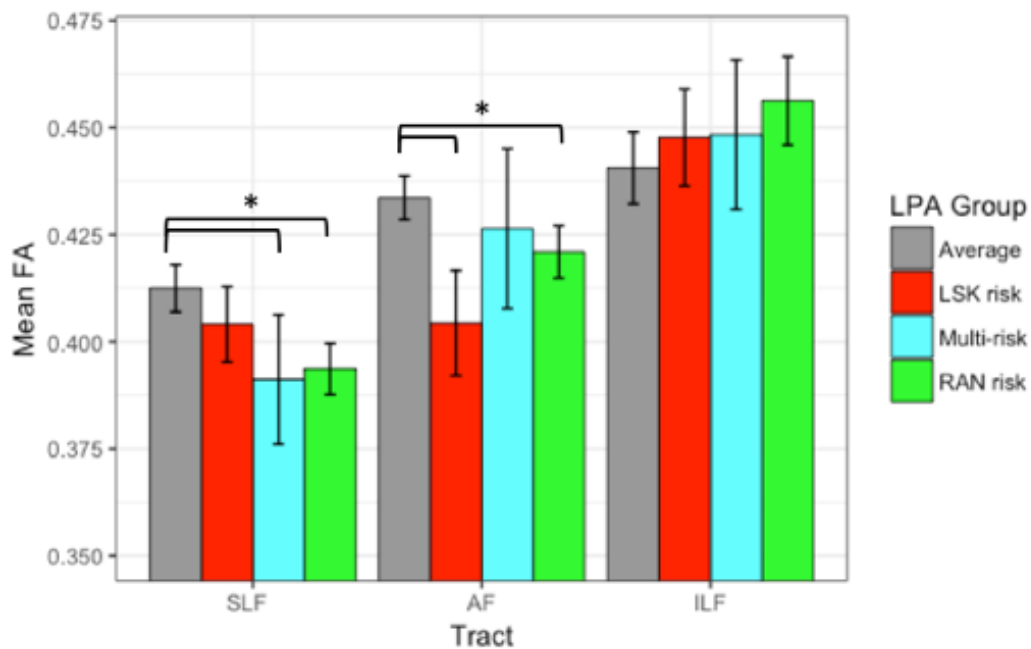


**Figure 3.1:** Grey matter volume differences as revealed by t-test comparisons ( $p = 0.001$ ) between each risk group and the average group for cluster size  $>20$  voxels [red=Avg.>LSK risk, cyan=Avg.>Multi-risk, green=Avg.>RAN-risk, yellow=Avg.>PA-risk] (A) Lateral and ventral views (B) Horizontal slice view

### 3.4.5 *KG DWI results*

Three factorial ANOVAs were conducted to compare the mean FA values in the left ILF, AF, and SLF tracts among the groups, while controlling for age in months, gender, handedness, and SNR of the DWI images (a proxy for motion effects). (The PA group was excluded from this analysis due to small  $n$  with usable data.) There was a significant main effect of LPA on the mean value of left SLF ( $F(3, 52) = 4.376, p < 0.05$ ) and of LPA on the mean value of the left AF ( $F(3, 52)$

= 2.803,  $p < 0.05$ ). The main effect of LPA on left ILF was not significant. Post hoc comparisons using the Games-Howell test indicated that the Average group had significantly higher mean FA values in the left SLF than the Multi-risk and RAN risk groups ( $p < 0.05$ ,  $d = 0.587$  and  $d = 0.6$ , respectively). For the AF tract, the Average group had significantly higher FA than the LSK risk group and the RAN group ( $p < 0.05$ ,  $d = 0.856$  and  $d = 3.84$ , respectively).



**Figure 3.2.** Average fractional anisotropy of left SLF, AF, and ILF by LPA group. Error bars indicate  $\pm 1$  SEM per group. \* $p < 0.05$ .

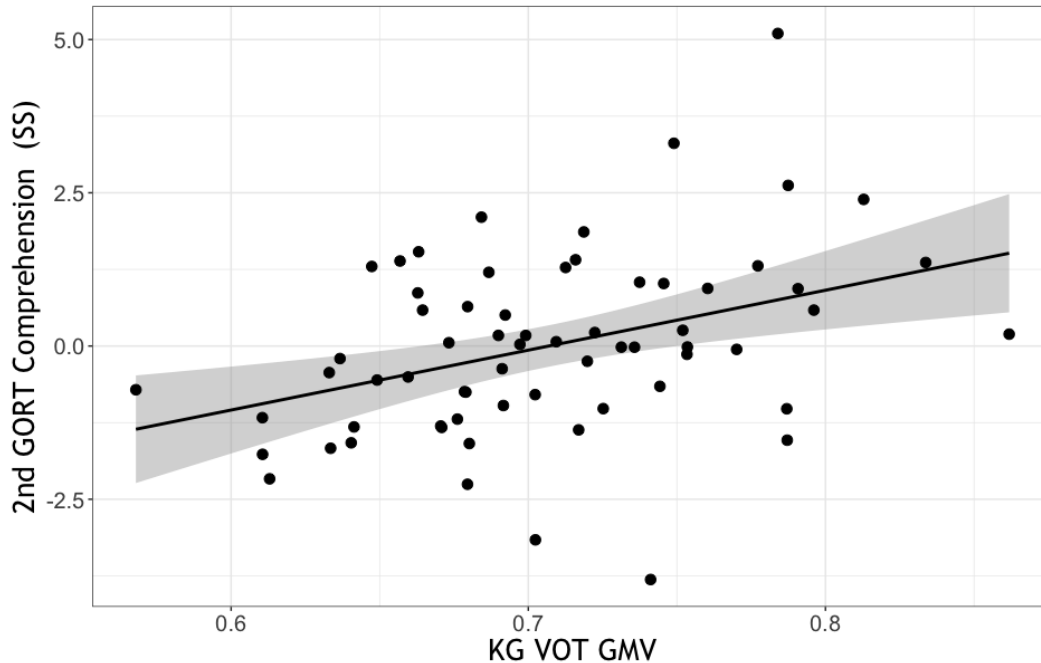
### 3.4.6 Longitudinal prediction analysis

In order to investigate whether neuroimaging measures significantly contribute to prediction of poor reading performance, a stepwise regression model (Table 3) was constructed in the subset of the children who have complete longitudinal data ( $n = 67$ ) with 2<sup>nd</sup> GORT Comprehension and Fluency scores as

the outcome. The brain measures were selected for this analysis are the ones that significantly differed between the Average and the Multi-risk groups. Age, gender, handedness were entered in the first block, kindergarten behavioral measures (i.e. PA, RAN, VSTM, and LSK) in a second block, mean FA in the left SLF tract as a third block, and mean GMVI extracted from three regions showing significant GMV reductions in the Multi-risk group (i.e. left occipito-temporal, right medial temporal, paracentral) as the final block. As depicted in Table 3, the overall regression model predicting 2<sup>nd</sup> grade Reading Fluency was significant ( $F(10,56) = 3.684, p < 0.001$ ) and accounted for 39.7% of variance. The model for Comprehension was also significant ( $F(10,55) = 6.81, p < 0.001$ ), accounting for 54.6% of total variance. The neuroanatomical measures significantly accounted for unique variance in reading performance, above the behavioral measures with the mean SLF FA accounting for unique 6.1% in fluency and 6.9% in comprehension and with GMV in the three ROI's accounting for 11.5% in reading comprehension, with left ventral occipito-temporal and right medial temporal regions being the significant predictors ( $p < 0.05$ ).

	Comprehension					Fluency				
	B	SE B	$\beta$	t	p	B	SE B	$\beta$	t	p
<b>Step 1</b>	<b>R<sup>2</sup>=0.181, p=0.002</b>					<b>R<sup>2</sup>=0.22, p&lt;0.001</b>				
Age	-.154	.051	-.347	-3.020	.004	-.221	.055	-.449	-4.020	.000
Gender	-.797	0.442	-.207	-1.804	.076	-.369	.478	-.086	-0.771	.444
<b>Step 2</b>	<b>R<sup>2</sup>=0.362, p=0.005</b>					<b>R<sup>2</sup>=0.313, p=0.102</b>				
RAN	.036	.016	.184	2.272	.027	.045	.018	.279	2.510	.015
VSTM	.016	.120	-.002	0.124	.902	.109	.143	.128	.763	.448
PA	.290	.180	.227	1.545	.128	.122	.214	.103	0.569	.571
LSK	.039	.020	.222	2.095	.040	-.002	.023	-.009	-.074	.941
<b>Step 3: DWI</b>	<b>R<sup>2</sup>=0.431, p=0.01</b>					<b>R<sup>2</sup>=0.373, p=0.02</b>				
SLF	15.441	5.829	.276	2.649	.010	16.097	6744.000	.258	2.387	.020
<b>Step 4: GMVI</b>	<b>R<sup>2</sup>=0.546, p=0.006</b>					<b>R<sup>2</sup>=0.397, p=0.537</b>				
L VOT	7.840	3.272	.235	2.396	.020	14.530	4.163	.052	.470	.640
Paracentral	2.504	2.730	.089	.917	.363	-2.628	3.453	-.084	-0.761	.450
R MT	6.072	2.941	.209	2.065	.044	4.516	3.716	.140	1.215	.229

**Table 3.3.** Results of stepwise multiple regression analyses predicting reading comprehension and fluency. L VOT-left ventral occipito-temporal cluster; R MT-right medial temporal cluster



**Figure 3.3.** KG grey matter volume of the ventral occipito-temporal region predicting unique variance in GORT Comprehension

#### **3.4.7 Bias analysis**

Results indicate that the largest biases barely exceeded or distorted the mean estimates by 1%. Values across analyses ranged between 0 and 1.09%, suggesting that bias due to idiosyncrasies in the sample was unlikely in the present study.

### **3.5 Discussion**

The current MRI study investigated alterations in grey matter volume and white matter coherence among groups of children with risk profiles identified through latent profile analysis of a large kindergarten (KG) sample: the LSK, Multi, RAN and PA risk groups. Voxel-based morphometry (VBM) analysis revealed significantly reduced grey matter volume (GMV) indices in each of the risk groups, as compared to the Average group, in regions that support reading development. A diffusion-weighted imaging (DWI) analysis revealed that risk was associated with

significantly reduced mean fractional anisotropy in two tracts important for reading: the left superior longitudinal fasciculus and the left arcuate fasciculus. GMV and DWI measures in kindergarten each accounted for additional variance in 2<sup>nd</sup> grade fluency and comprehension above the contribution of kindergarten behavioral scores. Further, these group differences were not likely due to socioeconomic or home literacy measures, as these did not differ between groups.

### **3.5.1 Behavioral and grey matter group characteristics**

#### **3.5.1.1 LSK risk profile**

Behaviorally, the LSK risk group was characterized by reduced performance on letter sounds knowledge (LSK) and word ID (WID) and below-average phonological awareness (PA) performance at *KG*, although the PA deficit was milder than the PA or Multi-risk groups. At *2<sup>nd</sup> Grade* the group demonstrated poor performance on vocabulary, reading, and language comprehension measures. Initially, the profile was characterized as orthographic risk due to selective impairment on LSK and WID, but the low performance on *2<sup>nd</sup> Grade* reading and language comprehension measures is indicative of a general deficit in language comprehension. Children with similar deficits, who have typical word reading and decoding abilities, but a pervasive deficit in reading comprehension, have previously been described as “poor comprehenders” (Nation, Clarke, Marshall, & Durand, 2004; Catts, Adlof, Weismer, 2006). Similarly to the current results, in another study, a retrospective analysis of poor comprehenders revealed below-average performance on PA in kindergarten, but not in later grades (Catts & Weismer, 2006). In another study poor performance on LSK in kindergarten was



predictive of poor comprehension in tenth grade (Stanley, Petscher, Catts, 2018). This suggests that the LSK risk group's deficits in LSK, PA, and word reading at *KG* was likely due to language factors such as semantic knowledge influencing the early development of phonological awareness and visual representations of letters and words (Metsala & Walley, 1998).

The LSK risk profile demonstrated reduced GMV in temporo-parietal regions bilaterally, in accordance with previous findings that showed functional alterations in these regions in pre-reading children with poor letter-knowledge skills (Specht et al., 2009; Yamada, 2011). Meta analyses of older children and adults with dyslexia consistently report structural and functional atypicalities in bilateral temporo-parietal regions (Richlan 2011; 2013; Linkersdörfer et al., 2012). Additionally, in developing readers, temporo-parietal regions showed involvement in linking speech units to their letter-representations (Turkeltaub, 2003; Yamada, 2011). For example, a longitudinal fMRI study of kindergarten children from pre-reading to reading stage, demonstrated initial involvement of the left hemispheric region in audio-visual integration during a phonological task, with gradual decrease across more advanced stages of reading (Yu et al., 2018). This support the significance of the temporo-parietal region for early letter-to-sound mapping skill development.

There is also significant evidence for the involvement of the bilateral temporo-parietal regions in semantic processing. Neuroimaging studies of natural language processing have attributed the processing of multiple linguistic cues (e.g., syntactic, phonological, semantic) important for comprehension to the superior

temporal and parietal regions bilaterally (Cutting et al., 2006; Friederici, Meyer, & von Cramon, 2000; Landi, Frost, Mencl, Sandak†, & Pugh, 2013; Price, Bonner, Peelle, & Grossman, 2015). Additionally, reduced activation for sentence comprehension in these regions has been demonstrated in poor readers (Meyler et al., 2007) and specifically, in individuals with a reading comprehension deficit (Grossman et al., 2002). Thus, the location of the brain differences in the LSK group is consistent both with the initial behavioral deficit of the group on measures of letter-sound integration and the later deficits in language comprehension.

#### 3.5.1.2 Multi-risk profile

The Multi-risk profile was characterized by low performance across all *KG* measures, and deficits in reading and spelling at *2<sup>nd</sup> Grade*. Importantly, *2<sup>nd</sup> Grade* performance was within the range of 1 SD below the mean, consistent with the definition of dyslexia in previous literature (e.g., Perrachione, Del Tufo, & Gabrieli, 2011; Ramus, Pidgeon, & Frith, 2003). The Multi-risk group demonstrated reduced GMV in left occipito-temporal and right medial temporal regions relative to the average group. The left occipito-temporal region corresponds to the putative visual word form area (VWFA), a hub of a system that develops selectivity for letter and word form as the result of increased experience with reading (Cohen et al., 2002; Dehaene & Cohen, 2011). Children and adults with dyslexia have demonstrated attenuated functional activation, reduced cortical thickness and grey matter volume in this region (Shaywitz et al., 2007; Van der Mark et al., 2009). In a recent fMRI study, 5-year-old kindergarteners who could read, but not their peers who were pre-readers, demonstrated selectivity in the VWFA to orthographic stimuli (Saygin et

al., 2016). In another study, in beginning readers in kindergarten prior to reading instruction, letter specialization in the VWFA was associated with better word identification skills (Centanni et al., 2018). These findings suggest that current results could be due to reading performance differences between the Average and the Multi-risk groups at *KG*. Indeed, the Average group performed significantly better than the Multi-risk group on LSK and Word ID measures in *KG*. Alternatively, in another study, hypoactivation in the VWFA region in individuals with dyslexia compared with reading-matched controls was demonstrated, suggesting that this hypoactivation is uniquely related to dyslexia rather than to reading ability (Hoeft et al., 2007). Additionally, in a cross-sectional study that investigated functional activation during an implicit word processing task across development, neural specialization for reading involved the disengagement of the homologue right-hemispheric occipito-temporal regions, rather than increased activation in the left hemispheric region (Turkeltaub, 2003). Thus it is also possible that the differences in the region in the Multi-risk group are a marker of dyslexia that is present early in acquisition of literacy, and could be causal to reading failure.

The Multi-risk group also demonstrated reduced GMV in the right medial temporal region, including the hippocampal and parahippocampal areas that are known to support memory and learning (Squire, 1992). The declarative memory system, supported by the regions, underlies the acquisition of knowledge of events and facts, thereby playing a general, albeit important, role in learning many of the foundational skills required for reading development (Nicolson & Fawcett, 2007). These areas have also been implicated in rapid statistical learning and implicit

memory (Schapiro, Turk-Browne, Norman, & Botvinick, 2016; Turk-Browne, Scholl, Chun, & Johnson, 2009). Structural and functional alterations in these regions have been reported in previous studies of individuals with dyslexia (Brambati, Ogar, Neuhaus, Miller, & Gorno-Tempini, 2009; Casanova et al., 2005; Eliez et al., 2000) and their involvement in word recognition and reading has been demonstrated (Papanicolaou et al., 2003; Piai et al., 2016). Genetic studies linked hippocampal alterations with dyslexia; alterations of dyslexia susceptibility genes in rodents (e.g., *Dyx1c1*, *Dcdc2*, *Kiaa0319*) caused atypical hippocampal neuronal migration (Burbridge et al., 2008; Peschansky et al., 2010; Szalkowski et al., 2012).

It is possible that the pattern of deficit across multiple pre-literacy measures of the Multi-risk group reflects a general learning impairment that hampers the development of the reading prerequisite skills necessary for reading development, resulting in the dyslexia-like profile exhibited by the group in 2<sup>nd</sup> grade. Other studies, however, suggest that dyslexia is characterized by implicit and procedural deficits in extraction of stimuli regularity, rather than by an impaired declarative memory system (Nicolson & Fawcett, 2007; Vicari et al., 2005). Indeed, there is increasing evidence for the involvement of the hippocampal/parahippocampal regions in extraction of regularity (Turk-Browne et al., 2009; Schapiro et al., 2018). The findings of reduced GMV in these regions in the multi-risk group is in line with rapid learning deficit theories of dyslexia (Ahissar et al., 2006; Gabay et al., 2015; Perrachione et al., 2017). An alternative hypothesis is that the declarative memory system that underlies the development of semantic knowledge, for example, may play an important role in the development of compensation strategies in individuals

with dyslexia (Bookheimer, 2002; Nation & Snowling, 2004; Ullman & Pullman, 2015). In the dyslexia risk group, alterations in this system reflected their inability to develop compensatory mechanisms resulting in pervasive reading deficits. In support of this hypothesis, previous studies demonstrated an increase in GMV in medial temporal regions in individuals with dyslexia following intervention (Krafnick et al., 2011; Temple et al., 2003).

#### 3.5.1.3 RAN risk profile

The RAN risk group was characterized behaviorally by low performance on RAN at *KG*, and demonstrated typical reading performance on average at *2<sup>nd</sup> Grade*. This typical performance is somewhat unexpected, given that children with RAN deficit often demonstrate failure to develop automaticity in reading as evident by reduced performance on reading fluency and other timed reading measures (for a review see Norton & Wolf, 2012). Since most of the RAN deficit studies were performed in older children and adults, however, it is possible that the fluency deficit emerges in later grades, when reading automaticity is present in the typically-reading controls. Alternatively, it is also possible that RAN does not represent a persistent deficit for all children, with some children having slower naming speed during the task due to fatigue or distraction.

The RAN risk group demonstrated reduced GMV as compared to the Average group in the right anterior middle temporal lobe. This region is thought to support naming abilities as patients with lesions in the region demonstrate deficits in name retrieval across input modalities and stimuli (for a review see Abel et al., 2016). Furthermore, better performance on a task requiring retrieval of conceptual

knowledge from visual input was associated with increased GMV in the region (Acres, Taylor, Moss, Stamatakis, & Tyler, 2009). Thus, our finding of reduced GMV in the anterior middle temporal region in children with a deficit on RAN, a task that requires rapid identification of items and retrieval of their name, is consistent with the previously established functional role of this region. The only other study that investigated the neuroanatomical underpinnings of RAN in young readers similarly found that RAN correlated with activity in a right anterior middle temporal region during an implicit in-scanner word-reading task, as well as in bilateral right posterior temporal and left ventral inferior frontal gyri (Turkeltaub, 2003). In the few neuroimaging studies of adults and older children conducted to date, RAN is often associated with a distributed network of regions and there is little agreement across studies of which structures support rapid naming and are altered in individuals with a RAN deficit subtype (Norton & Wolf, 2012). Two regions that have most consistently emerged from these studies are left inferior frontal and right cerebellar regions (Eckert, 2003; He, Xue, Chen, Chen, et al., 2013; Jednoróg et al., 2014; Norton et al., 2014). To account for discrepancies in findings between the two studies in early-readers (including the current study) and studies of already-reading individuals, it is possible that the neuroanatomical profiles of mature readers who demonstrate deficits in fluency and rapid naming (i.e. RAN deficit) are different from those who show an initial deficit in rapid naming early in reading development (i.e. RAN risk).

### **3.5.2 PA risk profile**

This profile was characterized by a deficit in PA and VSTM performance at *KG* and reduced performance on VSTM and vocabulary skills, as compared to the Average group, at *2<sup>nd</sup> Grade*. Despite the pervasive deficit in different aspects of phonological processing (i.e., short-term memory and awareness), this group demonstrated no deficits in reading on average. Neuroanatomically, the group was characterized by reduced GMV in the right medial temporal (including parahippocampal), fusiform, and inferior parietal regions. Both the behavioral deficits and the location of the neural differences in this group indicate a deficit in general systems supporting working memory. Inferior parietal regions bilaterally have been associated with working memory for verbal and non-verbal information (Beneventi, Tønnessen, Ersland, & Hugdahl, 2010; Nee et al., 2013; Owen, McMillan, Laird, & Bullmore, 2005). Additionally, in an fMRI study of working memory performance on an n-back task, children with dyslexia demonstrated reduced activation in multiple regions, including in the bilateral inferior parietal lobule (Beneventi, Tønnessen, et al., 2010).

The region of the GMV reduction in the medial temporal lobe of the PA risk group, overlapped with that of the Multi-risk group, suggesting the possibility that both short-term and long-term memory systems may be impaired in this group affecting the development of phonological skills. Since this group didn't demonstrate GMV reductions, evident in the Multi-risk group, in the left occipito-temporal region they were able to develop typical reading abilities despite a phonological deficit. Indeed it has been suggested that impairments in aspects of

phonology (e.g., PA and VSTM) is often associated with dyslexia risk regardless of reading outcomes (Moll, 2013). Indeed, significantly higher paternal history of dyslexia in this group supports the idea that the observed phonological impairment is an endophenotype of familial risk.

### **3.5.3 *White matter alterations***

Three left white matter tracts were selected for the current analysis due to their established role in reading and reading related functions (Vandermosten, Boets, Wouters, et al., 2012; Yeatman, 2012a). In previous studies, white matter atypicalities in these tracts were observed in children and adults with dyslexia (Klingberg et al., 2000; Niogi & McCandliss, 2006; Scerri, Darki, Newbury, Whitehouse, Peyrard-Janvid, Matsson, Ang, Pennell, Ring, & Stein, 2012; Vandermosten, Boets, Poelmans, et al., 2012). In the current analysis, differences in mean FA between the Average group and the risk groups were evident for the SLF and AF, but not the ILF tracts. Failure to find group differences in ILF is in agreement with a cross-sectional study that investigated mean FA values in the three tracts in dyslexia risk and no-risk groups in pre-reading, beginning, and fluent reading stages (Wang et al., 2016). Differences in FA were not evident in ILF in kindergarten, but emerged in later grades (Wang et al., 2016). The left ILF tract connects the occipito-temporal regions to the anterior and medial temporal lobe and supports the mapping of visual word form representations to their lexical-semantic representations (Yeatman et al., 2011). According to the developmental account of neural specialization for reading, the occipito-temporal network that supports orthographic processing of words emerges after the temporo-parietal network that



supports phonological processing as the result of reading acquisition, and therefore isn't expected to be present in pre-readers (Pugh et al., 2001). A recent longitudinal study confirmed this hypothesis, demonstrating increased functional connectivity between the temporo-parietal and the occipito-temporal regions from pre-reading to reading stages (Yu et al., 2018). Thus, despite the current findings of GMV reductions in the occipito-temporal region in the Multi-risk group, suggesting alterations in the region even before it is specialized for reading and in line with previous descriptions of the emergence of the VWFA (Dehaene, Le Clec, Poline, Le Bihan, & Cohen, 2002; Dehaene et al., 2010), it is likely that the white matter connectivity between the VWFA region and the language network emerges with increased reading experience.

Whereas the LSK risk group demonstrated white matter reductions in the AF tract, the Multi-risk group demonstrated a lower mean FA in the SLF tract. Interestingly, the RAN risk group had a lower mean FA in both tracts. Prior studies often regarded AF to be a segment of the SLF due to their spatial proximity, causing difficulty separating voxels belonging to each of the tracts (Makris et al., 2005). Recent literature, however, suggests that the two tracts are distinct both neuroanatomically and functionally (Martino et al., 2013; Yeatman et al., 2011). The AF tract connects the middle and inferior temporal gyri with the precentral gyrus and posterior portion of the inferior and middle frontal gyri (Martino et al., 2012). It is thought to be involved in auditory-motor integration important for the manipulation and articulation of incoming phonological information (Hickok & Poeppel, 2004, 2007) and for articulatory-based processes that allow keeping

information active during working memory (Catani et al., 2007; Cunillera et al., 2009). Accordingly, AF has been associated with phonological skills (Catani & Jones, 2005; Friedmann & Gvion, 2003; Rilling et al., 2008; Saur et al., 2008; Yeatman et al., 2011) and word learning (Lopez-Barroso et al., 2013) across development. Furthermore, atypicalities in the AF tract have been reported in studies of at-risk pre-readers (Saygin et al., 2013; Vandermosten, 2015; Wang et al., 2016).

Alternatively, SLF connects the supramarginal gyrus and superior temporal gyrus with the precentral gyrus and has been implicated in verbal fluency and automaticity (Catani & Jones, 2005; Gold, Powell, Xuan, Jiang, & Hardy, 2007; Rauschecker et al., 2009). Although a case study demonstrated that a damage to this tract was associated with a failure in learning to read (Rauschecker et al., 2009), previous studies in pre-reading children at a behavioral (Saygin et al., 2013) and familial (Wang et al., 2016) risk for dyslexia didn't find group differences in this tract. Studies in older children and adults, consistently report reduced FA in SLF associated with dyslexia (Beaulieu et al., 2005; Carter et al., 2009; Deutsch et al., 2005; Klingberg et al., 2000; Niogi & McCandliss, 2006), but other studies in typical readers didn't find an association between the integrity of this tract and reading (Niogi & McCandliss, 2006; Odegard, Farris, Ring, McColl, & Black, 2009; Rollins et al., 2009). Thus, previous studies implicated AF in phonological processing and integration and, less consistently, the SLF tract in supporting automaticity processes important for reading.

The current findings are in accordance with the previously documented roles of each of the tracts reviewed above. Importantly, in line with the current findings, a longitudinal study in pre-readers at familial risk for dyslexia has demonstrated that the rate of FA change from the pre-reading to the reading stage in SLF was associated with reading fluency and the rate of change in AF was associated with reading comprehension (Wang et al., 2016). Thus, the low FA of the LSK group in the left AF seems to correspond to the initial deficit of the group on PA and LSK, and the later deficit in comprehension. The low FA of the Multi-risk group in SLF reflects, on the other hand, seem to reflect the more general deficits demonstrated by the group across various reading-related tasks. The RAN risk group had reduced FA in both of the tracts, suggesting altered connectivity in this group across systems that support both motor-speech integration and automaticity. This finding is in accordance with the notion that RAN performance requires rapid naming of stimuli and is thought to reflect the automaticity with which the various components of the reading circuit are integrated (Norton & Wolf, 2012; Wolf et al., 2002).

### **3.5.4 *Contribution of neural measures to prediction of reading outcomes***

In order to optimize educational outcomes and avoid the psychosocial consequences associated with failure to develop reading, an important goal of dyslexia research is to identify valid predictors of dyslexia prior to reading failure (Ozernov-Palchik & Gaab, 2016). Accordingly, in an effort to enhance the sensitivity and specificity of risk identification, previous studies in pre-readers combined behavioral and neuroimaging measures in reading outcome prediction

models and demonstrated that neuroimaging measures substantially enhance prediction accuracy (e.g., Bach, Richardson, Brandeis, Martin, & Brem, 2013; Hoeft, 2011a; Maurer et al., 2009; Preston et al., 2016; Wang et al., 2016). For example, in one study, functional activation and white matter integrity predicted reading gains over 2.5 years with 72% accuracy in children with dyslexia (Hoeft et al. 2011). In another study, volume changes in temporo-parietal white matter, together with preliteracy measures, accounted for 56% of the variance in reading outcomes (Myers et al. 2014). Consistent with these findings, the current model combined KG behavioral, GMV, and DWI measures and accounted for a large variance of 73.3% of 2<sup>nd</sup> reading comprehension and 68.9% of reading fluency performance. Importantly, mean FA in the SLF accounted for unique variance in both outcome measures, and mean GMV in left VWFA accounted for unique variance in reading comprehension. These findings reinforce previous research in suggesting that early brain alterations are important structural markers for dyslexia and could be used to enhance the accuracy of early risk identification.

### **3.5.5 Limitations**

The main challenge of conducting a neuroimaging study of multiple subtypes, rather than collapsing across risk groups, is that this approach limits the group sizes available for analysis. In the present study, the sample sizes for the DWI and the longitudinal analyses were relatively small and thus some of the findings could potentially reflect Type-II errors. If such errors were produced they would point to the fact that the findings that exceeded conventional levels of significance were rather robust and reflective of true relationships in the population of interest.

We attempted to overcome this limitation and increase confidence over stability of the estimated parameters using bootstrapping (Efron & Tibshirani, 1994). The method pointed to a negligible bias that was present around the mean estimates of the dependent variables. It is important, however, to acknowledge well-known limitations of bootstrapping. For example, bootstrapping can fail if the sample is not representative of the population with the resampling process merely reproducing the biases of the sample (Stine, 1989). Further work has suggested that bootstrapping using resampling fails to reproduce a sample's maximum estimate (Bickel & Freedman, 1981).

For the VBM analysis, uncorrected thresholds were employed for most comparisons, since in addition to smaller group sizes in the current risk groups, uncorrected thresholds are often used in pediatric data sets due to lower signal-to-noise ratio and high inter-individual variance (e.g., Thomason, Burrows, Gabrieli, & Glover, 2005). Additionally, in the current analysis we relied on cluster size permutation for cluster size estimates and these estimates may be influenced by intrinsic attributes of the data or processing techniques (e.g., resolution, smoothing kernel, minimum cluster level) (Strawn et al., 2015). Despite these limitations, the locations of the GMV alterations are consistent with previous literature and with the theoretical conceptualization of the different dyslexia risk profiles.

### **3.5.6 *Theoretical implications of findings***

Taken together, our findings demonstrate that different subtypes of dyslexia risk profiles are associated with distinct grey and white matter alterations in structures that support early reading development. These differences provide the

first imaging evidence for the heterogeneity of neuroanatomical attributes of dyslexia before children begin to read and suggest a distinct developmental trajectory of dyslexia risk profiles. Our findings have the potential to inform theories of dyslexia that aim to elucidate the etiological basis of this disorder. One leading theory is the *double deficit hypothesis* (Wolf & Bowers, 1999) that proposes that RAN and PA are two *independent* deficits of dyslexia. In contrast, according to the single-deficit models of dyslexia, RAN deficit is a failure to fluently access and retrieve phonologically based information and thus is an extension of the phonological deficit (Lervåg & Hulme, 2009; Ramus, 2003). There is mixed evidence from behavioral studies of dyslexia for the stochastic independence of PA and RAN constructs (Compton et al., 2001; Cronin, 2013; Schatschneider et al., 2002; Torgesen et al., 1997; Vukovic & Siegel, 2006; Wagner & Barker, 1994; Wolf & Bowers, 1999). Several studies to date have provided neuroimaging evidence for the dissociation of the PA and RAN skills and deficits in the brain (He, Xue, Chen, Chen, et al., 2013; Jednoróg et al., 2014; Norton et al., 2014). These studies were conducted in already reading children and adults, and therefore offer limited information about the independent neural development of the RAN and PA deficits in the brain. The dissociation of the neural substrates of the PA and RAN risk groups in the current study provides the first evidence that the development of these deficits is independent.

Furthermore, we demonstrated that children who have both deficits in RAN and PA (i.e., double-deficit group) have neural atypicalities that differ from those of the single deficit group. This suggests that this group's deficit is unique rather

than, as suggested by the double deficit hypothesis, additive and may represent underlying aberrations in both semantic and orthographic memory systems. Finally, this is the first study to identify neuroanatomical precursors of the poor comprehenders' profile. The findings provide information on the early cognitive and neuroanatomical signatures of this language-based deficit and the distinct developmental nature of this disorder. To evaluate the longitudinal significance of these findings for reading development, however, a longer follow-up of children from kindergarten to later grades is necessary. Overall, our findings illuminate the neurobiological nature of early dyslexia risk profiles and confirm the feasibility and importance of early identification of dyslexia risk and of targeted interventions that are tailored to the specific profile of risk.

	<b>Average Performers</b> Mean ± SD	<b>LSK Risk</b> Mean ± SD	<b>Multiple Risk</b> Mean ± SD	<b>RAN Risk</b> Mean ± SD	<b>PA Risk</b> Mean ± SD
<b><i>Kindergarten</i></b>					
Maternal Education	5.17±1.25	4.96±1.15	4.69±1.43	5.34±1.11	4.77±0.83
Paternal Education	5.24±1.09	4.28±1.33	4.15±1.62	5.18±1.29	3.77±0.83***
Total household income	8.04±1.59	7.55±2.13	8.25±3.37	8.17±1.87	6.88±2.84
Number of adults' books	3.53±2.31	3.6±2.14	2.5±1.08	3.72±2.29	1.5±0.53**
Number of children's books	4.34±1.69	3.9±1.68	3.5±1.43	3.77±1.43	3.25±1.98
Hours spent reading with child	5.12±3.03	3.82±2.23	3.83±3.24	4.93±3.93	4±1.78
Maternal ARHQ	.29±.19	.25±.14	.27±.11	.29±.12	.32±.16
Paternal ARHQ	.35±.11	.45±.23	.41±.12	.26±.09	.33±.11

**Supplemental Table 3.4:** Comparison of family and home environment measures across LPA groups

Group mean significantly below average performers with Games-Howell post-hoc comparison: \*  $p \leq .05$ ; \*\*  $p \leq .01$ ; \*\*\*  $p \leq .001$

## **4 Study 3: Relationships Between Early Literacy and Non-Linguistic Rhythmic Processes in Kindergarten Children<sup>9</sup>**

### **4.1 Abstract**

A growing number of studies report links between non-linguistic rhythmic abilities and certain linguistic abilities, particularly phonological skills. The current study investigated the relationship between non-linguistic rhythmic processing, phonological abilities, and early literacy abilities in kindergarteners. A distinctive aspect of the current work was the exploration of whether processing of different types of rhythmic patterns is differentially related to kindergarteners' phonological and reading-related abilities. Specifically, we examined the processing of metrical vs. nonmetrical rhythmic patterns, i.e., patterns capable of being subdivided into equal temporal intervals or not (Povel & Essens, 1985). This contrast is of interest because metrical sequences are characteristic of much music, in which rhythm often

---

<sup>9</sup> Portions of this chapter were originally published as Ozernov-Palchik, O., Wolf, M., & Patel, A. D. (2018). Relationships between early literacy and nonlinguistic rhythmic processes in kindergarteners. *Journal of experimental child psychology*, 167, 354-368.



involves an underlying temporal grid of isochronous units. In contrast nonmetrical sequences are arguably more characteristic of speech rhythm, which is temporally structured but does *not* involve an underlying grid of equal temporal units. A rhythm discrimination app with metrical and nonmetrical patterns was administered to 74 kindergarteners in conjunction with cognitive and pre-literacy measures. Findings support a relationship among rhythm perception, phonological awareness, and letter-sound knowledge (an essential precursor of reading). A mediation analysis revealed that the association between rhythm perception and letter-sound knowledge is mediated through phonological awareness. Furthermore, metrical perception accounted for unique variance in letter-sound knowledge above all other language and cognitive measures. These results point to a unique role for temporal regularity processing in the association between musical rhythm and literacy in young children.

## **4.2 Introduction**

In research on cognitive relations between music and language a topic of long-standing interest is the link between musical abilities (or training) and linguistic processes involved in reading. This issue has been investigated for over 40 years (e.g., Hurwitz, Wolff, Bortnick, & Kokas, 1975), due in large part to its relevance to real-world issues such as finding new ways to enhance reading development. This is especially relevant for children struggling with language-based deficits, such as developmental dyslexia. The significance of this topic for educational practice parallels its theoretical relevance to issues such as the modularity of language, i.e., the extent to which the cognitive and neural

mechanisms supporting language are domain-specific vs. shared with other domains, such as instrumental (non-verbal) music (Patel, 2008a).

A central question that emerges is how *non-linguistic* music processing---which utilizes non-verbal auditory and motor skills---can be related to reading, which involves mapping stored phonological and lexical representations onto visual representations. Several researchers suggest that one link may lie in the temporal processing of sound (e.g., Goswami, 2011; Lehongre, Ramus, Villiermet, Schwartz, & Giraud, 2011, Ch. 3; Tallal & Gaab, 2006; Tierney & Kraus, 2011). Reading acquisition depends heavily on phonological awareness (PA): that is, the ability to segment the seemingly continuous flow of speech into a sequence of perceptually discrete speech sounds, from words and syllables to phonemes. PA underlies the ability to blend phonemes and manipulate segmented speech sounds. These abilities, in turn, draw on auditory processes involved in analyzing the temporal structure of sound patterns. Indeed, relations between linguistic stress sensitivity and reading abilities have been found in several studies (Holliman, Wood, & Sheehy, 2008, 2010; Whalley & Hansen, 2006; Wood, 2006).

Like speech, instrumental music also relies on fine timing distinctions and structured patterns of duration in complex sound sequences that unfold rapidly in time (Patel, 2008b). Thus even though speech and instrumental music have many salient differences as acoustic patterns (Ding et al., 2017), a growing body of evidence suggests that some of the temporal processing mechanisms humans apply to these domains may be shared (see Kraus & Chandrasekaran, 2010 for one review). This could help explain why links between musical temporal abilities, PA,

and reading in children have been demonstrated in multiple studies (David, Wade-Woolley, Kirby, & Smithrim, 2007b; Flaugnacco, Lopez, Terribili, Zoia, Buda, Tilli, Monasta, Montico, Sila, Ronfani, et al., 2014; Holliman et al., 2010; Whalley & Hansen, 2006). For example, Moritz et al. (2012) found that the non-linguistic rhythmic abilities of kindergarteners (their ability to reproduce or discriminate short rhythmic patterns made with bongo drum sounds) predicted their phonological and reading abilities in 2<sup>nd</sup> grade, even when partial correlation was used to control for overall cognitive ability (see also David, Wade-Woolley, Kirby, & Smithrim, 2007). In other studies, non-linguistic rhythmic processing deficits in children with dyslexia were reported, with the severity of these deficits predicting variance in phonological and reading abilities (e.g., Flaugnacco, Lopez, Terribili, Zoia, Buda, Tilli, Monasta, Montico, Sila, Ronfani, et al., 2014; Goswami, Huss, Mead, Fosker, & Verney, 2013). Importantly, longitudinal studies demonstrated positive impact of rhythmic training on phonological processing and reading, both with typically - developing children (e.g., Rautenberg, 2015) and with struggling readers (e.g., Bhide, Power, & Goswami, 2013b; Flaugnacco, Lopez, Terribili, Zoia, Buda, Tilli, Monasta, Montico, Sila, Ronfani, et al., 2014). Yet not all studies have found such relations (Anvari, Trainor, Woodside, & Levy, 2002; Gordon et al., 2015). Anvari et al. (2002) for example, found that musical pitch processing (same/different melody and chord discrimination), but not rhythm (same different rhythm discrimination and vocally produced rhythm repetition), was correlated with early reading abilities in 5-year olds. Gordon et al., (2012) demonstrated that the association between PA and rhythm in 6-year olds was no longer significant after

non-verbal IQ was partialled out (Gordon et al., 2015). It is likely that heterogeneity in testing methods and sample characteristics accounts for some of these discrepancies (Swaminathan & Schellenberg, 2016).

As research on relations between non-linguistic rhythmic skills and reading abilities proceeds, one issue that merits attention (and which motivates the current study) is whether performance on *different types* of rhythmic patterns is differentially related to reading-related linguistic abilities. ‘Rhythm’ is not an undifferentiated category. In research on music cognition, one basic distinction is between rhythmic patterns that are metrical and nonmetrical (Povel & Essens, 1985). As noted by Povel and Essens (1985), a metrical sequence is one that “is mapped onto a frame formed of equal time intervals. For example, the interval sequence 22312114 (numbers indicating intervals between tone onsets in arbitrary time units), with a total duration of 16 time units, may be perceived metrically as having a metrical framework with intervals of 4 time units.” By this Povel and Essens (1985) mean that the above sequence of intervals can be subdivided into successive subgroups of equal duration (in the above example: [22][31][211][4], in which the numbers in each set of brackets sums to 4). In contrast, a nonmetrical sequence is “one that cannot be subdivided into equal time intervals, such as the interval sequence of 13214, with a total duration of 11 time units.” In this case, there is no way to subdivide intervals such that successive groups are of equal duration (e.g., the subdivision [13][21][4], the groups have durations of 4, 3, and 4).

The current study investigates whether the processing of metrical sequences has a different relationship to reading-related abilities than does the processing of nonmetrical sequences. This contrast is of interest because metrical sequences are characteristic of much music, in which rhythm often involves an underlying temporal grid of isochronous units (such as musical ‘measures’ or the time intervals between beats). In contrast nonmetrical sequences are arguably more characteristic of speech rhythm, which is temporally structured but does not involve an underlying grid of equal temporal units (Nolan & Jeon, 2014; Patel, 2008a, Ch. 3; Turk & Shattuck-Hufnagel, 2013). (Note that linguists often refer to speech as having a ‘metrical hierarchy’, but this refers to the organization of prominence at different levels of prosodic structure, such as the syllable, word, and phrase, not to an underlying isochronous temporal grid which organizes these units [See Patel, 2008a Ch. 3 for an extended discussion]. It is the ‘isochronous grid’ meaning of ‘metrical’ that we employ in the current work). Thus we hypothesized that processing abilities for metrical rhythms would contribute no unique variance to predicting PA and reading skills over and above processing abilities for nonmetrical rhythms.

To test this hypothesis, we measured kindergarteners’ ability to discriminate non-linguistic rhythmic patterns that were metrical or nonmetrical according to the above definition of these terms, and related performance on these tasks to measures of PA and reading while also measuring other cognitive abilities. We focused on kindergarteners because the investigation of rhythm-reading links at this age has significant theoretical and pragmatic interest. In the United States, it is at this age

that children are required to enter the school system and when initial reading instruction begins. From a theoretical perspective, demonstrating that links with rhythm exist early in reading development, before such links are confounded by instructional practices and musical training, points to the existence of shared cognitive and neural mechanisms between the two domains. From a translational perspective, if rhythm-reading links can be demonstrated in kindergarteners, this would be an ideal age to capitalize on such links by providing structured musical activities to boost phonological processing and reading acquisition. At this age, children engage with music with enthusiasm and the brain's microarchitecture is still developing rapidly (Giedd et al., 1999). Furthermore, reading interventions are most effective when begun in kindergarten and first grade (Hiebert, 2000). Thus, if rhythmic and reading abilities are linked at this age, rhythm-based interventions could potentially benefit both typically-developing children and provide an unobtrusive boost to phoneme-related skills in children who struggle with reading.

We tested a number of language abilities in kindergarteners, specifically chosen because they are known to be precursors to literacy. These measures include letter-sound knowledge, PA, verbal working memory, and rapid automatized naming (RAN) (Norton & Wolf, 2012; Scarborough, 1998; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004). Letter-sound knowledge measured in kindergarten is the most robust, but ephemeral predictor of reading outcomes (i.e., it loses its predictive accuracy beyond kindergarten) (McBride-Chang, 1999; Wagner & Barker, 1994). At the intersection of phonology and written language, this ability reflects the cumulative knowledge of letter names, the sounds they

make, and their visual representations. PA is the meta-understanding of the sound units of oral language, measured by the ability to identify and manipulate linguistic sounds independent of meaning. PA has a reciprocal relationship with reading development (e.g., the orthographic knowledge of ‘cat’ may enhance the segmentation of the word into sounds), and its role in reading changes through the years as reading becomes more automatic and less reliant on decoding abilities (Scarborough, Dobrich, & Hager, 1991). WM is a separate, but closely related construct to PA that measures the capacity to maintain and process information (e.g., digits, pseudowords) for a short period of time (Siegel & Linder, 1984; Stanovich et al., 1984). RAN is the ability to rapidly retrieve the name of visually presented familiar items in a serial array (e.g., objects, colors, numbers, or letters), in order to understand the automaticity with which visual information can be integrated with language processes (Denckla & Rudel, 1976; Wolf & Denckla, 2005). In addition to measures of letter-sound knowledge, PA, verbal working memory, and RAN, we also included a test of grammatical processing, motivated by a recent study which found that rhythm discrimination skills in 6-year-old children were associated with grammatical skills, but not PA (Gordon et al., 2015).

While there are standardized tests for the early literacy measures described above, tests of rhythmic ability vary widely between studies, and no prior research with kindergarteners has used rhythm patterns that vary in metrical structure. Thus there were two considerations in constructing the rhythm tasks: 1) to optimize child engagement, we developed a novel tablet-based rhythm discrimination game (described in Methods, Figure 1 shows a screen shot of the app); 2) to manipulate

metrical structure we drew on the research from the music cognition literature, in which temporal patterns (with no intensity or pitch variation) are created in systematic ways that either conform well or poorly to an underlying metrical structure per the definition of Povel and Essens (Grahn & Rowe, 2009; Povel & Essens, 1985).

### **4.3 Methods**

#### **4.3.1 *Participants and general procedures***

Kindergarten students ( $n = 74$ , 46.9% girls) from 5 schools in the Boston-area participated in this study (three Archdiocese Catholic schools, two private schools). All children whose parents provided consent were included in the study. The mean age of the sample was 69.63 months or 5.8 years ( $SD = 4.27$  months). Each school provided information on their percentage of students who qualify or receive free/reduced lunch as the proxy for the school-level socioeconomic status. Reduced lunch qualification at the five schools ranged from 0% to 40%, indicating a range of socioeconomic status across the schools.

All testing was completed in schools early in the spring of the academic year. Language and rhythm assessments were administered in one or two sessions depending on the testing schedule. Total testing time varied between 40 minutes to an hour. A team of research assistants administered all tests on a one-to-one basis with the children in a designated quiet area. Standard scores were calculated for all assessments using the test publisher age norms.



#### **4.3.2 *Language and cognitive measures***

The language and cognitive battery was administered by trained researchers in one 40-60 minute session. Standard administration and scoring rules were followed for all the tests. Four tests were administered, with the measure of early literacy (letter-sound knowledge) being embedded in the third (Phonological Awareness) Test:

**The comprehensive test of phonological processing (CTOPP;** Wagner et al., 1999). Two subtests were administered (1) Blending Words (2) Nonword Repetition (NWR).

**Rapid Automatized Naming/Rapid Alternating Stimulus (RAN/RAS;** Wolf & Denckla, 2005b). The Colors and Objects subtests were administered.

**Kaufman brief intelligence test (KIT-2)** (KBIT-2, Kaufman & Kaufman, 2004). The Matrices subtest was administered as a measure of nonverbal IQ.

**The phonological awareness test (PAT)** (PAT, Robertson & Salter, 1997). The following subtests were administered to measure phonological awareness skills: 1) Rhyming Discrimination: the child decides if two words rhyme; 2) Rhyming Production: the child produces a rhyme for a given word; 3) Segmentation of Sentences: the child claps hands for each word in the sentence as they repeat the sentence aloud; 4) Segmentation of Syllables: the child claps hands for each syllable in a word as they repeat it aloud; 5) Deletion of Sounds: Compounds/Syllables: the child repeats the word while deleting a given syllable(s); and 6) Graphemes: the child identifies the sounds that correspond to different

consonants, vowels, consonant blends, consonant digraphs, r-controlled vowels, vowel digraphs, and diphthongs. Although the Graphemes subtest was administered as part of the phonological battery of assessments, it is designed to measure letter-sound knowledge and represents an early literacy measure.

**The grammar and phonology screening (GAPS) test** (Van der Lely, Gardner, McClelland, & Froud, 2007). The Grammar subtest was administered to measure morpho-syntactic skills. In this test, the child repeats sentences that target aspects of morphological and syntactic structure known to be deficient in children with specific language impairment.

For all measures, raw scores were converted to standard scores using age-referenced norms from the publisher. The two CTOPP measures have a mean of 10 and a standard deviation of 3, and all other measures have a mean of 100 and a standard deviation of 15.

#### **4.3.3 *Rhythm perception measures***

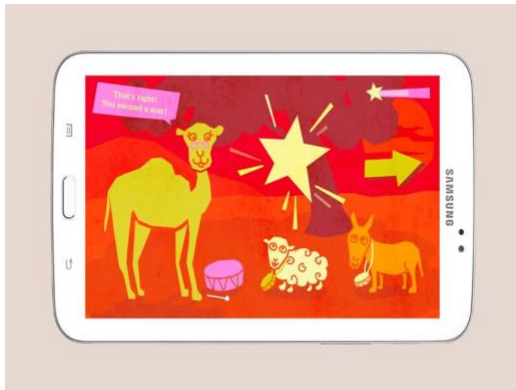
The tablet-based “Rhythm School” App was developed to measure musical rhythm discrimination skills in children (Figure 1). The app implements child-driven design principles and collects rhythm discrimination accuracy and response time data. In the app, children listen to two short rhythmic patterns (made from a cowbell sound), and decide if the two patterns are the same or different. The rhythms were administered on a Samsung Galaxy Tab 3 Lite (T110) tablet as part of a game in which children help animals learn how to drum. The procedures were first explained to the children by the researcher who assessed understanding by

illustrating same and different rhythm sequences by tapping on the desk. The children were then handed the tablet and they played the app following instructions by one of the characters (the Camel teacher). There were 22 trials in total, with the first 2 trials being practice trials that included performance feedback and repeated instructions.

On each trial, the child first heard the ‘teacher’s rhythm’ (played by the camel) twice, followed 2 seconds later by a second rhythm. The task was to decide if the second rhythm was played by ‘Sandy the sheep’ (who always imitates the teacher, reproducing exactly what the teacher played) or ‘Dudley the donkey’ (who always plays a rhythmic pattern different from the teacher). None of the characters depicted on the screen moved during a trial (i.e., the screen image was still), so that the same/different decision could only be based on auditory information. This task was modeled on a similar task reported in Wieland et al. (2015) which in turn drew from (Grahn & McAuley, 2009) in the design of rhythmic stimuli.

Our temporal patterns were metrical or nonmetrical according the definition of Povel and Essens (1985) (cf. the Introduction). These patterns are shown as sequences of intervals between tone onsets in Table 1, expressed as a multiple of 200 ms and expressed in musical notation in Supplemental Figure A. In metrical sequences it is always possible to subdivide the interval patterns into two groups of equal length (e.g., for pattern a: [1111][31], for pattern b: [112][211], etc.), thus satisfying Povel and Essens (1985) definition of a metrical sequence. In the nonmetrical sequences, it is not possible to subdivide the patterns in this way. Our metrical and nonmetrical patterns were adapted from patterns 1-15 and 21-35

respectively, in Povel and Esssen (1985), Table 2. The original sequences in their Table 2 were adapted for children by truncating the sequences to 4 to 6 intervals. Our sequences were made up of identical short cowbell sounds of equal amplitude and duration, pitched at approximately 494 Hz (B4). Due to truncation, the number of sounds per pattern varied from 5 to 7.



**Figure 4.1:** Screen shot of the rhythm school app

	<b>Metrical</b>	<b>Nonmetrical</b>
1)	1 1 1 1 3 1	1 1 1 1 2 1
2)	1 1 2 2 1 1	1 1 1 2 3
3)	2 1 1 2 1 1	1 1 3 1 2 1
4)	2 2 1 1 1 1	2 1 1 3
5)	3 1 2 2	2 3 1 1 1
6)	1 1 2 1 1 2	1 1 1 2 2
7)	2 1 1 1 2 1	1 2 1 1 2
8)	1 3 1 1 1 1	1 2 3 1 1
9)	1 3 2 1 1	2 1 1 1 2
10)	2 1 1 2 1 1 <sup>10</sup>	3 1 1 1 1
11)	1 1 2 1 3	1 1 1 2 1 1
12)	1 2 1 1 1 2	1 1 1 3 1
13)	1 2 1 2 1 1	1 2 1 1 1
14)	1 3 1 2 1	1 2 3 1 1
15)	3 1 1 2 1	2 3 1 1

<sup>10</sup> Due to truncation of the original Povel & Essen patterns, metrical patterns 3 and 10 are identical. These patterns were not paired in any of the trials.

**Table 4.1:** Temporal intervals between note onsets  
Note. The intervals are in multiples of 200 ms.

In our rhythm discrimination task, the 20 experimental trials were evenly divided so that 10 pairs consisted of metrical rhythms and 10 of nonmetrical rhythms (5 “same” pairs and 5 “different” pairs in each group). To create the 10 metrical pairs, 5 metrical rhythms were randomly chosen from the pool of 15 metrical patterns to create the 5 “same” pairs, and the remaining 10 metrical patterns were used to create the 5 “different” pairs. A similar procedure was followed for creating the 10 nonmetrical pairs. Metrical and nonmetrical pairs were presented in a different random order for each participant in a single block. After each pair of rhythms was presented, children were prompted to respond at their own pace, and the app proceeded to the next trial after the child responded. The testing phase was preceded by two example trials (adapted from Povel & Essens 1985) using rhythms not presented in the experimental trials.

## **4.4 Results**

### **4.4.1 *Exploratory factor analysis***

Performance on the literacy and rhythm measures across the entire sample is summarized in Table 2. Since the multiple measures of PA were correlated (see Supplemental Table A), we conducted an exploratory factor analysis with a promax rotation technique (Aad et al., 2010) on the six measures of PA (i.e., Rhyming Discrimination, Rhyming Production, Segmentation of Sentences, Segmentation of Syllables, Deletion of Sounds, and Blending Words). Cattell’s Scree Plot was used

to determine the adequate number of factors and 2 were selected. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.55, within the range considered suitable for factor analysis (Gorsuch, 1983; Hair, Anderson, Tatham, & Black, 1995), and Bartlett's test of sphericity was significant ( $\chi^2(10) = 24.35, p < 0.01$ ). Two factors were extracted and explained 100% of the total variance. The first factor was interpreted as rhyming (represented both of the rhyming measures accounting for 62% of the variance), the second factor was interpreted as phonemic manipulation (represented by the four measures: Segmentation of Sentences, Segmentation of Sounds, Deletion of Sounds, and Blending Words) accounting for 38% of the variance. The two factor scores were saved and included in the subsequent analyses.

#### **4.4.2 *Rhythm and literacy association***

There were significant relationships between rhythmic processing, phonological abilities, and the measure of early literacy (letter-sound knowledge, as measured by the Graphemes test). Standard scores on the measures of early literacy and phonology were positively correlated with discrimination sensitivity for metrical rhythms and, to a weaker extent, for nonmetrical rhythms (Table 3). Specifically, there was a significant association of Graphemes performance with metrical ( $r = .42, p < 0.001$ ) and marginally significant with nonmetrical ( $r = .22, p = 0.06$ ). Both metrical ( $r = .24, p = 0.05$ ) and nonmetrical ( $r = .31, p = 0.012$ ) rhythm performance were significantly associated with factor 2 (phonemic manipulation). Unexpectedly, metrical was also significantly associated with RAN

colors ( $r = .24, p = 0.036$ ). After Bonferroni correction for multiple comparisons was applied, only the metrical association with Graphemes remained significant.

To probe the effect of general cognitive abilities on the relationships among rhythm, early literacy, and phonological performance, the correlation analysis was repeated with non-verbal IQ and verbal working memory (NWR) partialled out. The partial correlation between RAN colors and metrical was no longer significant. The association of metrical rhythm performance with Graphemes remained significant ( $r = .39, p = 0.005$ ) and the correlation between Graphemes and nonmetrical performance became significant ( $r = .26, p = 0.036$ ). Only the association between Graphemes and metrical remained significant after correcting for multiple comparisons.

Consistent with prior work with children (D. McAuley personal communication, July 6, 2016), a paired t-test of the  $d'$  values for each condition revealed no significant differences discrimination performance on metrical vs. nonmetrical rhythms ( $t = .26, p = 0.81, d = -0.03$ ). Mean discrimination performance was low (metrical:  $M = .595, SD = .204$ ; nonmetrical:  $M = .59, SD = .18$ ), indicating that the task was challenging for the children. Despite this low mean value, there was a wide range in discrimination performance, and this variance did correlate with language measures (examples of  $d'$  values for the metrical and nonmetrical task are shown in Figure 2, regressed on the measure of early literacy, i.e., the Graphemes standard score.) Furthermore, one sample t-test revealed that both for metrical ( $t = 4.27, p < 0.001, CI = [.275, .756], d = 0.47$ ) and nonmetrical ( $t = 4.43, p < 0.001, CI = [.27, .7], d = 0.49$ ) rhythm discrimination, the population

mean was significantly different from zero. To evaluate whether the relationship between metrical and nonmetrical rhythm discrimination and the pre-literacy tasks is driven by children who performed below chance on the rhythm tasks, the bivariate correlation analyses were repeated excluding children who scored  $d' < 0$  on metrical or nonmetrical ( $n = 17$ ). The association between metrical and Graphemes ( $r = .32, p = 0.017$ ) and the association between nonmetrical and factor 2 ( $r = 0.2, p = 0.046$ ) remained significant in the remaining children ( $n = 57$ ).

In order to rule out that difference in rhythm performance related to socioeconomic status, rhythm performance was compared across the five schools, which ranged considerably in SES (cf. Methods). Differences in performance were not significant for metrical ( $F(4, 72) = .82, p = 0.51, \eta_p^2 = 0.006$ ) or nonmetrical ( $F(4,71) = 1.35, p = 0.26, \eta_p^2 = 0.005$ ) rhythms.

	Minimum	Maximum	Mean	Std. Deviation
<hr/> Kindergarten <hr/>				
Nonverbal IQ	83.00	138.00	99.81	11.74
Rhyming Discrimination	61.00	115.00	104.28	11.63
Rhyming Production	66.00	119.00	101.52	14.88
Segmentation of Sentences	55.00	117.00	104.72	13.38
Segmentation of Syllables	67.00	123.00	99.92	14.74
Deletion compounds/syllables	78.00	122.00	103.11	10.08
Graphemes	70.00	134.00	100.84	12.07
Non-Word Repetition	3.00	16.00	8.93	2.70
Blending Words	3.00	15.00	9.19	2.73
Grammar	3.00	100.00	71.24	36.00
RAN Objects	62.00	130.00	101.21	16.65
RAN Colors	55.00	135.00	100.14	16.51
Metrical	10.00	90.00	59.51	20.43

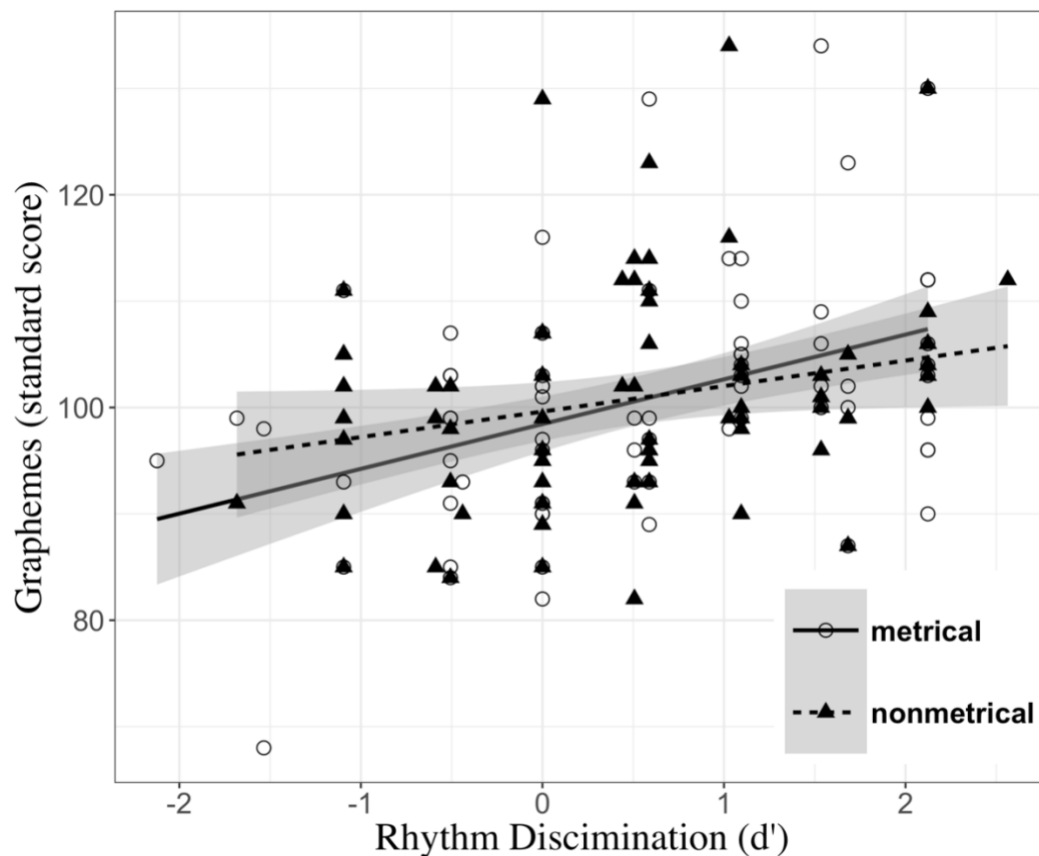


Nonmetrical	20.00	100	59.14	18.32
-------------	-------	-----	-------	-------

**Table 4.2:** Language (standard scores) and rhythm (percentage accuracy) performance. *Note.* Standard scores are reported for language and cognitive measures. Percent accuracy is reported for the rhythm measure

	Non-verbal IQ	Factor 1	Factor 2	Graphemes	Grammar	RAN Objects	RAN Colors
Bivariate Correlation							
SB	.186	-.056	0.242*	.417**	.075	.114	0.235*
WB	-.047	.182	0.307*	.220	.134	-.018	.025
KBIT, Gender, Age, and School Partialled Out							
SB	N/A	-.030	.116	.392**	.071	.078	.175
WB	N/A	.201	.202	.259*	.167	-.010	.030

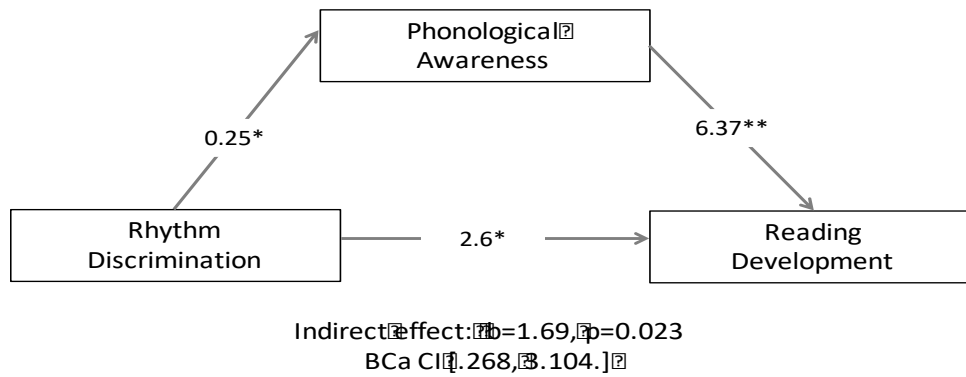
**Table 4.3:** Correlation coefficients between early literacy measures and rhythm variables \* $< 0.05$  \*\* $< 0.001$ , Bonferroni corrected for multiple comparisons



**Figure 4.2:** Scatterplots showing the relationship between the PAT Graphemes scores and the  $d'$  values representing performance on metrical (red) and nonmetrical (blue) patterns.

#### 4.4.3 *Mediation analysis*

To evaluate whether and to what extent the association between early literacy and rhythm discrimination is mediated through phonological awareness, a mediation analysis was conducted using the lavaan package in R (Rosseel, 2012). The package estimates several association effects: a) the direct effect of the independent variable on the dependent variable (i.e., Graphemes) independently of the mediator; b) the indirect effect of the independent variable on the dependent variable through the mediator (i.e., phonological awareness); and c) the total effect, configured as the sum of the direct and indirect effects (Hayes, 2012). A mean score of metrical and nonmetrical rhythm discrimination was computed and included as a predictor for the analysis. The mediator for this model was factor 2 (phonemic manipulation), which showed a significant association with rhythm performance. The model was fit using a robust maximum-likelihood criteria in the lavaan package (Rosseel, 2012) and bootstrapped standard error estimates were computed to account for potential deviation from multivariate normality and for the known normality problems when testing defined mediation coefficients. Overall the model provided a good fit to the data (Satorra–Bentler  $\chi^2(2) = 35.234$ ,  $p < 0.001$ ; CFI = 1.00; TLI = 1.00; RMSEA = 0) and Figure 3 displays the estimated coefficients. Results revealed significant direct ( $b = 2.6$ ,  $p = 0.024$ , BCa CI [.341, 4.866]) and indirect via phonological awareness factor 2 ( $b = 1.69$ ,  $p = 0.023$ , BCa CI [.268, 3.104]) effects of rhythm discrimination on Graphemes scores.



**Figure 4.3:** Rhythm-phonology-literacy mediation model

Effects of rhythm (mean  $d'$ ) on phonological awareness (the mediator-factor 2), direct effects on Reading (outcome-Graphemes), and indirect effects on Reading via phonological awareness.  $*p < 0.05$ ,  $**p < 0.001$ .

#### 4.4.4 Regression analysis

In order to assess the independent contribution of metrical (vs. nonmetrical and PA) patterns to the association between rhythm discrimination and literacy, a regression model was constructed with the Graphemes as the outcome measure (Table 4). Stepwise regression revealed that metrical discrimination accounted for a significant ( $p = 0.02$ ) and unique variance (6.16%) in Graphemes scores after IQ and NWR (1.74%,  $p = 0.018$ ), PA (factors 1 and 2: 12.13%,  $p < 0.001$ ), and nonmetrical scores (0.04%,  $p = 0.457$ ) were entered into the model. The overall model accounted for 40.83% of Graphemes performance.

	<b>B</b>	<b>SE B</b>	<b><math>\beta</math></b>	<b>t</b>	<b>p</b>
<b>Step 1</b>	<b><math>R^2=0.017, p=0.018</math></b>				
IQ	.041	.109	.116	.377	0.701
NWR	.585	.454	.119	1.289	0.203
<b>Step 2</b>	<b><math>R^2=0.121, p &lt; 0.001</math></b>				
PA factor 1	.519	1.164	.106	.446	.658
PA factor 2	5.601	1.745	.114	3.21	0.002
<b>Step 3</b>	<b><math>R^2=0.0004, p=0.457</math></b>				
nonmetrical	-0.24	1.17	.105	-.206	0.838
<b>Step 4</b>	<b><math>R^2=0.062, p=0.02</math></b>				
metrical	2.629	1.099	.109	2.393	0.02

**Table 4.4:** Hierarchical regression predicting Graphemes scores

## 4.5 Discussion

The current study examined relations among non-linguistic rhythmic processing, phonological awareness (PA), and reading abilities in early-reading kindergarten children. The study builds on a growing body of research suggesting that certain basic temporal processing mechanisms are shared by speech and non-linguistic rhythmic processing, mechanisms relevant to phonological processing and hence to the acquisition of reading skills (Kraus & White-Schwoch, 2016). A distinctive aspect of this study was the exploration of whether processing of *different types* of rhythmic patterns is differentially related to kindergartener's reading-related linguistic abilities. Specifically, we compared the processing metrical vs. nonmetrical patterns, i.e., patterns capable of being subdivided into equal temporal intervals or not (Povel & Essens, 1985). This contrast is of interest because metrical sequences are characteristic of much music, in which rhythm often involves an underlying temporal grid of isochronous units (such as musical

‘measures’ or time intervals between beats). In contrast nonmetrical sequences are arguably more characteristic of speech rhythm, which is temporally structured but does *not* involve an underlying grid of equal temporal units (Nolan & Jeon, 2014; Patel, 2008a, Ch. 3; Turk & Shattuck-Hufnagel, 2013). Thus we predicted that discrimination performance on both types of rhythms would be associated with PA, but that metrical rhythms would not predict unique variance in phonological and reading abilities over and above performance on nonmetrical rhythms. Consistent with our hypothesis, PA was associated with both metrical and nonmetrical rhythm discrimination. However, contrary to our expectation, metrical processing uniquely predicted letter-sound knowledge above general intelligence, auditory working memory, phonological awareness, and nonmetrical processing.

#### **4.5.1 *Relationship between rhythm and phonological awareness***

As expected based on previous studies (e.g., Holliman, Wood, & Sheehy, 2010; Moritz et al., 2012), rhythm was significantly associated with PA skills. Specifically, when performance on PA skills was broken down into two distinct factors (using factor analysis), both metrical and nonmetrical rhythm processing was associated with the phonemic manipulation factor. These results are congruent with our initial hypothesis that since the flow of speech in English is characterized by varied time intervals, processing of metrical patterns should not be uniquely associated with PA.

Importantly, PA partially mediated the association between rhythm processing and reading. This indirect relationship supports the tentative pathway proposed to explain the association between musical rhythm and reading through

speech perception. Both speech and non-linguistic rhythm processing require accurate processing of the temporal structure in acoustic stimuli. Temporal patterns in speech contain important cues to phonological units, such as phonemes, syllables, and stresses. These collective cues help language learners (including infants) segment syllables and words from the acoustic stream and develop a robust phonological template (Kuhl, 2004). Phonological awareness is a foundational skill for reading acquisition as learning to read depends on mapping visual letter representations to their acoustic counterparts (Bruck, 1992; Ramus, 2003; Stanovich, 1986). Consequently, performance on phonological awareness tasks is a strong predictor of later reading success or impairment (Morris, Stuebing, et al., 1998; Ozernov-Palchik, Norton, et al., 2016).

We found that rhythmic abilities were not associated with the rhyming factor. This suggests that rhyming activities, which often occur as part of early reading instruction (Phillips, Clancy-Menchetti, & Lonigan, 2008), do not drive our observed relationship between rhythmic skills and early reading abilities. Additionally, there was no significant difference in rhythm or literacy performance across schools, further suggesting that the observed association between the two skills is not due to environmental variables such as home literacy and musical training. The lack of association between the rhyming factor and rhythm skills in the current study is supported by a previous investigation in young children that found that rhythm perception and production were associated with Segmentation and Deletion subtests of the PAT, but not with Rhyming (Moritz et al., 2012). Furthermore, rhyming does not require explicit phonological awareness and is

acquired early in development, and therefore is not considered a strong predictor of reading outcomes beyond the preschool years (Muter, Hulme, Snowling, & Stevenson, 2004; Nation, 1997).

#### **4.5.2 *Relationship between metrical processing and early literacy***

Metrical, but not nonmetrical processing was significantly associated with early literacy (i.e., letter-sound knowledge as measured by the Graphemes subtest). Furthermore, contrary to our expectation, metrical processing uniquely predicted early literacy abilities above general intelligence, auditory working memory, phonological awareness, and nonmetrical processing. Given that speech does not use a temporal grid based on equal-duration intervals, how can this finding be understood? Our metrical rhythms had an underlying temporal structure that our nonmetrical patterns did not, i.e., a structure based on equal-duration units. It is possible that children who were better at detecting these underlying regularities in metrical sequences may also be the same children who are better at picking up statistical regularities in speech (e.g., phonotactic, syllabic, or prosodic patterns), which would, in turn, aid in their segmenting the speech stream into its underlying phonological units (Schön & Tillmann, 2015). Indeed, it has been suggested that domain-general mechanisms, key for language specialization, guide the acquisition of Western cultural preference for metrical rhythms in infancy (Trehub & Hannon, 2006). There is emerging evidence for positive links between statistical learning, in both visual and auditory modalities, and reading skill, suggesting that processing of implicit structure is an important mechanism for literacy development (Apfelbaum, Hazeltine, & McMurray, 2013; Arciuli & Simpson, 2012; Spencer,

Kaschak, Jones, & Lonigan, 2015). Additionally, studies of individuals with dyslexia report cross-domain deficits in tasks requiring perceptual learning of and adaptation to stimulus regularity (Ahissar et al., 2006; Gabay et al., 2015; Perrachione et al., 2017). Thus, it is possible that sensitivity to stimulus regularities could explain our findings of a link between metrical rhythmic skills and early literacy, particularly since performance on metrical rhythms accounted for unique variance in early literacy above auditory measures such as working memory and phonological skills.

#### **4.5.3 *Considerations and implications***

Unlike prior studies that demonstrated a connection between rhythmic processing and grammar (Gordon et al., 2015) or RAN (David, Wade-Woolley, et al., 2007), our results showed a unique association between rhythmic and phonological skills. Additionally, while some have suggested that auditory working memory could be an important link between phonological awareness and music perception (Jaffe-Dax, Raviv, Jacoby, Loewenstein, & Ahissar, 2015; Peynircioglu, Durgunoglu, & Üney-Küsefog˘lu, 2002), even after controlling for auditory working memory (Nonword Repetition) and general cognitive abilities (IQ) the significance of the association between rhythm and phonology remained. Furthermore, since there were differences in the number of events between some of the metrical and nonmetrical items (metrical patterns had more events on average), the lack of correlation between metrical or nonmetrical performance with Nonword Repetition suggest that our results cannot be attributed to differences in working memory load posed by the two tasks.



It is possible that differences between our study and studies that found correlations between rhythm and other cognitive and literacy abilities are due to differences in methodology or subject characteristics (David, Wade-Woolley, Kirby, & Smithrim, 2007; Gordon et al., 2015). Our study differed from past studies of childhood relationships between rhythmic and linguistic abilities in a number of respects: larger sample size and the types of measures administered (e.g., the Structured Photographic Expressive Language Test was used in Gordon et al., 2014 to evaluate grammatical abilities). Furthermore, the age range in the current study was intentionally restricted to kindergarten students, while other studies included children from more advanced grades and reading abilities.

Somewhat surprisingly, there was a lack of a performance advantage in discrimination of metrical vs. nonmetrical patterns in the current study (Grahn & Rowe, 2009). While such an advantage has been found in research with adults, our results are consistent with prior work with young children (D. McAuley personal communication, July 6, 2016). One factor that may have promoted comparable performance on metrical vs. nonmetrical rhythms in our study is that (unlike previous work) we did not change the underlying tempo of the patterns from trial to trial. This may have encouraged the perception of a beat in the nonmetrical patterns, which were interspersed among the metrical patterns in the discrimination test, thus weakening the perceptual contrast between metrical and nonmetrical patterns. Nevertheless, we did find differences in how metrical and nonmetrical pattern discrimination related to linguistic abilities, suggesting that the patterns were not processed in an identical way. We suspect that these differences in

processing reflect the greater degree of temporal regularity in our metrical patterns.

Our findings add to an increasing body of literature regarding the cognitive and neural overlap between music and language (e.g., Bhide et al., 2013b; Gordon et al., 2015; Huss et al., 2011; Moritz et al., 2012; Tierney & Kraus, 2013c), and suggest that it is worth investigating the impact of rhythmic training on reading development in young children. It is important to note that due to the correlational nature of our study, causality or direction of the relationship between rhythm and literacy could not be established. While there is some evidence from behavioral and neuroimaging studies that rhythm training improves skills related to literacy (Flaugnacco, Lopez, Terribili, Zoia, Buda, Tilli, Monasta, Montico, Sila, Ronfani, et al., 2014; Kraus et al., 2014; Moreno, Friesen, & Bialystok, 2011; Zhao & Kuhl, 2016), larger longitudinal studies of children with follow-up are greatly needed (Jaschke, Honing, & Scherder, 2018). An intriguing question for such studies concerns the possible bi-directional effects of training, i.e., whether interventions targeting phonological awareness are associated with improved non-linguistic rhythm processing abilities.

## **5 Study 4: The Relationship Between Socioeconomic Status and White Matter Coherence in Pre-Reading Children: A Longitudinal Investigation**

### **5.1 Abstract**

Reading is a learned skill crucial for educational attainment. Children from families of lower socioeconomic status (SES) tend to have poorer reading outcomes and this gap widens across years of schooling. Reading relies on the orchestration of multiple neural systems integrated via specific white-matter pathways, but there is limited understanding about whether these pathways relate differentially to reading performance depending on SES background. Kindergarten white-matter coherence and second grade reading outcomes were investigated in an SES-diverse sample of 121 children that was divided into higher-SES ( $n = 61$ ) and lower-SES ( $n = 60$ ) groups. The three left-hemispheric white-matter tracts most associated with reading were examined: arcuate fasciculus (AF), superior longitudinal fasciculus (SLF), and inferior longitudinal fasciculus (ILF). Children from lower-SES families had significantly reduced fractional anisotropy (FA) in the occipito-temporal segment of the left ILF in kindergarten. In lower-SES children, but not in

higher-SES children, higher FA in this segment in kindergarten was associated with better second-grade reading outcomes. Random forests classification revealed that the parental reading history, IQ, home literacy environment, and FA in the right SLF discriminated with 78% accuracy between lower-SES children who developed into good versus poor readers in second grade. These results have implications for understanding the role of the environment in the development of the neural pathways that support reading, and the possible neural mechanisms of successful reading development in children from lower-SES backgrounds.

## **5.2 Introduction**

Reading is a learned skill crucial for successful educational attainment. There are well-documented socioeconomic status (SES) disparities in reading achievement across development (Peterson & Pennington, 2015; Reardon, 2011). Neural specialization for reading is experientially driven and occurs through utilizing and repurposing brain structures that support functions such as vision, audition, and language (Dehaene, 2004). The efficient integration across these spatially disparate brain regions is made possible by long-range white matter connections that form across development (Wandell, Rauschecker, & Yeatman, 2012). Three white matter tracts have been consistently linked to reading: (1) the left arcuate fasciculus (AF), connecting the superior temporal lobe with the inferior frontal gyrus (IFG); (2) the left superior longitudinal fasciculus (SLF), connecting the inferior parietal with the inferior frontal/premotor regions; and (3) the left inferior longitudinal fasciculus (ILF), connecting the posterior inferior temporal gyrus with the ventral anterior and medial temporal lobe (Myers, 2014; Saygin et

al., 2013; Vandermosten, Boets, Wouters, et al., 2012; Wang et al., 2016). The left ILF passes in close proximity to the visual word form area (Yeatman et al., 2013), a cortical region involved in word recognition, whose specialization for reading is experientially driven (Cohen et al., 2002; Dehaene & Cohen, 2011). The current study investigated, for the first time, whether there is a relationship between SES and these three white-matter pathways in pre-reading children; whether SES modulates the links between these pathways and longitudinal reading outcomes; and which familial, behavioral, and white-matter factors distinguish between good and poor reading outcomes in lower-SES children.

A confluence of genetic and environmental influences interact reciprocally to affect children's reading development (Ozernov-Palchik, Yu, et al., 2016). Hereditary risk is a strong predictor of reading disability, as approximately 40-60% of children who have a parent who is reading-disabled will have reading problems themselves (Gilger, Hanebuth, Smith, & Pennington, 1996; Snowling, Gallagher, & Frith, 2003). Environmental factors are also significant in determining reading skill and explain up to 30% of individual differences in reading, with parental SES exerting the most influence (Olson, Keenan, Byrne, & Samuelsson, 2014; Petrill, Deater-Deckard, Thompson, De Thorne, & Schatschneider, 2006; Taylor, Roehrig, Hensler, Connor, & Schatschneider, 2010). SES is a multidimensional construct that encompasses parental education levels, economic resources such as income, and social status (Tomalski & Johnson, 2010). SES is also a proxy for quantifying the quality of the prenatal and postnatal environment to which a child is exposed. Parental education and occupation are considered to be stable indicators of SES and

are correlated with parental involvement in children's educational attainment (Bradley & Corwyn, 2002; Friend et al., 2008).

Gaps in reading achievement between lower- and higher-SES children are one of the most consistent findings in the educational literature (Reardon, 2011). Lower SES has been associated with worse performance in vocabulary, phonological awareness, single word decoding, reading comprehension, and grammar (Bowey, 1995; Noble & McCandliss, 2005). Children from lower-SES families are 2.5 times more likely to read below grade level and more likely to meet the criteria for reading disability than children from higher-SES backgrounds (Peterson & Pennington, 2015). Crucially, these gaps in reading achievement begin even before children enter school (Coley, 2002; Reardon & Portilla, 2016), widen across the years of schooling (Feinstein, 2003), and have been widening over the past decades (Brooks-Gunn & Duncan, 1997; Reardon & Portilla, 2016; Reardon, 2011). This underscores the need to understand the early impact of SES on reading development and its neurobiology in order to prevent and remediate the spiraling effects of SES on educational attainment.

SES in childhood has profound implications for brain development, with language and reading-related brain structures being particularly affected (Noble & McCandliss, 2005; Noble, Tottenham, & Casey, 2005; Romeo et al., 2017; Rowe & Goldin-Meadow, 2009). SES is thought to affect brain development, and subsequent behavioral outcomes, through a range of mediating factors such as maternal stress and cognitive stimulation (Hackman et al., 2010). Animal studies have demonstrated the effects of these factors on mechanisms underlying neural

development and plasticity such as dendritic branching, gliogenesis, synaptogenesis, neurogenesis, and the integration of newly generated neurons into functional circuits (for a review see Hackman et al., 2010).

Indeed, the association between SES and brain structure and function has been demonstrated in multiple ways (e.g., Betancourt et al., 2015; D'Angiulli, Herdman, Stapells, & Hertzman, 2008; Hackman, 2009; Hanson et al., 2013; Jednoróg et al., 2012; Luby et al., 2013; Mackey et al., 2015; Noble et al., 2015; Noble, Houston, Kan, & Sowell, 2012; but see Lange et al., 2010; Eckert et al., 2001; Brain Development Cooperative Group, 2012; Raizada, Richards, Meltzoff, & Kuhl, 2008). Specifically, lower SES has been linked to reduced grey matter volume (Hanson et al., 2013; Hanson et al., 2011; Jednoróg et al., 2012; Luby et al., 2013), reduced cortical thickness (Mackey et al., 2015), reduced degree of cortical gyrification (Jednoróg et al., 2012), and reduced surface area (Natalie & Noble, 2014) in occipito-temporal, temporo-parietal, and inferior frontal regions that support reading development (Booth, 2001; Martin, 2015; Ozernov-Palchik, 2016). Functional MRI and electrophysiological studies have reported decreased specialization for reading and language in task-relevant regions in children from lower-SES backgrounds (see a review by Pavlakis, Noble, Pavlakis, Ali, & Frank, 2015).

Few studies to date have investigated the association between structural or white-matter connectivity and SES (Chiang et al., 2011; Noble, Korgaonkar, Grieve, & Brickman, 2013; Piras, Cherubini, Caltagirone, & Spalletta, 2011), with even fewer studies in children (Chiang et al., 2011; Gullick, Demir-Lira, & Booth,

2016; Jednoróg et al., 2012). The results of these investigations in school-age children are mixed, with some studies finding SES-related differences in white matter pathways including left SLF and ILF (Dufford & Kim, 2017; Gullick, Demir-Lira, & Booth, 2016; Ursache & Noble, 2016), but others finding no association between white matter and SES (Chiang et al., 2011; Jednoróg et al., 2012). Only one study examining the relations among SES, white matter, and reading ability (Gullick et al., 2016). In that study of 42 children ages 8-14, higher SES was associated with higher FA in several white matter tracts involved in reading, and SES modulated the association between tract measures and reading ability, such that brain-reading ability links were stronger in lower-SES children. Interpretation of that study is difficult, however, because the commonly observed relation between SES and reading ability (e.g., Reardon, 2011) was not observed in this study in which reading ability was unrelated to SES.

The present study differs from prior research relating SES, reading, and the brain in two fundamental ways. First, prior studies examined children after multiple years of formal education, which confounds factors related early home environment and later school quality. We examined children before or early in kindergarten. Second, we examined variation in reading outcomes among children from lower SES environments longitudinally following these children to determine reading ability at the end of second grade. In terms of neuroimaging, findings from studies of individuals with dyslexia have demonstrated the importance of right-hemispheric networks, specifically prefrontal regions and right SLF, for improved reading outcomes (Constable et al., 1998; Hoeft, 2011; Milne, Syngienotis,



Jackson, & Corballis, 2002; Powers et al., 2016). In these studies, however, SES was not characterized, and most likely represented children from higher-SES environments whose families are most likely to participate in research. It is unknown, therefore, whether such right-hemisphere characteristics support better reading outcomes among children from lower SES environments.

The current study examined, for the first time, the interaction between white matter coherence, SES, and longitudinal reading outcomes in a large sample of children from pre-reading to reading age. First, we compared pre-literacy and literacy performance in kindergarten children from lower- and higher-SES families. We asked whether at-risk kindergarten children have better reading outcomes if they are in the lower-SES group compared to the higher-SES group. Second, we asked whether there were significant SES-related differences in fractional anisotropy (FA, a proxy for white matter tract coherence; Lebel, Treit, & Beaulieu, 2017) in the left AF, ILF, and SLF tracts important for reading. Second, we investigated whether SES modulates the longitudinal association between kindergarten white matter FA and second grade reading outcomes. To probe the direction of the modulatory effect, we investigated how kindergarten white matter predicted second grade reading separately in lower-SES and higher-SES groups. Third, we divided the lower-SES group based on second grade reading outcomes and investigated whether cognitive (e.g., vocabulary), parental (e.g., home literacy environment), and white matter factors could differentiate the lower-SES children who became typical readers from those who became poor readers. We included the

right SLF in this analysis due to its established role in compensation for poor reading.

Based on prior findings, we hypothesized that children from lower-SES families would demonstrate significantly poorer pre-reading and reading skills as compared to children from higher-SES families. We hypothesized that there would be reduced coherence in one or more of the three left-hemispheric tracts associated with reading. Based on previous findings, we expected modulatory effects on the FA-reading relationship with a stronger association between white matter FA and reading outcomes in lower-SES children. Finally, we predicted that some or all of the following factors would contribute to second grade outcomes in the low-SES group: IQ, vocabulary, home literacy environment, and parental reading history. Due to the scarcity of neuroimaging research on brain differences in lower-SES children in relation to reading outcomes, we relied on the reading disability literature and predicted that improved reading outcomes would be associated with higher (more typical) left-hemispheric FA or the recruitment of the right SLF.

## **5.3 Methods**

### **5.3.1 Overview**

Children from 20 diverse schools in New England completed a short battery of pre-reading assessments administered by trained researchers in their schools, administered individually by trained researchers (for details see Ozernov-Palchik, Norton, et al., 2016). These assessments took place in the spring of pre-kindergarten or fall of kindergarten. A subset of children were contacted and screened for eligibility, and eligible children participated in a follow-up neuroimaging study and

additional behavioral assessments. Because the study focused on risk for reading difficulty, children invited for neuroimaging over-represented children at apparent higher risk due to a family history of reading difficulty or low scores on pre-reading assessments. Children who participated in the neuroimaging session came in for behavioral follow-up sessions at the ends of first and second grade. This study was approved by the institutional review boards at the Massachusetts Institute of Technology and Boston Children's Hospital. Parents gave written consent and children gave verbal assent to participate.

### **5.3.2 *Participants***

Children with complete MRI diffusion and behavioral data ( $n = 129$ , 53% female, mean age = 66.98 months,  $SD = 4.16$ , range 58-80,) were included in the current analysis. Four participants were excluded due to poor quality diffusion data. SES was determined by parent report in kindergarten using the Barratt questionnaire (see details below). A median split was performed on SES scores, and of the 125 participants and four participants with exactly the median score were excluded, yielding a final  $n$  of 121 (higher-SES = 61, lower-SES = 60). The longitudinal analysis included 114 children who had longitudinal reading data available.

Inclusion criteria, based on parental report, were normal hearing, no neurological or psychiatric disorders, American English as a native language, and a full-term birth (>36 weeks). All children had KBIT-2 Matrices (nonverbal IQ) subtest standard scores above 80. Parents indicated which ethnic and racial category they identified their child with. The higher-SES group reported the following racial

and ethnic identities: 8% African American, 86% White, 2% Asian, 2% American Indian or Alaska Native, 2% did not report race; 94% not Hispanic, 6% Hispanic. The lower-SES group reported the following racial and ethnic identities: 31% African American, 2% Asian, 54% White, 3% American Indian or Alaska Native, 2% multiple races, 8% did not report race; 76% not Hispanic, 24% Hispanic. The lower-SES group contained a larger percentage of ethnic and racial minorities than the higher-SES group, mirroring demographic distributions in the United States (He, Goodkind, & Kowal, 2016).

### **5.3.3 Kindergarten behavioral measures**

All participants completed a comprehensive psychometric battery assessing cognitive and language skills in kindergarten/pre-k (see Ozernov-Palchik, Norton, et al., 2016). Based on the extensive literature on the importance of these pre-literacy skills for reading development, we focused on letter knowledge, phonological awareness, rapid naming, and vocabulary (e.g., Ozernov-Palchik, Norton, et al., 2016; Scarborough, 1998; Schatschneider et al., 2004b). Nonverbal IQ was further included to rule out differences due to general cognitive abilities.

**Phonological awareness (PA).** The Elision, Nonword Repetition, and Blending Words subtests were administered from the Comprehensive Test of Phonological Processing (CTOPP, Wagner et al., 1999). The mean of the standard scores of the subtests was used as the PA composite score.

**Rapid automatized naming (RAN).** The Colors and Objects subtests of the Rapid Automatized Naming/Rapid Alternating Stimulus (RAN/RAS) tests

(Wolf & Denckla, 2005b) were administered. The mean of Colors and Objects standard scores was used as the RAN composite score.

**Letter sound knowledge (LSK).** The Letter Sound Knowledge subtest from the York Assessment of Reading for Comprehension (YARC, Stothard et al., 2010) was administered.

**Nonverbal IQ (IQ).** The Matrices subtest from the Kaufman Brief Intelligence Test, Second Edition (KBIT-2, Kaufman & Kaufman, 2004b) was administered.

**Vocabulary.** The Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4, Dunn, 2007) assesses vocabulary knowledge.

#### **5.3.4 2nd Grade behavioral measures**

An extensive battery of assessments including phonological awareness, RAN, cognition, language comprehension, vocabulary, and reading measures was administered at the end of second grade. For the current study, the following measures were used to characterize children's second-grade language and reading outcomes:

**Untimed Single Real and Nonword Word Reading.** The Word Identification (WID) and Word Attack (WA) subtests of the Woodcock Reading Mastery Tests (WRMT-III, Woodcock, 2011). The child is asked to read aloud as many single words or nonwords as possible of increasing difficulty.

**Timed Single Real and Nonword Word Reading.** The Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subtests of the Test of Word Reading Efficiency (TOWRE, Torgesen, Wagner, & Rashotte, 1999).

### **5.3.5 *Measures of reading outcomes***

To first investigate whether SES modulates the relationship between white matter and reading ability, a linear regression analysis with second grade reading was conducted. For this continuous analysis, a composite word-reading score (WR) of all four second-grade measures was created. To investigate possible mechanisms of compensation, factors that discriminate between lower-SES children who became good readers from those who became poor readers in second grade were explored. For this categorical analysis, the two groups of readers were defined as follows. Children who had standard scores below 90 (below the 25th percentile for age) on at least two of the four subtests (WID, WA, SWE, or PDE) were categorized as poor readers. Children who scored 90 or above on all four subtests were categorized as typical readers.

### **5.3.6 *Socioeconomic status***

The Barratt Simplified Measure of Social Status (BSMSS) Questionnaire (Barratt, 2006) was completed by one of the child's parents. The SES score (SES) was derived from maternal and paternal years of education (ranging from a score of 3 for less than 7th grade education to a score of 21 for graduate degree) and an occupation prestige score for each parent (ranging from 5 (e.g., manual labor) to 45 (e.g., higher executive)). The overall SES score was then calculated as ((mean of

parent education) + (mean of parent occupation)), ranging from 8 to 66, with higher scores reflecting higher SES.

### **5.3.7 *Home literacy environment***

Information about the home literacy environment was collected via a parent report questionnaire. The questionnaire was adapted from Senechal (1997) and included the following questions that have been shown to directly contribute to the acquisition of early reading skills (Powers et al., 2016; Sénéchal, LeFevre, Hudson, & Lawson, 1996; Senechal, 2002; Senechal, 1997): number of children's books at home, age of the child when first read to, frequency of the child being read to, frequency of the child looking at books, and direct instruction of writing and the alphabet. A composite score (HLE) of the six questions was created for the analysis.

### **5.3.8 *Parental history of reading difficulties***

Parental history of reading difficulties was evaluated using the Adult Reading History Questionnaire (ARHQ, Lefly & Pennington, 2000). The questionnaire is scored by summing the responses to all questions and dividing by the total number of questions. Greater scores on the ARHQ indicate more reading impairment during childhood. For the current analysis, a composite of maternal and paternal ARHQ scores was created. ARHQ scores were missing for one of the parents for a subset of the participants ( $n = 39$ ).

### **5.3.9 *Imaging procedures***

As described by Raschle and colleagues (Raschle, Zuk, Ortiz-Mantilla, et al., 2012), children practiced in a mock scanner area at the beginning of each MRI session, with child-friendly equipment (e.g., pediatric headphones, head padding,

etc.) and procedures (e.g., strategies for motion reduction) that were also used during actual data acquisition. MRI sequences were acquired on a Siemens 3T Trio whole-body MRI scanner at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT, using a standard 32-channel head coil. The T1-weighted MPRAGE scan used the following specifications: 176 slices, TR= 2350 ms; TE= 1.64 ms; flip angle=9°; FOV= 256 mm; voxel size 1.0 x 1.0 x 1.0 mm. As in (Saygin et al., 2013), an online prospective motion correction algorithm was implemented to reduce the effect of motion artifacts during the structural scan, and 10 selective reacquisition time points were acquired and included to replace time points that were affected by head motion (Tisdall, 2012). The diffusion-weighted (DW) MRI scan included 10 non-diffusion-weighted volumes ( $b=0$ ) and 30 diffusion-weighted volumes acquired with non-colinear gradient directions ( $b=700 \text{ s/mm}^2$ ), all at 128x128 base resolution and isotropic voxel resolution of  $2.0 \text{ mm}^3$ . Scans were evaluated for motion and scanner-induced outliers using DTI prep software (Lui, Hansen, & Kriegstein, 2011). Motion parameters were found by rigidly registering the interleaved subvolumes. The translation threshold was set to 2.0 mm and the rotation threshold to  $0.5^\circ$ . From the original sample of 129 participants, four participants with 10 or more motion outliers were excluded from the analysis. To examine the possibility that the results were biased by differences in motion between the two SES groups, an independent samples  $t$ -test was used to compare the number of motion outliers in the lower- and higher-SES groups; this analysis revealed no significant differences between the two groups ( $p > 0.43$ ).



### 5.3.10 Identification of key white matter tracts

The Automatic Fiber Quantification (AFQ; <http://github.com/jyeatman/AFQ>) software package (Yeatman, 2012b) was used to identify each of the white matter tracts of interest chosen for their documented role in reading development (i.e., left ILF, SLF, and AF, as well as right SLF) and quantify the diffusion parameters along the tract. The AFQ pipeline includes the following steps (for an overview see Yeatman, 2012b): 1) fiber tracts are estimated using a deterministic streamlines tracking algorithm (STT) (Basser, Pajevic, Pierpaoli, Duda, & Aldroubi, 2000; Mori, Crain, Chacko, & Van Zijl, 1999) with an FA threshold of 0.2 and angle threshold of 40°; 2) fiber tracts are segmented using a region of interest (ROI) approach and fiber tract probability maps; 3) fiber groups are cleaned into a compact bundle using an iterative statistical outlier rejection algorithm; and 4) diffusion characteristics are calculated at each node, or spatial location, along the trajectory of the fiber. Each fiber was sampled at 100 equidistant nodes that can be used to compute the FA value at each node along the fiber. AFQ segments the whole-brain fiber group into 20 white matter tracts that are defined in the white matter atlas (Wakana et al., 2007). Diffusion parameters were computed using a weighted sum of each fiber's value at a given node where a fiber is weighted based on its Mahalanobis distance from the core or mean location of the tract (Johnson et al., 2013). This improves detection power for group differences. For each tract, the 100 nodes along the tract were resampled to 50 nodes by discarding the most anterior and posterior portions of fiber tract, as this is where individual fibers separate from the core fascicle toward their destination in the cortex. This approach normalizes the fiber endpoints across participants and

improves the co-registration of each fiber tract among all participants (Yeatman et al., 2011).

Our analyses focused on three left-hemispheric tracts that have been linked to reading ability in young children: AF, SLF, and ILF (Ben-Shachar, Dougherty, & Wandell, 2007; Saygin et al., 2013; Wang et al., 2016; Yeatman et al., 2011; Yeatman, 2012a). The right SLF was also examined due to evidence of its compensatory involvement in the reading development of at-risk and poor readers (Barquero, Davis, & Cutting, 2014; Hoeft, 2011a; Martin, 2015; Wang et al., 2016). Consistent with prior studies, fractional anisotropy (FA) was used as the primary parameter of interest to estimate microstructural integrity, but axial diffusivity (AD) and radial diffusivity (RD) were also computed and used in follow-up analyses of SES differences. FA is a summative measure of the three diffusion directions and indicates microstructural integrity of the white matter (Beaulieu, 2002; Pierpaoli & Basser, 1996), whereas AD and RD can inform understanding of the mechanisms of FA differences; AD has been related to changes in axon integrity and mechanisms of axon degeneration and RD is thought to reflect axonal myelination and density (Song et al., 2003; Song et al., 2005; Tyszka, Readhead, Bearer, Pautler, & Jacobs, 2006; Zhang et al., 2009).

## **5.4 Statistical analyses**

All analyses were executed in the statistical package R (Ihaka & Gentleman, 1996).

### **5.4.1 *Pre-literacy and literacy performance by SES***

To determine whether our study was consistent with the behavioral literature on performance gaps between lower- and higher-SES children, we

conducted a series of t-tests controlling for gender and IQ on kindergarten pre-literacy as well as second-grade literacy skills. False discovery rate (FDR) correction was used to adjust for multiple comparisons among these tests (Benjamini, Drai, Elmer, Kafkafi, & Golani, 2001).

#### **5.4.2 *Differences between SES groups in tract FA***

To test whether there were SES-based differences in FA between the two SES groups in the three tracts, a two-sample t-test was conducted, controlling for gender and IQ, with each of the nodes in left ILF, SLF, and AF. The Shapiro-Wilk test of normality was performed on all FA values to determine whether to use a two-sample t-test or the Mann-Whitney U-test for each paired comparison. Due to the high degree of correlation among the nodes on a specific tract, the traditional Bonferroni method is overly conservative and could lead to type 2 errors. Instead, permutation-based multiple correction (the AFQ\_MultiCompCorrection function based on Nichols & Holmes, 2001) was applied in Matlab (The Mathworks Inc., 2007) to determine the appropriate  $p$ -value. Significance was set at  $p = 0.05$  for all analyses and  $p = 0.05$  for cluster-based permutation corrections.

#### **5.4.3 *Differences between SES groups in AD and RD***

To reveal possible mechanisms underlying SES-related differences in FA, a series of t-tests was conducted to compare RD and AD between the two SES groups. Following the procedures described above, group comparisons were executed within tracts and nodes showing significant differences in FA.

#### **5.4.4 *Modulating effects of SES on FA/risk-reading relationship (longitudinal)***

To test whether SES modulates risk-reading outcomes, 2<sup>nd</sup> grade reading outcomes were compared in children who were at-risk based on low PA, RAN, and LSK performance in kindergarten in lower- and higher-SES groups. To test whether SES modulated the association between kindergarten left ILF and second-grade reading abilities, a multiple linear regression model was constructed as follows:  $WR\ scores = \beta_0 + \beta_1 \times gender + \beta_2 \times IQ + \beta_3 \times FA + \beta_4 \times SES + \beta_5 \times (reading \times SES) + \varepsilon$ . The Shapiro-Wilk test was performed to determine whether each variable was normally distributed. To probe the nature of the moderation effects identified in the regression model, correlations between ILF FA and second-grade reading were computed separately in each group after controlling for gender and IQ.

#### **5.4.5 *Classification of reading outcomes in lower- SES***

To test which protective factors are associated with good reading outcomes in lower-SES children, children from the lower-SES group were classified into typical and poor readers based on criteria outlined above. Mean FA across all four tracts (left ILF, SLF, AF, and right SLF) was computed. A Random Forests classification algorithm implemented in R statistics (Liaw & Wiener, 2002) was used to identify which of the following behavioral variables and pathways could differentiate among lower-SES children who became good versus poor readers: IQ, HLE, ARHQ, mean FA in left ILF, SLF, and AF and right SLF.

Random forests is an ensemble learning regression trees algorithm that constructs many decision trees at each level and outputs the mean prediction of the decision trees as the classification result. We generated 10,000 decision trees with

three levels. Each tree used a bootstrapped subsample of participants and a randomly selected subset of three predictor variables (mtry=3-based on Díaz-Uriarte & De Andres, 2006). Prediction error and variable importance were estimated using out-of-bag (OOB) samples. The OOB sample is a set of observations not used for building the current tree. OOB error estimates are derived by training the model on trees from bootstrapped subsamples and obtaining a test-set classification error on a third of the cases not included in the construction of the trees. Random forests uses the Gini Index as a measure for the best split selection, which evaluates the impurity of a given node with respect to the parent node. Models including combinations of all variables in different configurations were estimated. The combination of variables that yielded the lowest classification error was selected for inclusion in the final model and a confusion matrix was estimated. The most important features found from the variable importance map were then used in a leave-one-out cross validation procedure to estimate the final accuracy of the model and the confusion matrix.

## 5.5 Results

Using median split score of 51 on the BSMSS (mean = 48.43, SD = 11.95, range: 15-66), participants were divided into higher-SES ( $n = 61$ , 59.0% female) and lower-SES groups ( $n = 60$ , 47.5% female). There were no significant differences in age ( $t(118.82) = -0.33$ ,  $p = 0.94$ ,  $d = 0.01$ ) or gender ( $\chi^2(1) = 1.18$ ,  $p = 0.28$ ) between the groups. There were marginally significant differences in IQ, with lower-SES children exhibiting marginally lower scores ( $t(118.47) = 1.79$ ,  $p = 0.075$ ,  $d = 0.29$ ). Children from the lower-SES group had lower HLE composite

scores indicating a less rich home literacy environment ( $t(112.12) = 2.22, p = 0.03, d = 0.41$ ), but not higher ARHQ scores, indicating no group differences in parental history of reading difficulties ( $t(71.34) = -1.5, p = 0.14, d = 0.34$ ). In terms of SES, the educational range of the lower-SES group was high school/GED to master's degree (maternal mean years of education (BSMSS values) = 16.33, SD = 2.94; paternal mean = 14.21, SD = 3.43) and in the higher-SES group it was bachelor's degree to doctorate (maternal mean years of education = 19.69, SD = 1.81; paternal mean = 18.65, SD 2.59). Maternal occupational prestige scores ranged from 2 to 9 in the lower-SES group (mean = 5.55, SD = 2.09) and 5 to 9 in the higher-SES group (mean = 7.88, SD = 5.1). The paternal occupational prestige range in the lower-SES group was 5 to 40 (mean = 4.1, SD = 2.63) and in the higher-SES group it was 25 to 45 (mean = 7.77, SD = 1.12).

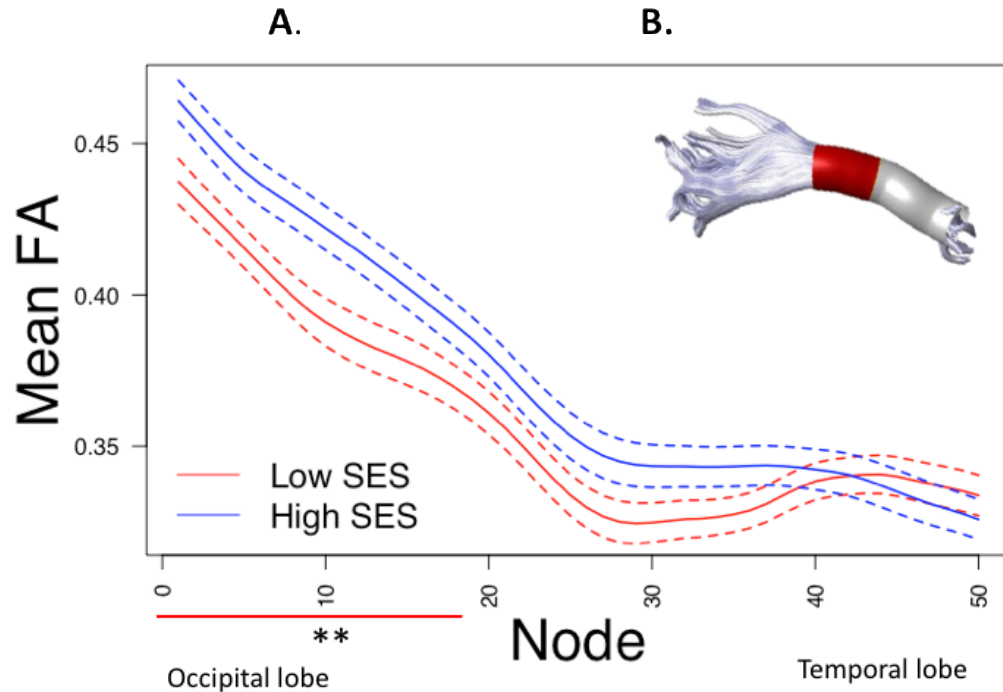
### **5.5.1 Differences by SES in (pre)literacy performance**

In kindergarten, the higher-SES group scored significantly better than the lower-SES group on phonological awareness ( $t(117.76) = 3.4, p < 0.001, d = 0.62$ ), vocabulary ( $t(116.38) = 3.83, p < 0.001, d = 0.7$ ), and letter-sound knowledge ( $t(118.5) = 2.61, p = 0.01, d = 0.47$ ), but not on RAN ( $t(117.73) = 0.75, p = 0.45, d = 0.14$ ). In second grade, the higher-SES group scored significantly better than the lower-SES group on all of the individual reading measures: WID ( $t(101.18) = 3.27, p = 0.001, d = 0.62$ ), WA ( $t(104.36) = 2.73, p = 0.007, d = 0.52$ ), PDE ( $t(103.74) = 3.36, p = 0.001, d = 0.63$ ), SWE ( $t(106.61) = 2.06, p = 0.04, d = 0.39$ ) and the second grade WR composite ( $t(99.01) = 2.93, p = 0.004, d = 0.55$ ). The same pattern was maintained after controlling for IQ: the higher-SES group scored higher on

phonological awareness ( $t(117.74) = 3.04, p = 0.003, d = 0.56$ ), vocabulary ( $t(118.13) = 3.39, p < 0.001, d = 0.62$ ), letter-sound knowledge ( $t(118.5) = 2.35, p = 0.02, d = 0.43$ ), as well as second grade WID ( $t(100.84) = 2.95, p = 0.003, d = 0.56$ ), WA ( $t(103.95) = 2.42, p = 0.02, d = 0.46$ ), PDE ( $t(99.95) = 3.05, p = 0.003, d = 0.57$ ), and WR ( $t(96.3) = 2.58, p = 0.01, d = 0.49$ ). Groups did not differ on RAN ( $t(117.39) = 0.48, p = 0.63, d = 0.09$ ) or SWE ( $t(103.26) = 1.7, p = 0.09, d = 0.32$ ). See *Supplemental Table 1* for behavioral descriptive statistics across the two SES groups and FDR-adjusted significance.

### **5.5.2 Differences between SES groups in tract FA**

We examined differences in FA between higher- and lower-SES groups in left ILF, SLF, and AF. After controlling for participant gender and IQ, the higher-SES group had significantly higher FA values in the occipito-temporal nodes 1-19 of the left ILF (Figure 1). There were no significant differences between the SES groups in the other tracts.



**Figure 5.1:** Fractional anisotropy differences between lower- and higher-SES children. A) Tract profile of the left ILF depicting FA values of lower- and higher-SES children at each of the 50 nodes. The solid line represents the mean FA and the dashed line represents the standard error. B) Nodes in left ILF in which higher-SES children exhibit significantly higher FA than lower-SES children are marked in red along the axis.  $**p < 0.05$ .

### 5.5.3 Differences between SES groups in AD and RD

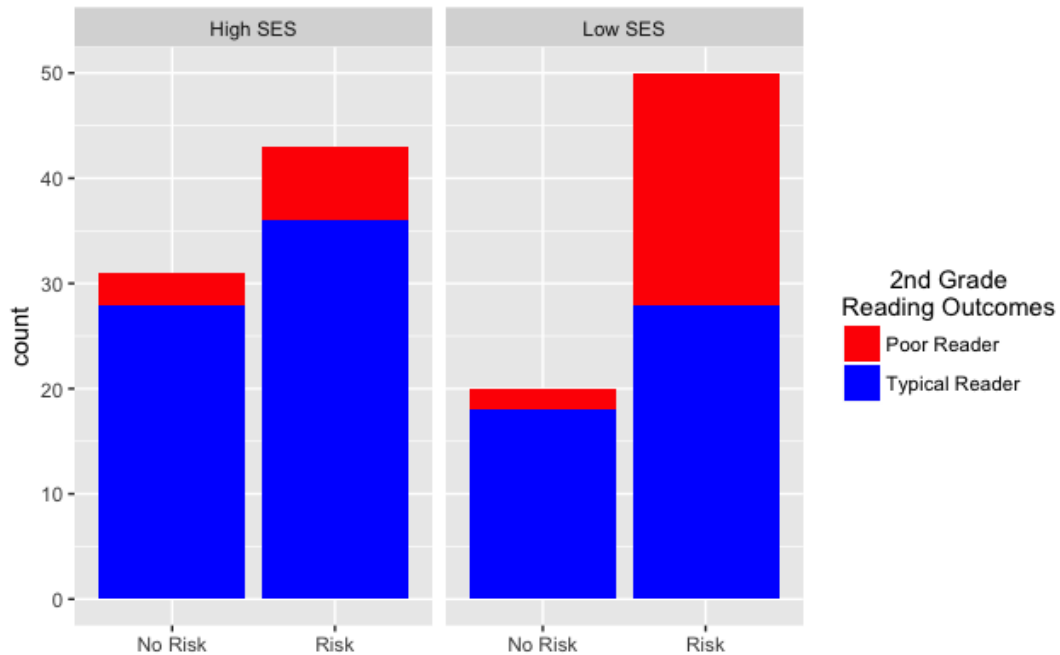
A follow-up analysis in nodes 1-19 of the left ILF found significantly lower RD in the higher-SES group in nodes 1-12 ( $p < 0.05$ ). There were no significant differences in AD between the groups.

### 5.5.4 Modulating effects of SES on the risk-reading relationship (longitudinal)

Using the criteria of lower 25<sup>th</sup> percentile on kindergarten PA, RAN or LSK, lower- and higher-SES children were divided into risk versus non-risk groups. According to the criteria of scoring below 90 on 2 of the 4 2<sup>nd</sup> grade word reading subtests, children were divided into Typical and Poor Readers. At-risk children from the lower-SES group were significantly more likely to end up in the Poor



Reader group as compared to at-risk children from higher-SES families (44% versus 16%; ( $\chi^2(4) = 22.29, p < 0.001$ ).

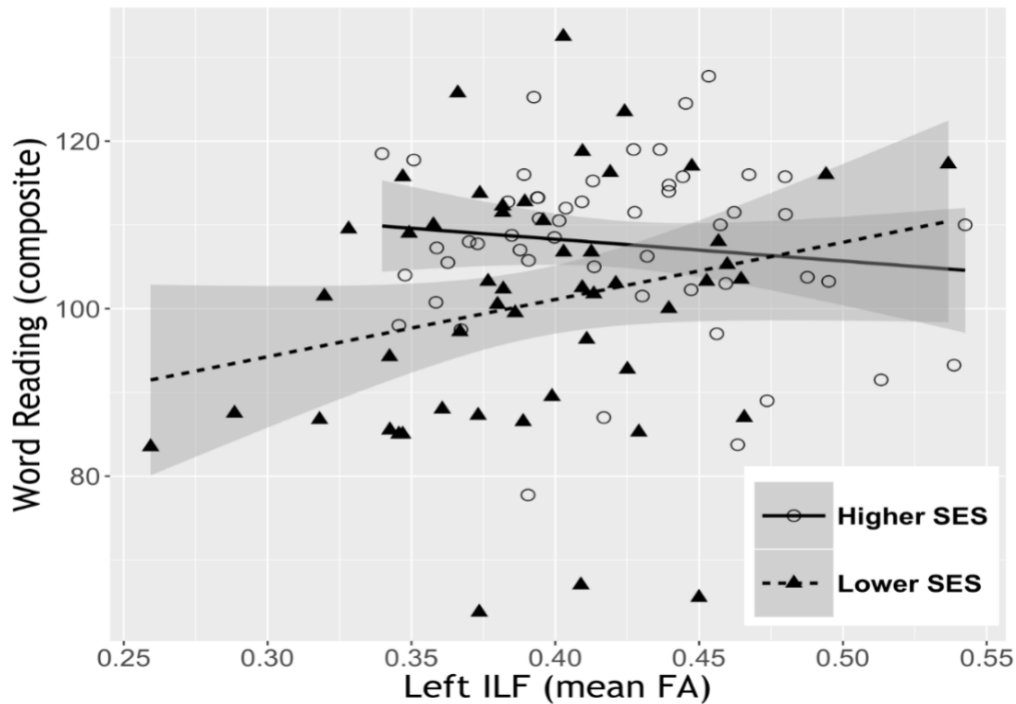


**Figure 5.2:** Differences in frequencies by KG risk, 2<sup>nd</sup> grade reading outcomes, and SES. Children from lower-SES families who were at-risk in kindergarten were significantly more likely to become poor readers as compared to children from the higher-SES group.

##### 5.5.5 *Modulating effects of SES on the FA-reading relationship (longitudinal)*

Mean FA was calculated for the nodes that were found to be significantly associated with SES. The Shapiro-Wilk test revealed that the WR variable was not normally distributed ( $W = 0.967, p = 0.007$ ). Accordingly, the *lmPerm* package (Wheeler, Torchiano, & Torchiano, 2016) in R was used to calculate permuted linear regression in order to test the interaction between left ILF FA and SES in predicting second-grade reading (the WR composite). Results indicated that after controlling for IQ ( $R^2 = 0.058, p = 0.023$ ) and gender ( $R^2 = 0.035, p = 0.784$ ), SES ( $R^2 = 0.052, p = 0.033$ ), the interaction between SES and mean left ILF FA ( $R^2 =$

0.03,  $p = 0.039$ ) accounted for unique and significant variance in WR. Mean ILF FA did not significantly predict WR overall ( $R^2 = 0.046$ ,  $p = 0.144$ ). Greater FA in the left ILF was associated with better reading outcomes in the lower-SES ( $R^2 = 0.062$ ,  $p = 0.028$ ), but not the higher-SES group ( $R^2 = 0.002$ ,  $p = 0.516$ ).



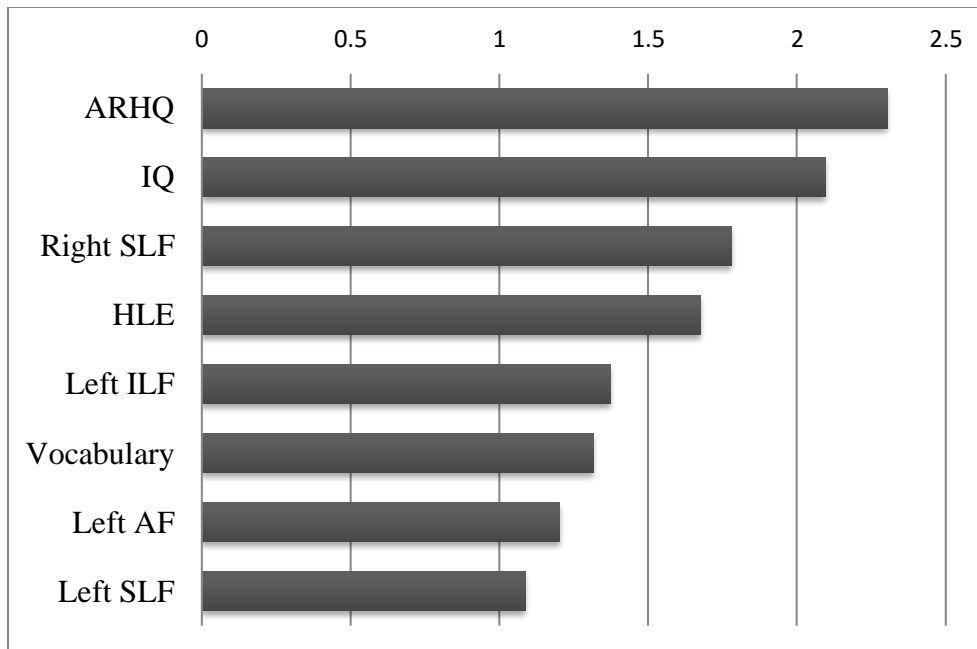
**Figure 5.3:** Association between tract FA (nodes 1-19) and reading by SES. Scatter plot depicts the significant positive association between FA and second grade reading composite in lower- but not higher-SES group. Word reading composite is a standard score based around 100.

#### 5.5.6 *Classification of reading outcomes in low SES group*

Children from the lower-SES group who had longitudinal data were divided into poor readers ( $n = 17$ ) and typical readers ( $n = 31$ ) in the second grade. Mean FA across the 50 nodes was computed for each of the four tracts. Random forests classification analysis was conducted on a subset of participants who had complete data (typical = 15, poor = 12). The children who were classified as typical readers

were younger ( $t(24.99) = 2.04, p = 0.053, d = 0.77$ ) and had higher SES scores ( $t(24.99) = -2.94, p = 0.009, d = 1.2$ ) than those who became poor readers.

The overall model including all variables yielded an estimated out-of-bag (OOB) error rate of 48.15%. Variables were ordered based on their significance (Figure 3), with parent history of reading difficulty (ARHQ) being the most significant in discriminating between the two groups. The top ten models with the lowest error rates are depicted in the Appendix (Table 1A). The model with the lowest error rate (43.3%) included ARHQ, HLE, IQ and right SLF (Table 1). The random forests analysis was repeated with these variables only and demonstrated a classification accuracy of 77.78%. The model was validated using leave-one-out cross validation. A model including only the behavioral measures yielded an accuracy of 63%. Due to significant differences in age across the two groups, age was included in a follow-up analysis to investigate whether it was a significant discriminator. Models including age yielded an OOB rate of 44.44%, with ARHQ, IQ, right SLF, and HLE remaining the most significant variables.



**Figure 5.4:** Variables ranked by relative importance. Relative importance represents mean decrease in Gini; HLE=home literacy environment, ARHQ=adult reading history questionnaire

Final Model: HLE, FHD, IQ, R_SLF 77.78% classification accuracy		Predicted		Error
		Low SES Poor Readers	Low SES Typical	
Observed	Low SES Poor Readers	9	3	0.25
	Low SES Typical Readers	3	12	0.20

**Table 5.1:** Model with the highest reading-outcome classification performance in lower-SES group.

## 5.6 Discussion

The present study found the expected relation between SES and both pre-literacy kindergarten measures and second-grade reading, such that children from higher-SES environments had higher pre-literacy and reading scores than children from lower-SES environments. At-risk children from lower-SES group were more likely to become poor readers in 2<sup>nd</sup> grade than children from higher-SES group. One novel finding was that children from lower-SES environments appeared to have weaker structural connectivity in the left inferior longitudinal fasciculus (ILF) as measured by reduced coherence (i.e., fractional anisotropy (FA)) and higher radial diffusivity (RD) in kindergarten. A second novel finding was there was a differential relation between white-matter microstructure in kindergarten and reading outcomes in second grade, such that in children from lower-SES, but not higher-SES environments, better reading in second grade was associated with greater FA in the left ILF in kindergarten. Further, among children from lower-SES environments, microstructural characteristics of the right SLF contributed to

classifying which children would become typical versus poor readers by the end of second grade. In combination, these findings reveal that white-matter characteristics in young kindergartners have a differential relation to long-term reading outcomes that depend upon their SES environments.

### **5.6.1 *Association between SES and (pre-) literacy performance***

Consistent with previous findings (Reardon, 2011), we observed significant associations between SES and reading-related performance in kindergarten and second grade. All pre-literacy skills examined, except RAN, were associated with SES. These pre-literacy skills (phonological awareness, letter knowledge, vocabulary knowledge) are all directly related to language. RAN involves rapid retrieval of color and object names and therefore indexes both lexical retrieval and domain-general speed of processing, both of which are important for reading ability. This independence of rapid naming from familial environmental factors has been demonstrated elsewhere (Heath et al., 2014) and may point to the unique contribution of this measure to predicting reading outcomes (Wolf & Bowers, 1999). The achievement gap between lower- and higher-SES children has been one of the most robust patterns in educational research (Reardon, 2011), and the confirmation of this gap in the current sample points to its representativeness.

### **5.6.2 *Left ILF tract coherence and SES***

Overall, the findings of reduced left ILF FA and increased RD in children from lower-SES families are in line with documented SES-related differences in neural structure and function across development (Hanson et al., 2013; Jednoróg et al., 2012; Lawson, Duda, Avants, Wu, & Farah, 2013; Mackey et al., 2015; Noble

et al., 2012; Noble, Wolmetz, Ochs, Farah, & McCandliss, 2006; Raizada & Kishiyama, 2010). Consistent with the present findings, lower parental SES has been associated with decreased FA in multiple white matter clusters including in the left ILF tract in older children (Gullick et al., 2016). To our knowledge, no studies to date have investigated differences in white matter based on SES in younger kindergarten children. Thus, this study is the first to reveal which reading-related pathways are affected in lower-SES children very early in reading development. Our results speak to the contribution of the early familial environment, rather than that of the quality of schooling, to observed SES-related disparities.

The left ILF is a component of the ventral reading pathway that transmits information between the visual word form area (VWFA) and the anterior and medial temporal lobe. This pathway is involved in orthographic processing and mapping printed words to meaning (Epelbaum et al., 2008; Wandell et al., 2012). The specific cluster that showed significant association with SES was located in the first 19 nodes of the pathway, within the cortical proximity of the occipito-temporal region of the VWFA. The VWFA is experience-dependent and develops specialization for print through learning to read a particular orthography (Dehaene & Cohen, 2011; Dehaene, Cohen, Morais, & Kolinsky, 2015; Dehaene et al., 2010; Heim, Pape-Neumann, van Ermingen-Marbach, Brinkhaus, & Grande, 2015). Specialized patterns of white-matter connectivity to the VWFA have been shown to predate, and possibly drive, the region's specialization for word recognition (Saygin et al., 2016). Taken together, this suggests that factors associated with

being raised in a lower-SES environment affects the connectivity patterns of the VFWA in a manner that could constrain reading development.

What could be the mechanisms that link SES and changes in these tracts? Fractional anisotropy is influenced by two opposing and partially environmentally-driven processes: myelination that increases FA, and pruning of axons that decreases FA (Yeatman, 2012a), as well as other factors. Studies in animal models have indicated that RD is more sensitive to myelination, whereas AD is more sensitive to axonal degradation (Song et al., 2003; Song et al., 2002). Findings of increased RD in the lower-SES group, but absence of differences in AD, therefore suggest that differences in FA are due to increased myelination of the left ILF in the higher-SES group. Myelination speeds conduction between distant cortical regions. Across development, myelination of axons is guided both by intrinsic genetic codes and extrinsically driven mechanisms, such as neuronal activity along a particular axon (Emery, 2010; Mount & Monje, 2017). For example, it has been demonstrated that neuronal activity stimulates the proliferation of committed oligodendrocyte progenitor cells (OPCs), which generate those glial cells, the oligodendrocytes, that form myelin (Mount & Monje, 2017). Developmentally, higher FA in childhood indicates maturation, as FA tends to increase through childhood, reaching a plateau in early-to-mid-adulthood (depending on the tracts) and then starting to decrease (Lebel, Treit, & Beaulieu, 2017). Thus, taken together, evidence suggests that reduced left ILF FA in the lower-SES group reflects decreased myelination of this pathway with negative consequences for their reading development.



Our current understanding of how SES affects brain development is modest. It is hypothesized that the quality of parental care varies depending on SES factors and can influence, neural development through epigenetic processes, and subsequently, the neural activity that regulates cognition (Hackman et al., 2010). In the present study, there was evidence that children from higher-SES environments had enhanced home literacy environments (HLE scores) and vocabulary relative to children from lower-SES environments. These enhanced language and pre-literacy experiences may have spurred increased myelination of the left ILF as measured by higher FA and lower RD values.

There is considerable evidence that experience or training can alter white-matter FA. For example, increased FA in target tracts has been demonstrated in response to juggling training (Scholz, Klein, Behrens, & Johansen-Berg, 2009), meditation training (Tang et al., 2010), cognitive training (Mackey, Whitaker, & Bunge, 2012), learning to read (Carreiras et al., 2009; Thiebaut de Schotten et al., 2014), and following reading intervention (Keller & Just, 2009). For example, in study of adults, learning Morse code was associated with significant increase in the left ILF FA (Schlaffke, Leemans, Schweizer, Ocklenburg, & Schmidt-Wilcke, 2017). Importantly, the effects of parental practices (e.g., quality of nutrition, conflict in the household, and verbal abuse) on white matter structure in early childhood have been shown (Choi, Jeong, Rohan, Polcari, & Teicher, 2009; Dufford & Kim, 2017; Lebel et al., 2016; Ou et al., 2014). Therefore, the present finding of the region's sensitivity to influences related to familial SES is in

accordance with its general malleability in response to environmental input and experience.

### **5.6.3 *Modulating effects of SES on left ILF-reading relationship***

Consistent with previous studies that demonstrated modulatory influences of SES on the brain-reading relationship (Brito, Piccolo, & Noble, 2017; Gullick et al., 2016; Noble et al., 2006; Romeo et al., 2017), left ILF FA in kindergarten was associated with better second-grade reading outcomes in lower- but not higher-SES children. The longitudinal nature of the current findings is a notable difference from other studies that only examined the SES x brain association with reading concurrently. One possible interpretation of these findings is that SES can exaggerate or mitigate the links between early brain connectivity and reading outcomes. For example, higher quality childcare and schools, better housing, and higher quality and quantity of cognitive and linguistic stimulation could lead to successful reading outcomes even in the face of neuroanatomical variations—namely reduced white matter coherence of left ILF (Brito et al., 2017). Another possible interpretation is related to differences in heritability of reading across SES groups. It has been shown that environmental differences accounted for more variance in reading performance in lower-SES children than in higher-SES children (Friend et al., 2008; Kirkpatrick, Legrand, Iacono, & McGue, 2011). This raises the possibility that the neurobiological basis of reading varies based on SES, with stronger experiential effects on brain-reading relationships in children from lower-SES families. Future studies are needed to evaluate these possibilities and

determine the precise mechanisms of the modulatory influences of SES on the brain-reading links.

#### **5.6.4 *Classification of reading outcomes in the lower-SES group***

We discovered that microstructural characteristics of the right SLF, specifically greater FA, in kindergarten contributed significantly to predicting which children from lower-SES environments would progress to typical versus poor reading in second grade. Prior research has shown that family history of reading difficulty (e.g., Pennington & Lefly, 2001), home literacy environment (Burgess, Hecht, & Lonigan, 2002; Sénéchal, 2002), and IQ (Cardon, Dilalla, Plomin, DeFries, & Fulker, 1990) are all related to reading outcomes, but this is the first evidence that white-matter microstructure is also predictive of future reading ability in pre-literate children. Indeed, all of these factors were predictive of future reading ability in the present study, but it is noteworthy that FA of the right SLA significantly improved the accuracy of the prediction model. This is consistent with other studies reporting that brain measures can enhance prediction accuracy beyond conventional familial and behavioral measures (Gabrieli et al., 2015; Ozernov-Palchik & Gaab, 2016).

It is also noteworthy that it was a right dorsal white-matter tract (the right SLF) that predicted future reading outcomes rather than the three left-hemisphere tracts (AF, SLF, and ILF) that are most often associated with reading. Prior studies of dyslexia that did not consider SES (and most likely were largely representing higher-SES families) have noted that greater right prefrontal activation and increased SLF coherence were associated with improved reading outcomes.

Specifically, increased task-related activation in the inferior frontal gyrus predicted reading gains in children with dyslexia prior to treatment (Hoeft et al., 2011) and differentiated responders from non-responders following treatment (Odegard, Ring, Smith, Biggan, & Black, 2008). Further, interventions aimed at helping poor readers have yielded increased activations in the right prefrontal cortex (see a meta analysis by Barquero et al., 2014). Notably, right SLF FA predicted longitudinal reading gains in children with dyslexia (Hoeft et al., 2011) and higher rates of FA development were reported for children with familial risk of dyslexia who became good, versus poor, readers (Wang et al., 2016). The present study, therefore, suggests for the first time that the microstructure of the right SLF is relevant not only for reading outcomes in children or adults with dyslexia, but also for children from lower SES environments.

The linear regression analysis and the random forests analysis reveal different variables associated with reading outcomes, with the left ILF being correlated with reading among lower-SES children, but the right SLF being more diagnostic of future reading outcomes in random forests analysis. This dissociation reflects inherent differences between the two statistical approaches. Unlike regression methods, random forests analysis is highly sensitive for modeling non-linear and complex interactions between variables, and is robust to small sample sizes and outliers (Grömping, 2007). Thus, our results suggest that lower-SES children recruit additional pathways for reading to compensate for reduced connectivity of the left-hemispheric ventral reading system.

### **5.6.5 *Limitations and future directions***

First, although the sample of lower-SES children in the current study was racially and ethnically diverse, it was not truly representative of the lower segment of the United States population in terms of SES. All mothers and 95% of fathers in the current lower-SES sample had completed high school, whereas the national high school completion rate stands at 83% (Aud, 2012). Nevertheless, the mean SES of the sample is comparable to previous neuroimaging studies of SES in children (e.g., Gullick et al., 2016; Noble et al., 2006; Skoe, Krizman, & Kraus, 2013) and even greater disparities between the two SES groups would be expected were the lower-SES group less advantaged. Second, because white matter development begins in utero and continues through adolescence, in order to gain a deep, mechanistic understanding of the role of parental SES in a child's brain development and to disentangle hereditary from environmental influences, future work needs to examine these effects in infancy and longitudinally. Finally, whereas the random forests analysis is robust to small group sizes, it usually requires a large number of variables (Breiman, 2001). Therefore, a more comprehensive study with multiple variables is needed to better understand the cognitive, familial, and neural factors that play a compensatory role in reading attainment in low socioeconomic environments.

### **5.7 *Significance***

Our paper is the first to demonstrate SES-related brain differences in a large, heterogeneous sample of children before reading instruction and the longitudinal significance of these differences for reading development. This study extends

previous findings on the socioeconomic disparities in education and reveals potential neurodevelopmental mechanisms that underlie these disparities.

	Low SES			High SES					
<b>Kindergarten</b>	mean	SD	range	mean	SD	range	t value		cohen's d
Age	67.81	3.83	60-75	66.94	4.24	60-80	-0.33		0.06
IQ	97.81	9.25	80-122	97.81	10.33	80-131	1.79		0.33
PA Mean	8.92	1.77	13-Jun	10.75	1.68	7.33-15.67	3.40	*	0.62
RAN Mean	93.13	13.24	59-115	98.53	13.96	68.33-121.33	0.75		0.14
PPVT	108.26	12.45	81-133	119.81	10.69	101-146	3.83	*	0.7
LSK	96.36	14.33	67-126	102.46	13.08	79-134	2.61	*	0.48
<b>Second Grade</b>									
WID	100.23	12.78	69-126	110.51	11.08	80-132	3.27	*	0.62
WA	97.34	13.07	77-128	107.06	12.04	78-133	2.73	*	0.52
SWE	98.54	12.6	55-124	105.98	11.32	74-126	2.06	*	0.39
PDE	91.34	14.1	66-119	102.91	11.96	73-124	3.36	*	0.63
<b>Background</b>									
HLE	5.15	0.79	3.67-7.4	5.41	0.83	2.83-7.4	2.22	*	0.41
ARHQ	0.33	0.09	0.12-0.58	0.3	0.11	0.14-0.57	-1.50		0.34

**Supplemental Table 5.1:** Descriptives by SES group.

\* *FDR adjusted  $p < 0.05$*

<b>Models</b>	<b>OOB Error</b>
ARHQ, hle, right SLF	0.43
ARHQ, hle, left ILF, right SLF	0.45
ARHQ, hle, right SLF, IQ	0.45
ARHQ, left AF, left SLF, Vocabulary, right SLF, IQ	0.45
ARHQ, left AF, left SLF, right SLF, IQ	0.45
ARHQ, left AF, right SLF, IQ	0.45
ARHQ, left ILF, right SLF, IQ	0.45
ARHQ, Vocabulary, right SLF, IQ	0.45
ARHQ, right SLF, IQ	0.45
ARHQ, hle, left AF, left ILF, right SLF, IQ	0.50
ARHQ, left AF, Vocabulary, right SLF, IQ, age	0.50
ARHQ, left AF, right SLF, IQ, age	0.50

**Supplemental Table 5.2:** Top ten models tested with random forests.

## **6 Conclusions and Future Directions**

Reading represents a nexus of perceptual and cognitive neural systems, which means that any neural component could underlie a form of dyslexia. Indeed, impairments associated with dyslexia have been demonstrated across multiple domains. In attempts to reconcile this wide range of findings, micro-models of dyslexia have been proposed, each accounting for a small fraction of dyslexia-related deficits. All-encompassing theories have also been proposed to explain all of the deficits observed in dyslexia. The former attempts led to suggestions of abandoning the term ‘dyslexia’ altogether (Elliott, 2014), while the latter attempts have been targeted with a barrage of “falsification bullets” that aim to shoot down one theory after another with counter evidence. Rather than rushing to replace one theory with a better one, it seems more productive to work together as a field to build cumulatively upon previous research, and to nurture and update existing theories. In other words, it makes sense to treat theories regarding dyslexia like graduate students:



“Once admitted, one tries hard to avoid flunking them out (of course, not at all costs) and one spends much time and effort on their development so that they may become long-term contributors to science.” (Roelofs, 2005, *p.* 314).

Thus, I believe that the field should collectively move towards collaboratively and critically developing a comprehensive theory of dyslexia etiology by applying rigorous methodology and practicing open science. Replication of previous studies would allow for either the strengthening or correction of previous models. Replication studies are not possible without methodological transparency, and futile if null results are not reported. Data sharing practices would support larger data points, resulting in increased robustness and reliability of results. It is my personal goal to apply these principles to extending the current work in advancing the development of a comprehensive and accurate theory of dyslexia.

## **6.1 Theories of Dyslexia Etiology**

Goswami (2012) outlined several research designs that are necessary to establish the etiological basis of dyslexia: 1) studying pre-readers; 2) using reading-level matched controls; 3) testing effects on other cognitive systems 4) longitudinal follow-up; 5) training studies; 6) cross-language studies; and 7) research with illiterate adults. The current study applied the first three designs to evaluate the most common deficits in dyslexia from a developmental cognitive neuroscience perspective. I hope the findings make a contribution, however modest, in support of several existing theories of dyslexia, namely Wolf & Bowers’ *double-deficit hypothesis*, Pennington’s *multi-deficit model*, and Ahissar’s *anchoring hypothesis*. Importantly, Wolf and Bowers conceptualized the double-deficit hypothesis as a

placeholder for what would emerge as the multi-deficit view of dyslexia. The findings presented here support this view by demonstrating multiple deficit profiles associated with dyslexia risk. The applicability of the current findings for many of the other theories reviewed in Chapter 1 is limited by the scope of the measures selected. For example, visual skills were not examined in the current study. Therefore, it is possible that deficits in these skills will characterize some of the children in the study. Thus, this work can only support certain hypotheses, but is inconclusive about the multitude of other hypotheses proposed for dyslexia.

The current thesis revealed distinct cognitive and neural markers of developmental dyslexia prior to receiving formal reading instruction. Study 1 demonstrated that cognitive profiles of risk are heterogeneous and stable. Children who are in the PA, RAN, double-deficit, or LSK risk profile, remained in their respective profile two years later. Study 2 demonstrated that the double-deficit and the LSK risk profiles were associated with significant impairments in reading fluency and comprehension, respectively, at the end of 2nd grade (three years later). Interestingly, the PA and RAN risk groups showed average reading performance on both comprehension and fluency reading measures, suggesting that each deficit independently is not sufficient for causing reading difficulties. Additionally, since these two profiles were overrepresented in lower SES schools, it is likely that they are more related to environmental factors that could affect their stability. Study 3 demonstrated rhythm discrimination impairments in the PA, but not RAN risk groups. Importantly, the impairment was uniquely related to processing of metrical rhythms, suggesting a putative role of impairment in extraction of temporal

regularity in poor phonological development. The mediation model confirmed the direction of causality from rhythm processing to phonological awareness and to reading skills. These findings therefore support the theory of a core deficit in the processing of rhythmic cues in music and in speech leading to phonological impairments in dyslexia.

Neuroimaging evidence also supports the dissociation of risk profiles early in development. Each of the risk profiles had unique patterns of grey matter volume and white matter coherence reductions. These patterns of reductions were consistent with the underlying impairments proposed for each profile. In the LSK profile, the dorsal network was implicated; in the double-deficit risk, memory and learning networks; in RAN risk, regions that support lexical retrieval; and in PA risk, regions that support working memory. Three white matter tracts were examined in studies 2 and 4 using two different diffusion methods. Results showed that left AF and SLF, but not ILF, are associated with dyslexia risk early in reading development. Interestingly, RAN risk was associated with reduced coherence in both tracts, pointing to a more general connectivity impairment in the group. Suggesting more experientially driven deficits, left ILF, however, was associated with socioeconomic background in Study 4, and with risk for reading failure in children from lower socioeconomic background.

## **6.2 Multi-Factorial Model of Risk**

These findings support the existence of multiple cognitive, neurobiological, and environmental pathways to reading impairment. Consistent with the multi-componential models of dyslexia, these findings indicate that dyslexia risk is multi-

factorial and probabilistic — that is, multiple genetic, neural, environmental, and cognitive factors interact reciprocally to increase risk for dyslexia. The nature of interaction among these factors is not completely understood. Large epidemiological studies are needed to comprehensively characterize these factors in using behavioral, neuroimaging, and genetic methods. Pending such studies, the complexity of these factors stands in opposition to attempts to identify a single-deficit theory of dyslexia. Most recently, for example, a theory has been proposed that dyslexia stems from a deficit in neural hyper-excitability (Hancock, Pugh, Hoeft, 2018). This theory was able to successfully account for a wide range of dyslexia findings of noisier and less reliable processing of stimuli across different domains, such as less categorical speech processing and slower naming. This theory, however, can-not successfully account for the findings of allophonic perception in dyslexia that suggest increased precision of discrimination (Banai & Ahissar, 2018). Characterizing these complex interactions in pre-reading children at-risk for dyslexia, longitudinally, can move us forward to identifying the origins of such deficits. For example, the current finding from a small number of children, showed brain atypicalities in the neural regions supporting memory systems (systems that have also been implicated in short-term statistical learning-Turk-Browne et al., 2009; Schapiro et al., 2017) in the double-deficit profile. These findings, together with the results of unique deficit in processing temporally regular structure in musical rhythm in at-risk pre-reading children, both suggest that double-deficit could represent a deficit in learning from regularity.

### **6.3 Early Identification of Risk**

Since dyslexia is defined as a failure to develop typical reading skills, it is currently not diagnosed until reading failure is substantial and chronic, usually in the second grade or later. As a result, children diagnosed with dyslexia at this point must make an enormous effort to close the gap between them and their typically-reading peers (Hiebert, 2000). Years of failure to read can lead to reduced self-esteem, depression and other psychological and clinical implications (Valas, 1999; Riddick, 2009). Moreover, targeted interventions are most effective when administered early — in kindergarten and first grade (Torgesen, 2000).

In order to prevent the detrimental consequences of the waiting-for-failure approach, it is important to identify which measures and, as shown in the current thesis, which patterns of performance on these measures, in kindergarten, are strong predictors of future reading success and failure. The use of data driven methods yields heterogeneous and stable profiles of risk. In educational settings, however, such an approach is not feasible as it requires large samples of data and sophisticated analytical methods. Instead, cut-off scores are used to separate children at little risk from those at high-risk for prolonged reading difficulty. If the cut-off is conservative (i.e., only a very low score is used to identify a child at risk), then too many children who will read poorly are missed (low sensitivity) and they will not be helped. If the cut-off is liberal (i.e., a somewhat higher score is used to identify a child at risk), then nearly all children at true risk will be identified, but many children who are not at true risk will be over-identified (low specificity). This over-identification is problematic because remediation resources, such a reading

specialists who can work one-on-one or in small groups of children, are scarce in most schools. Another challenge in using cut-off scores is that they yield a binary risk classification with little information for how intervention can be tailored for each individual child's profile. As demonstrated here, different risk profiles develop distinct reading impairments, and would likely benefit from different types of intensive instruction (e.g., one targeting accuracy/fluency and another comprehension).

Current research joins a growing number of studies that have demonstrated that neural markers of dyslexia are already evident in pre-reading children and infants who are at risk for dyslexia. Furthermore, consistent with the current findings, longitudinal studies have shown that brain measures make a unique contribution and even outperform behavioral tests in predicting reading gains and outcomes. Does neuroimaging hold the key for early and more precise identification of dyslexia? Much work lies ahead in identifying reliable biomarkers of dyslexia. Most studies to date, including the one reported here, were conducted in a small number of children and are based on group-level differences. Biomarkers must be robust enough to withstand neuroanatomical differences across individuals. Additionally, even with neuroimaging measures included, the best models reported so far achieved 94% accuracy in predicting reading outcomes (Bach et al., 2013). As pointed out by Poulsen (2018), although this sounds impressive, given a population prevalence of dyslexia at about 10%, a model can accomplish similar classification accuracy by simply predicting that nobody would develop dyslexia. There is room for optimism, however. Neuroimaging research into dyslexia is still

in its infancy, as compared to behavioral research, and has already made significant contributions toward understanding the elusive nature of dyslexia etiology. The first cohort of neuroimaging studies revealed that dyslexia is a neurobiological disorder. The second cohort has shown that the neural characteristics of dyslexia are present prior to reading onset and could be causal to reading failure. It is hoped that the next cohort of studies will “stand on the shoulders of giants,” relying on large data sets of data already collected to characterize the individual neural differences in dyslexia. This, in turn, will inform both the etiological understanding of dyslexia and the behavioral classification of risk, resulting in earlier and more precise interventions.

#### **6.4 Socioeconomic Disparities in Reading**

Environmental disparities have significant developmental implications across a range of domains, with language and reading being particularly affected (Hackman et al., 2010). In the current research, the prevalence of certain profiles of risk varied depending on the socioeconomic status (SES) of children’s schools. The specificity of pre-literacy measures and profiles of risk that are sensitive to environmental influences remains to be established. A contradiction in findings emerged between *Study 1* and *Study 4*, with the former reporting increased prevalence of the RAN risk profile in low-SES schools, but the latter finding no significant association between RAN in kindergarten and SES. This discrepancy is likely due to the type of SES measured (school-level versus parental) and populations sampled. There is a significant genetic component to RAN (Compton et al., 2001), but the knowledge of stimuli is influenced by environmental variables,

particularly early in development. Overall, RAN is considered a highly predictive and environmentally neutral measure, particularly later in development (i.e., beyond kindergarten).

Even in kindergarten, children from lower-SES backgrounds showed significant behavioral and neural disparities in reading-related skills and networks as compared to children from more privileged backgrounds. Importantly, quality of home environment (e.g., literacy environment) and compensatory neural systems mitigated the effects of SES on reading development. These findings suggest that SES effects on reading-development likely start even before the first day of kindergarten. Therefore, in order to narrow the SES-related educational attainment gaps, policy should target the early environments into which children are born by enriching the quantity and quality of home language and literacy practices.

The modulatory effects of SES on the risk-to-outcome pathway established in literature, mean that the quality of education matters the most for children from low SES backgrounds. Current policies consider SES as exclusionary criteria for eligibility for intervention services; as children receive specific learning disability diagnosis if their reading problems

“[. . . ]are primarily the result of visual, hearing, or motor disabilities, intellectual disability, emotional disturbance, or environmental, cultural, or economic disadvantage”

(from Individuals with Disabilities Education Act (IDEA) of 2004, [34 CFR §300.8(c)(10)].



Such policies deprive services from children who need them the most. Current evidence suggests that the profiles of risk in lower-SES children reflect those in higher-SES children. In the absence of an appropriate educational intervention, however, the outcomes are significantly worse in lower-SES children.

## **6.5 Bridging Research and Practice**

An important goal of science is the betterment of human life. The translational potential of neuroscience research for education, however, has been called “a bridge too far” (Bowers, 2016; Bruer, 1997). In his response, Gabrieli (2016) stated that:

“The immediate goal of such human cognitive neuroscience has not been the betterment of thought, emotion, or social interaction, but discovery of how the brain empowers such human endowments.”

Indeed, I hope that the discoveries presented in this thesis can help inform educational practices by propelling knowledge of *how* reading failure risk presents itself early in development, both cognitively and in the brain. These discoveries can thereby contribute to the betterment of outcomes for children with dyslexia.

## References

- Aad, G., Abbott, B., Abdallah, J., Abdelalim, A. A., Abdesselam, A., Abdinov, O., . . . Zutshi, V. (2010). Search for new particles in two-jet final states in 7 TeV proton-proton collisions with the ATLAS detector at the LHC. *Phys Rev Lett*, 105(16), 161801.
- Abel, T. J., Rhone, A. E., Nourski, K. V., Ando, T. K., Oya, H., Kovach, C. K., . . . Tranel, D. (2016). Beta modulation reflects name retrieval in the human anterior temporal lobe: An intracranial recording study. *Journal of neurophysiology*, jn. 00012.02016.
- Aboud, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental science*, 19(4), 632-656.
- Achal, S., Hoeft, F., & Bray, S. (2016). Individual differences in adult reading are associated with left temporo-parietal to dorsal striatal functional connectivity. *Cerebral cortex*, 26(10), 4069-4081.
- Acres, K., Taylor, K., Moss, H., Stamatakis, E., & Tyler, L. (2009). Complementary hemispheric asymmetries in object naming and recognition: a voxel-based correlational study. *Neuropsychologia*, 47(8), 1836-1843.
- Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., & Reiss, A. L. (2002). A developmental fMRI study of the Stroop color-word task. *Neuroimage*, 16(1), 61-75.
- Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in cognitive sciences*, 11(11), 458-465.
- Ahissar, M., Lubin, Y., Putter-Katz, H., & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature neuroscience*, 9(12), 1558-1564.
- Allor, J. H., Mathes, P. G., Jones, F. G., Champlin, T. M., & Cheatham, J. P. (2010). Individualized research-based reading instruction for students with intellectual disabilities: Success stories. *Teaching exceptional children*, 42(3), 6-12.

- Amitay, S., Ben-Yehudah, G., Banai, K., & Ahissar, M. (2002). Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain*, 125(10), 2272-2285.
- Anthoni, H., Sucheston, L. E., Lewis, B. A., Tapia-Paez, I., Fan, X., Zucchelli, M., . . . Kere, J. (2012). The Aromatase Gene CYP19A1: Several Genetic and Functional Lines of Evidence Supporting a Role in Reading, Speech and Language. *Behav Genet*, 42(4), 509-527.
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of experimental child psychology*, 83(2), 111-130.
- Apfelbaum, K. S., Hazeltine, E., & McMurray, B. (2013). Statistical learning in reading: Variability in irrelevant letters helps children learn phonics skills. *Developmental psychology*, 49(7), 1348.
- Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive science*, 36(2), 286-304.
- Ashburner, J., Barnes, G., Chen, C.-C., Daunizeau, J., Flandin, G., Friston, K., . . . Moran, R. (2014). SPM12 Manual.
- Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage*, 11(6), 805-821.
- Aso, K., Hanakawa, T., Aso, T., & Fukuyama, H. (2010). Cerebro-cerebellar interactions underlying temporal information processing. *Journal of cognitive neuroscience*, 22(12), 2913-2925.
- Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X., and Zhang, J. . (2012). *The condition of education*. Retrieved from Washington D.C.:
- Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *Neuroimage*, 82, 605-615.
- Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, 45(13), 2883-2901.
- Bakker, D. (1972). Temporal order in disturbed reading. *Bulletin of the Orton society*, 22(1), 80-83.
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral cortex*, 19(11), 2699-2707.
- Barquero, L. A., Davis, N., & Cutting, L. E. (2014). Neuroimaging of reading intervention: a systematic review and activation likelihood estimate meta-analysis. *PLoS One*, 9(1), e83668.
- Barratt, W. (2006). The Barratt simplified measure of social status (BSMSS). *Indiana State University*.
- Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magnetic resonance in medicine*, 44(4), 625-632.

- Beattie, R. L., & Manis, F. R. (2012). Rise Time Perception in Children With Reading and Combined Reading and Language Difficulties. *J Learn Disabil*.
- Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system—a technical review. *NMR in Biomedicine*, 15(7-8), 435-455.
- 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.
- Ben-Shachar, M., Dougherty, R. F., & Wandell, B. A. (2007). White matter pathways in reading. *Current opinion in neurobiology*, 17(2), 258-270.
- Beneventi, H., Tønnessen, F. E., Ersland, L., & Hugdahl, K. (2010). Executive working memory processes in dyslexia: behavioral and fMRI evidence. *Scandinavian journal of psychology*, 51(3), 192-202.
- Beneventi, H., Tønnessen, F. E., Ersland, L., & Hugdahl, K. (2010). Working memory deficit in dyslexia: behavioral and FMRI evidence. *International journal of neuroscience*, 120(1), 51-59.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., & Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural brain research*, 125(1-2), 279-284.
- Betancourt, L. M., Avants, B., Farah, M. J., Brodsky, N. L., Wu, J., Ashtari, M., & Hurt, H. (2015). Effect of socioeconomic status (SES) disparity on neural development in female African-American infants at age 1 month. *Developmental science*.
- Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: a comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, 7(2), 113-123.
- Bickel, P. J., & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The annals of statistics*, 1196-1217.
- 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.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state. A functional MRI study. *Journal of cognitive neuroscience*, 11(1), 80-95.
- Birch, H. G. (1962). Dyslexia and the maturation of visual function. *Reading disability: Progress and research needs in dyslexia*, 161-169.
- Bishop-Liebler, P., Welch, G., Huss, M., Thomson, J. M., & Goswami, U. (2014). Auditory temporal processing skills in musicians with dyslexia. *Dyslexia*, 20(3), 261-279.
- Bishop, D. V., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: same or different? *Psychological bulletin*, 130(6), 858-886.
- Blachman, B. A. (2000). Phonological awareness. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 483-502). Mahwah, NJ: Lawrence Erlbaum.

- Black, J. M., Myers, C. A., & Hoeft, F. (2015). The Utility of Neuroimaging Studies for Informing Educational Practice and Policy in Reading Disorders. *New directions for child and adolescent development*, 2015(147), 49-56.
- Black, J. M., Tanaka, H., Stanley, L., Nagmine, M., Zakerani, N., Thurston, A.,... & Hoeft, F. (2012). Maternal history of reading difficulty is associated with reduced language related gray matter in beginning readers. *NeuroImage*, 59(3), 3021-3023.
- Black, J. M., Xia, Z., & Hoeft, F. (2017). Neurobiological bases of reading disorder part II: The importance of developmental considerations in typical and atypical reading. *Language and linguistics compass*, 11(10).
- Blomert, L., & Mitterer, H. (2004). The fragile nature of the speech-perception deficit in dyslexia: Natural vs. synthetic speech. *Brain and language*, 89(1), 21-26.
- Blumenfeld, H. K., Booth, J. R., & Burman, D. D. (2006). Differential prefrontal-temporal neural correlates of semantic processing in children. *Brain and language*, 99(3), 226-235.
- Boets, B., de Beeck, H. P. O., Vandermosten, M., Scott, S. K., Gillebert, C.R., Mantini, D., .. & Ghesquiere, P. (2013). Intact but less accessible phonetic representation in adults with dyslexia *Science*, 342(6166), 1251-1254.
- Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., & Ghesquière, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Research in developmental disabilities*, 32(2), 560-570.
- Bogliotti, C., Serniclaes, W., Messaoud-Galusi, S., & Sprenger-Charolles, L. (2008). Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of experimental child psychology*, 101(2), 137-155.
- Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual review of neuroscience*, 25(1), 151-188.
- Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Gaillard, W., & Theodore, W. (1995). Regional cerebral blood flow during object naming and word reading. *Human brain mapping*, 3(2), 93-106.
- Booth, J. R., Bebko, G., Burman, D. D., & Bitan, T. (2007). Children with reading disorder show modality independent brain abnormalities during semantic tasks. *Neuropsychologia*, 45(4), 775-783.
- Booth, J. R., Burman, D.D, Van Santen, F.W., Harasaki, Y., Gitelamn, D.R., Parrish, T.R., Mesulam, M.M. (2001). The development of specialized brain systems in reading and oral language. *Child Neuropsychology*, 7(3), 119-141.
- Boscardin, C. K., Muthén, B., Francis, D. J., & Baker, E. L. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of educational psychology*, 100(1), 192.
- Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, 104, 198-230.

- Bosse, M. L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance: a cross-sectional study. *Journal of research in reading*, 32(2), 230-253.
- Bowers, J. S. (2016). Psychology, not educational neuroscience, is the way forward for improving educational outcomes for all children: Reply to Gabrieli (2016) and Howard-Jones et al.(2016).
- Bowey, J. A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. *Journal of educational psychology*, 87(3), 476.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual review of psychology*, 53(1), 371-399.
- Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of experimental child psychology*, 35(2), 345-367.
- Brambati, S., Ogar, J., Neuhaus, J., Miller, B., & Gorno-Tempini, M. (2009). Reading disorders in primary progressive aphasia: a behavioral and neuroimaging study. *Neuropsychologia*, 47(8), 1893-1900.
- Breier, J. I., Gray, L., Fletcher, J. M., Diehl, R. L., Klaas, P., Foorman, B. R., & Molis, M. R. (2001). Perception of voice and tone onset time continua in children with dyslexia with and without attention deficit/hyperactivity disorder. *Journal of experimental child psychology*, 80(3), 245-270.
- Breier, J. I., Gray, L. C., Fletcher, J. M., Foorman, B., & Klaas, P. (2002). Perception of speech and nonspeech stimuli by children with and without reading disability and attention deficit hyperactivity disorder. *Journal of experimental child psychology*, 82(3), 226-250.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., . . . Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proceedings of the national academy of sciences*, 107(17), 7939-7944.
- Brennan, C., Cao, F., Pedroarena-Leal, N., McNorgan, C., & Booth, J. R. (2013). Reading acquisition reorganizes the phonological awareness network only in alphabetic writing systems. *Human brain mapping*, 34(12), 3354-3368.
- Brewer, C. L. (1967). Presentation time, trials to criterion, and total time in verbal learning. *Journal of experimental psychology*, 73(1), 159.
- Breznitz, Z. (2006). Fluency in reading: synchronization of processes. *Routledge*.
- Breznitz, Z. (2008). The origin of dyslexia: The asynchrony phenomenon. *The Sage handbook of dyslexia*, 11-29.
- Breznitz, Z., & Misra, M. (2003). Speed of processing of the visual-orthographic and auditory-phonological systems in adult dyslexics: The contribution of "asynchrony" to word recognition deficits. *Brain and language*, 85(3), 486-502.
- Brito, N. H., Piccolo, L. R., & Noble, K. G. (2017). Associations between cortical thickness and neurocognitive skills during childhood vary by family socioeconomic factors. *Brain and cognition*, 116, 54-62.

- Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *The future of children*, 55-71.
- Bruck, M. (1992). Persistence of dyslexic's phonological awareness deficits. *Developmental psychology*, 28(5), 874.
- Bruer, J. T. (1997). Education and the brain: A bridge too far. *Educational researcher*, 4-16.
- Burbridge, J., T., Wang, Y., Volz, A. J., Peschansky, V. J., Lisann, L., . . . Rosen, G. D. (2008). Postnatal analysis of the effect of embryonic knockdown and overexpression of candidate dyslexia susceptibility gene homolog Dcdc2 in the rat. *Neuroscience*, 152(3), 723-733.
- Burgess, S. R., Hecht, S. A., & Lonigan, C. J. (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading research quarterly*, 37(4), 408-426.
- Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *The journal of educational research*, 90(5), 269-277.
- Cao, F., Bitan, T., & Booth, J. R. (2008). Effective brain connectivity in children with reading difficulties during phonological processing. *Brain and language*, 107(2), 91-101.
- Cardon, L. R., Dilalla, L. F., Plomin, R., DeFries, J., & Fulker, D. W. (1990). Genetic correlations between reading performance and IQ in the Colorado Adoption Project. *Intelligence*, 14(2), 245-257.
- Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, 461(7266), 983-986.
- Carroll, J. M., Solity, J., & Shapiro, L. R. (2016). Predicting dyslexia using prereading skills: the role of sensorimotor and cognitive abilities. *Journal of child psychology and psychiatry*, 57(6), 750-758.
- Carter, J. C., Lanham, D. C., Cutting, L. E., Clements-Stephens, A. M., Chen, X., Hadzipasic, M., . . . Kaufmann, W. E. (2009). A dual DTI approach to analyzing white matter in children with dyslexia. *Psychiatry research: Neuroimaging*, 172(3), 215-219.
- Casanova, M. F., Christensen, J. D., Giedd, J., Rumsey, J. M., Garver, D. L., & Postel, G. C. (2005). Magnetic resonance imaging study of brain asymmetries in dyslexic patients. *Journal of child neurology*, 20(10), 842-847.
- Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, 47(2), 149-180.
- Catani, M., Allin, M. P., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., & Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proceedings of the national academy of sciences*, 104(43), 17163-17168.
- Catani, M., Howard, R. J., Pajevic, S., & Jones, D. K. (2002). Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage*, 17(1), 77-94.

- Catani, M., & Jones, D. K. (2005). Perisylvian language networks of the human brain. *Annals of neurology*, 57(1), 8-16.
- Catts, H. W., Adlof, S. M., Hogan, T. P., & Ellis Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of speech, language, and hearing research*, 48(6), 1378-1396.
- Catts, H. W., McIlraith, A., Bridges, M. S., & Nielsen, D. C. (2017). Viewing a phonological deficit within a multifactorial model of dyslexia. *Reading and writing*, 30(3), 613-629.
- Catts, H. W., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of speech, language, and hearing research*, 49(2), 278-293.
- Centanni, T. M., Norton, E. S., Park, A., Beach, S. D., Halverson, K., Ozernov-Palchik, O., . . . Gabrieli, J. D. (2018). Early development of letter specialization in left fusiform is associated with better word reading and smaller fusiform face area. *Developmental science*. e12658.
- Che, A., Truong, D. T., Fitch, R. H., & LoTurco, J. J. (2015). Mutation of the Dyslexia-Associated Gene *Dcdc2* Enhances Glutamatergic Synaptic Transmission Between Layer 4 Neurons in Mouse Neocortex. *Cerebral cortex*, 26(9), 3705-3718.
- Chiang, M.-C., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Hickie, I., Toga, A. W., . . . Thompson, P. M. (2011). Genetics of white matter development: a DTI study of 705 twins and their siblings aged 12 to 29. *Neuroimage*, 54(3), 2308-2317.
- Choi, J., Jeong, B., Rohan, M. L., Polcari, A. M., & Teicher, M. H. (2009). Preliminary evidence for white matter tract abnormalities in young adults exposed to parental verbal abuse. *Biological psychiatry*, 65(3), 227-234.
- Chou, T. L., Booth, J. R., Bitan, T., Burman, D. D., Bigio, J. D., Cone, N. E., . . . Cao, F. (2006). Developmental and skill effects on the neural correlates of semantic processing to visually presented words. *Human brain mapping*, 27(11), 915-924.
- Clark, C. A. H., T., Specht, K., Narr, K. L., Manis, F. R., Toga, A. W., & Hugdahl, K. . (2014). Neuroanatomical precursors of dyslexia identified from pre-reading through age to 11. *Brain*, 137(12), 3136-3141.
- Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, 1(3), 98-101.
- Cohen, L., Lehericy, S., Chochon, F., Lemer C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125(Pt 5), 1054-1069.
- Coley, R. J. (2002). An Uneven Start: Indicators of Inequality in School Readiness. Policy Information Report.
- Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *Quarterly journal of experimental psychology*, 35(3), 469-495.
- Compton, D. L., Defries, J. C., & Olson, R. K. (2001). Are RAN-and phonological awareness-deficits additive in children with reading disabilities? *Dyslexia*, 7(3), 125-149.



- Compton, D. L., Davis, C. J., DeFries, J. C., Gayan, J., & Olson, R. K. (2001). Genetic and environmental influences on reading and RAN: An overview of results from the Colorado twin study. In Conference proceedings of the dyslexia research foundation conference in extraordinary brain series: Time, fluency, and developmental dyslexia (pp. 277-303).
- Conant, L. L., Liebenthal, E., Desai, A., & Binder, J. R. (2014). fMRI of phonemic perception and its relationship to reading development in elementary-to middle-school-age children. *Neuroimage*, 89, 192-202.
- Constable, R. T., Skudlarski, P., Mencl, E., Pugh, K. R., Fulbright, R. K., Lacadie, C., . . . Shaywitz, B. A. (1998). Quantifying and comparing region-of-interest activation patterns in functional brain MR imaging: methodology considerations. *Magnetic resonance imaging*, 16(3), 289-300.
- Cope, N., Eicher, J. D., Meng, H., Gibson, C. J., Hager, K., Lacadie, C., . . . Gruen, J. R. (2012). Variants in the DYX2 locus are associated with altered brain activation in reading-related brain regions in subjects with reading disability. *NeuroImage*, 63(1), 148-156.
- Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *cortex*, 45(1), 119-130.
- Cronin, V. S. (2013). RAN and double-deficit theory. *Journal of learning disabilities*, 46(2), 182-190.
- Cummine, J., Chouinard, B., Szepesvari, E., & Georgiou, G. (2015). An examination of the rapid automatized naming–reading relationship using functional magnetic resonance imaging. *Neuroscience*, 305, 49-66.
- Cunillera, T., Càmarà, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., . . . Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, 48(3), 541-553.
- Cutler, A. (2012). Native listening: Language experience and the recognition of spoken words. Mit Press.
- Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech*, 40(2), 141-201.
- Cutting, L., Clements, A., Courtney, S., Rimrodt, S., Schafer, J., Bisesi, J., . . . Pugh, K. (2006). Differential components of sentence comprehension: Beyond single word reading and memory. *Neuroimage*, 29(2), 429-438.
- Czamara, D., Bruder, J., Becker, J., Bartling, J., Hoffmann, P., Ludwig, K. U., . . . Schulte-Körne, G. (2011). Association of a rare variant with mismatch negativity in a region between KIAA0319 and DCDC2 in dyslexia. *Behavior genetics*, 41(1), 110-119.
- D'Angiulli, A., Herdman, A., Stapells, D., & Hertzman, C. (2008). Children's event-related potentials of auditory selective attention vary with their socioeconomic status. *Neuropsychology*, 22(3), 293-300.
- David, D., Wade-Woolley, L., Kirby, J., & Smithrim, K. (2007). Rhythm and reading development in school-age children: a longitudinal study. *Journal of research in reading*, 30(2), 169-183.

- de Bree, E., Wijnen, F., & Gerrits, E. (2010). Non-word repetition and literacy in Dutch children at-risk of dyslexia and children with SLI: Results of the follow-up study. *Dyslexia*, 16(1), 36-44.
- Dehaene, S. (2009). Reading in the brain: The new science of how we read. *Penguin*.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences*, 15(6), 254-262.
- Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nature reviews neuroscience*, 16(4), 234-244.
- Dehaene, S., Le Clec, H. G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, 13(3), 321-325.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., . . . Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, 330(6009), 1359-1364.
- Dellatolas, G., Watier, L., Le Normand, M. T., Lubart, T., & Chevrie-Muller, C. (2009). Rhythm reproduction in kindergarten, reading performance at second grade, and developmental dyslexia theories. *Archives of clinical neuropsychology*, 24(6), 555-563.
- Demb, J. B., Boynton, G. M., & Heeger, D. J. (1998). Functional magnetic resonance imaging of early visual pathways in dyslexia. *Journal of neuroscience*, 18(17), 6939-6951.
- Denckla, M. B. (1972). Color-naming defects in dyslexic boys. *Cortex*, 8(2), 164-176.
- Denckla, M. B., & Rudel, R. G. (1976). Rapid 'automatized' naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14(4), 471-479.
- Dennis, M. Y., Paracchini, S., Scerri, T. S., Prokunina-Olsson, L., Knight, J. C., Wade-Martins, R., . . . Monaco, A. P. (2009). A common variant associated with dyslexia reduces expression of the KIAA0319 gene. *PLoS genetics*, 5(3), e1000436.
- Denton, C., & Hocker, J. (2006). *Responsive reading instruction: Flexible intervention for struggling readers in the early grades*: Sopris west educational services.
- Deutsch, G. K., Dougherty, R. F., Bammer, R., Siok, W. T., Gabrieli, J. D., & Wandell, B. (2005). Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. *Cortex*, 41(3), 354-363.
- Devlin, J. T., Matthews, P. M., & Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of cognitive neuroscience*, 15(1), 71-84.
- Díaz-Uriarte, R., & De Andres, S. A. (2006). Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, 7(1), 3.

- Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal Modulations in Speech and Music. *Neuroscience & biobehavioral reviews*.
- Dobbins, I. G., Foley, H., Schacter, D. L., & Wagner, A. D. (2002). Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron*, 35(5), 989-996.
- Dufford, A. J., & Kim, P. (2017). Family Income, Cumulative Risk Exposure, and White Matter Structure in Middle Childhood. *Frontiers in human neuroscience*, 11, 547.
- Dunn, L. M., & Dunn, D.M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)*. San Antonio, TX: The Psychological Corporation, Inc.
- Eckert, M. A., Leonard, C. M., Richards, T. L., Aylward, E. H., Thomson, J., & Berninger, V. W. (2003). Anatomical correlates of dyslexia: frontal and cerebellar findings. *Brain*, 126(2), 482-494.
- Eckert, M. A., Leonard, C. M., Wilke, M., Eckert, M., Richards, T., Richards, A., & Berninger, V. (2005). Anatomical signatures of dyslexia in children: unique information from manual and voxel based morphometry brain measures. *Cortex*, 41(3), 304-315.
- Eckert, M. A., Leonard, C.M., Richards, T.L., Aylward, E.H., Thomson, J., & Berninger, V.W. (2003). Anatomical correlates of dyslexia: Frontal and cerebellar findings. *Brain*, 126, 482-494.
- Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., & Zeffiro, T. A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature*, 382(6586), 66.
- Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*: CRC press.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific studies of reading*, 9(2), 167-188.
- Eicher, J. D., & Gruen, J. R. (2013). Imaging-genetics in dyslexia: connecting risk genetic variants to brain neuroimaging and ultimately to reading impairments. *Molecular genetics and metabolism*, 110(3), 201-212.
- Elbro, C. (1996). Early linguistic abilities and reading development: A review and hypothesis *Reading & Writing: An interdisciplinary journal*, 8, 453--485.
- Eliez, S., Rumsey, J. M., Giedd, J. N., Schmitt, J. E., Patwardhan, A. J., & Reiss, A. L. (2000). Morphological alteration of temporal lobe gray matter in dyslexia: an MRI study. *Journal of child psychology and psychiatry*, 41(05), 637-644.
- Elliott, J. G., & Grigorenko, E. L. . (2014). The dyslexia debate. *Cambridge University Press*, 14.
- Emery, B. (2010). Regulation of oligodendrocyte differentiation and myelination. *Science*, 330(6005), 779-782.
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, 12(2), 121.
- Epelbaum, S., Pinel, P., Gaillard, R., Delmaire, C., Perrin, M., Dupont, S., . . . Cohen, L. (2008). Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. *Cortex*, 44(8), 962-974.

- Escoffier, N., Sheng, D. Y. J., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta psychologica*, 135(1), 12-16.
- Feinstein, L. (2003). Inequality in the early cognitive development of British children in the 1970 cohort. *Economica*, 70(277), 73-97.
- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement Gap in Reading Is Present as Early as First Grade and Persists through Adolescence. *The Journal of pediatrics*, 167(5), 1121-1125. e1122.
- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: empirical evidence for a definition of dyslexia. *Psychological science*, 21(1), 93-101.
- Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., . . . Ronfani, L. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in human neuroscience*, 8, 392.
- Fletcher, J. (2010). The prosody of speech: Timing and rhythm. *The Handbook of Phonetic Sciences, Second Edition*, 521-602.
- Fletcher, J. M., Foorman, B. R., Boudousquie, A., Barnes, M. A., Schatschneider, C., & Francis, D. J. (2002). Assessment of reading and learning disabilities a research-based intervention-oriented approach. *Journal of school psychology*, 40(1), 27-63.
- Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., . . . Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of educational psychology*, 86(1), 6.
- Flynn, L. J., Zheng, X., & Swanson, H. L. (2012). Instructing struggling older readers: a selective meta-analysis of intervention research. *Learning disabilities research & practice*, 27(1), 21-32.
- Franceschini, S., Gori, S., Ruffino, M., Viola, S., Molteni, M., & Facoetti, A. (2013). Action video games make dyslexic children read better. *Current biology*, 23(6), 462-466.
- Francis, D. J., Shaywitz, S.E., Stuebing, K.K., Shaywitz, B.A., & Fletcher, J.M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of educational psychology*, 88(1), 3-17.
- Friederici, A. D., Meyer, M., & von Cramon, D. Y. (2000). Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain and language*, 74(2), 289-300.
- Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: the syntax-semantic interface. *Brain research*, 1146, 50-58.
- Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain and language*, 86(1), 23-39.
- Friend, A., DeFries, J. C., & Olson, R. K. (2008). Parental education moderates genetic influences on reading disability. *Psychological science*, 19(11), 1124-1130.

- Frost, S. J., Mencl, W. E., Sandak, R., Moore, D. L., Rueckl, J. G., Katz, L., . . . Pugh, K. R. (2005). A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud. *Neuroreport*, 16(6), 621-624.
- Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of speech, language, and hearing research*, 58(3), 934-945.
- Gabel, L. A., Marin, I., LoTurco, J. J., Che, A., C.Murphy, Manglani, M., & Kass, S. (2011). Mutation of the dyslexia-associated gene *Dcdc2* impairs LTM and visuo-spatial performance in mice. *Genes, brain and behavior*, 10(8), 868-875.
- Gabrieli, J. D. (2016). The promise of educational neuroscience: Comment on Bowers (2016).
- Gabrieli, J. D., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron*, 85(1), 11-26.
- Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the national Academy of Sciences*, 95(3), 906-913.
- 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.
- Gascon, G., & Goodglass, H. (1970). Reading retardation and the information content of stimuli in paired associate learning. *Cortex*, 6(4), 417-429.
- Georgiou, G. K., Papadopoulos, T. C., Zarouna, E., & Parrila, R. (2012). Are auditory and visual processing deficits related to developmental dyslexia? *Dyslexia*, 18(2), 110-129.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., . . . Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature neuroscience*, 2(10), 861-863.
- Gilger, J. W., Hanebuth, E., Smith, S. D., & Pennington, B. F. (1996). Differential risk for developmental reading disorders in the offspring of compensated versus noncompensated parents. *Reading and writing*, 8(5), 407-417.
- Gold, B. T., Powell, D. K., Xuan, L., Jiang, Y., & Hardy, P. A. (2007). Speed of lexical decision correlates with diffusion anisotropy in left parietal and frontal white matter: evidence from diffusion tensor imaging. *Neuropsychologia*, 45(11), 2439-2446.
- Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental science*, 18(4), 635-644.
- Gorsuch, R. L. (1983). *Factor Analysis*. Hillsdale, NJ: Lawrence.
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in cognitive sciences*, 15(1), 3-10.

- Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nature reviews neuroscience*, 16(1), 43-54.
- Goswami, U., Fosker, T., Huss, M., Mead, N., & Szucs, D. (2011). Rise time and formant transition duration in the discrimination of speech sounds: the Ba-Wa distinction in developmental dyslexia. *Developmental science*, 14(1), 34-43.
- Goswami, U., Huss, M., Mead, N., Fosker, T., & Verney, J. P. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: significant longitudinal relations with word reading and reading comprehension. *Cortex*, 49(5), 1363-1376.
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory phonology*, 4(1).
- Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *Neuroimage*, 47(4), 1894-1903.
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of neuroscience*, 29(23), 7540-7548.
- Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2009). Neural systems for reading aloud: a multiparametric approach. *Cerebral cortex*, 20(8), 1799-1815.
- Green, J. A., & Mykytyn, K. (2014). Neuronal primary cilia: an underappreciated signaling and sensory organelle in the brain. *Neuropsychopharmacology*, 39(1), 244.
- Greenberg, S. (2006). A multi-tier framework for understanding spoken language. *Listening to speech: An auditory perspective*, 411-433.
- Grigorenko, E. L. (2004). Genetic bases of developmental dyslexia: A capsule review of heritability estimates.
- Grigorenko, E. L. (2006). Learning disabilities in juvenile offenders. *Child and adolescent psychiatric clinics of North America*, 15(2), 353-371.
- Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American statistician*, 61(2), 139-147.
- Grossman, M., Cooke, A., DeVita, C., Chen, W., Moore, P., Detre, J., . . . Gee, J. (2002). Sentence processing strategies in healthy seniors with poor comprehension: an fMRI study. *Brain and language*, 80(3), 296-313.
- Gullick, M. M., Demir-Lira, Ö. E., & Booth, J. R. (2016). Reading skill-fractional anisotropy relationships in visuospatial tracts diverge depending on socioeconomic status. *Developmental science*, 19(4), 673-685.
- Guttorm, T. K., Leppanen, P. H., Hamalainen, J. A., Eklund, K. M., & Lyytinen, H. J. (2010). Newborn event-related potentials predict poorer pre-reading skills in children at risk for dyslexia. *Journal of learning disabilities*, 43(5), 391-401.
- Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M. (2016). Music and dyslexia: a new musical training method to improve reading and related disorders. *Frontiers in psychology*, 7, 26.
- Hackman, D. A., & Farah, M.J. (2009). Socioeconomic status and the developing brain. *Trends in cognitive science*, 13, 65-73.

- Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, 11(9), 651-659.
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1995). Multivariable data analysis reading: USA: Prentice Hall International Editions.
- Hakvoort, B., de Bree, E., van der Leij, A., Maassen, B., van Setten, E., Maurits, N., & van Zuijen, T. L. (2016). The role of categorical speech perception and phonological processing in familial risk children with and without dyslexia. *Journal of speech, language, and hearing research*, 59(6), 1448-1460.
- Hamalainen, J., Leppanen, P. H., Torppa, M., Muller, K., & Lyytinen, H. (2005). Detection of sound rise time by adults with dyslexia. *Brain and language*, 94(1), 32-42.
- Hamalainen, J. A., Rupp, A., Soltesz, F., Szucs, D., & Goswami, U. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. *NeuroImage*, 59(3), 2952-2961.
- Hanson, J. L., Hair, N., Shen, D. G., Shi, F., Gilmore, J. H., Wolfe, B. L., & Pollak, S. D. (2013). Family poverty affects the rate of human infant brain growth. *PloS one*, 8(12), e80954.
- Hanson, M. J., Miller, A. D., Diamond, K., Odom, S., Lieber, J., Butera, G., . . . Fleming, K. (2011). Neighborhood community risk influences on preschool children's development and school readiness. *Infants & Young Children*, 24(1), 87-100.
- Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.
- Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., & Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit? *Journal of speech, language, and hearing research*, 52(6), 1510-1529.
- He, Q., Xue, G., Chen, C., Chen, C., Lu, Z.-L., & Dong, Q. (2013). Decoding the neuroanatomical basis of reading ability: a multivoxel morphometric study. *The journal of neuroscience*, 33(31), 12835-12843.
- He, W., Goodkind, D., & Kowal, P. (2016). An aging world: 2015. International population reports, P95/16-1.
- Heath, S. M., Bihop, D. V., Bloor, K. E., Boyle, G. L., Fletcher, J., Hogben, J. H., . . . Yeong, S. H. (2014). A spotlight on preschool: The influence of family factors on children's early literacy skills. *PloS one*, 9(4), e95255.
- Heim, S., Grande, M., Pape-Neumann, J., van Ermingen, M., Meffert, E., Grabowska, A., . . . Amunts, K. (2010). Interaction of phonological awareness and 'magnocellular' processing during normal and dyslexic reading: behavioural and fMRI investigations. *Dyslexia*, 16(3), 258-282.
- Heim, S., Pape-Neumann, J., van Ermingen-Marbach, M., Brinkhaus, M., & Grande, M. (2015). Shared vs. specific brain activation changes in dyslexia

- after training of phonology, attention, or reading. *Brain structure and function*, 220(4), 2191-2207.
- Heim, S., Tschierse, J., Amunts, K., Wilms, M., Vossel, S., Willmes, K., ... & Huber, W. (2008). Cognitive subtypes of dyslexia. *Acta neurobiologiae experimentalis*, 68(1), 73.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews neuroscience*, 8(5), 393.
- Hiebert, E. H., & Taylor, B. M. (2000). Beginning reading instruction: research on early interventions. *Handbook of reading research*, 3, 455-482.
- Hoeft, F., Hernandez, A., McMillon, G., Taylor-Hill, H., Martindale, J. L., Meyler, A., . . . Just, M. A. (2006). Neural basis of dyslexia: a comparison between dyslexic and nondyslexic children equated for reading ability. *The Journal of neuroscience*, 26(42), 10700-10708.
- Hoeft, F. M., B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., ... & Gabrieli, J. D. (2011). neural systems predicting long-term outcome in dyslexia. *Proceedings of the national Academy of Sciences*, 108(1), 361-366.
- Holliman, A., Wood, C., & Sheehy, K. (2010). The contribution of sensitivity to speech rhythm and non-speech rhythm to early reading development. *Educational Psychology: An international journal of experimental educational psychology*, 30(3), 247-267.
- Holliman, A. J., Wood, C., & Sheehy, K. (2008). Sensitivity to speech rhythm explains individual differences in reading ability independently of phonological awareness. *British journal of developmental psychology*, 26(3), 357-367.
- Hornickel, J., & Kraus, N. (2013). Unstable representation of sound: a biological marker of dyslexia. *Journal of neuroscience*, 33(8), 3500-3504.
- Hosseini, S. M., Black, J. M., Soriano, T., Bugescu, N., Martinez, R., Raman, M. M., . . . Hoeft, F. (2013). Topological properties of large-scale structural brain networks in children with familial risk for reading difficulties. *Neuroimage*, 71, 260-274.
- Howard Jr, J. H., Howard, D. V., Japikse, K. C., & Eden, G. F. (2006). Dyslexics are impaired on implicit higher-order sequence learning, but not on implicit spatial context learning. *Neuropsychologia*, 44(7), 1131-1144.
- Hurwitz, I., Wolff, P. H., Bortnick, B. D., & Kokas, K. (1975). Nonmusical effects of the kodaly music curriculum in primary grade children. *Journal of learning disabilities*, 8(3), 167-174.
- Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex*, 47(6), 674-689.
- Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, 5(3), 299-314.



- Im, K., Raschle, N. M., Smith, S. A., Grant, P. E., & Gaab, N. (2015). Atypical Sulcal Pattern in Children with Developmental Dyslexia and At-Risk Kindergarteners. *Cerebral cortex*, bhu305.
- Jaschke, A. C., Honing, H., & Scherder, E. J. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Frontiers in neuroscience*, 12, 103.
- Jaffe-Dax, S., Raviv, O., Jacoby, N., Loewenstein, Y., & Ahissar, M. (2015). A Computational Model of Implicit Memory Captures Dyslexics' Perceptual Deficits. *The journal of neuroscience*, 35(35), 12116-12126.
- Jednoróg, K., Altarelli, I., Monzalvo, K., Fluss, J., Dubois, J., Billard, C., . . . Ramus, F. (2012). The influence of socioeconomic status on children's brain structure. *PloS one*, 7(8), e42486.
- Jednoróg, K., Gawron, N., Marchewka, A., Heim, S., & Grabowska, A. (2014). Cognitive subtypes of dyslexia are characterized by distinct patterns of grey matter volume. *Brain structure and function*, 219(5), 1697-1707.
- Jenkins, J. R., & O'Connor, R.E. (2002). Early identification and intervention for young children with reading/learning disabilities. In L. D. R. Bradley, & D. Hallahan (Ed.), *Identification of learning disabilities: Research to practice* (pp. 185-250). Mahwah, NJ: Erlbaum.
- Jiménez, J. E., Rodríguez, C., & Ramírez, G. (2009). Spanish developmental dyslexia: Prevalence, cognitive profile, and home literacy experiences. *Journal of experimental child psychology*, 103(2), 167-185.
- Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental cognitive neuroscience*, 1(1), 7-21.
- Johnson, R. T., Yeatman, J. D., Wandell, B. A., Buonocore, M. H., Amaral, D. G., & Nordahl, C. W. (2013). Diffusion properties of major white matter tracts in young, typically developing children
- Joo, S. J., Donnelly, P. M., & Yeatman, J. D. (2017). The causal relationship between dyslexia and motion perception reconsidered. *Scientific reports*, 7(1), 4185.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of educational psychology*, 80(4), 437.
- Just, M. A., Carpenter, P. A., (2013). *Cognitive Processes in Comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Katz, R. B. (1986). Phonological deficiencies in children with reading disability: Evidence from an object-naming task. *Cognition*, 22(3), 225-257.
- Katzir, T., Kim, Y.-S., Wolf, M., Morris, R., & Lovett, M. W. (2008). The varieties of pathways to dysfluent reading comparing subtypes of children with dyslexia at letter, word, and connected text levels of reading. *Journal of learning disabilities*, 41(1), 47-66.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test*: Wiley Online Library.

- Keller, T. A., & Just, M. (2009). Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. *Neuron*, 64(5), 624-631.
- Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. (2008). Longitudinal predictors of word reading development. *Canadian psychology/Psychologie canadienne*, 49(2), 103.
- 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.
- Kirkpatrick, R. M., Legrand, L. N., Iacono, W. G., & McGue, M. (2011). A twin and adoption study of reading achievement: Exploration of shared-environmental and gene-environment-interaction effects. *Learning and individual differences*, 21(4), 368-375.
- Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J. D., 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(2), 493-500.
- Koster, C., Been, P. H., Krikhaar, E. M., Zwarts, F., Diepstra, H. D., & Van Leeuwen, T. H. (2005). Differences at 17 months: Productive language patterns in infants at familial risk for dyslexia and typically developing infants. *Journal of Speech, Language, and hearing research*, 48(2), 426-438.
- Kovelman, I., Norton, E. S., Christodoulou, J. A., Gaab, N., Lieberman, D. A., Triantafyllou, C., ... & Gabrieli, J. D. (2011). Brain basis of phonological awareness for spoken language in children and its disruption in dyslexia. *Cerebral Cortex*, 22(4), 754-764.
- Koyama, M. S., Di Martino, A., Zuo, X. N., Kelly, C., Mennes, M., Jutagir, D. R., Castellanos, F. X., & Milham, M. P. (2011). resting-state functional connectivity indexes reading competence in children and adults. *Journal of neuroscience*, 31(23), 8617-8624.
- Kraft, I., Schreiber, J., Cafiero, R., Metere, R., Schaadt, G., Brauer, J., . . . Wilcke, A. (2016). Predicting early signs of dyslexia at a preliterate age by combining behavioral assessment with structural MRI. *NeuroImage*, 143, 378-386.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature reviews neuroscience*, 11(8), 599-605.
- Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of neuroscience*, 34(36), 11913-11918.
- Kraus, N., & White-Schwoch, T. (2016). Neurobiology of Everyday Communication What Have We Learned From Music? *The Neuroscientist*, 1073858416653593.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience*, 5(11), 831-843.

- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science*, 9(2), F13-F21.
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, cognition and neuroscience*, 31(1), 32-59.
- Kuperberg, G. R., Sitnikova, T., & Lakshmanan, B. M. (2008). Neuroanatomical distinctions within the semantic system during sentence comprehension: evidence from functional magnetic resonance imaging. *NeuroImage*, 40(1), 367-388.
- Landi, N., Frost, S. J., Mencl, W. E., Sandak†, R., & Pugh, K. R. (2013). Neurobiological bases of reading comprehension: insights from neuroimaging studies of word-level and text-level processing in skilled and impaired readers. *Reading & Writing Quarterly*, 29(2), 145-167.
- Langer, N., Peysakhovich, B., Zuk, J., Drottar, M., Sliva, D. D., Smith, S., . . . Gaab, N. (2017). White Matter Alterations in Infants at Risk for Developmental Dyslexia. *Cerebral cortex*, 27(2), 1027-1036.
- Larsen, S., Hammill, D., & Moats, L. (1999). Test of Written Spelling-Revised, IV. *SanAntonio, Texas: Pearson*.
- Lawson, G. M., Duda, J. T., Avants, B. B., Wu, J., & Farah, M. J. (2013). Associations between children's socioeconomic status and prefrontal cortical thickness. *Developmental science*, 16(5), 641-652.
- Laycock, R., Crewther, D. P., Fitzgerald, P. B., & Crewther, S. G. (2009). TMS disruption of V5/MT+ indicates a role for the dorsal stream in word recognition. *Experimental brain research*, 197(1), 69-79.
- Lebel, C., Treit, S., & Beaulieu, C. (2017). A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR in Biomedicine*.
- Lebel, C., Walton, M., Letourneau, N., Giesbrecht, G. F., Kaplan, B. J., & Dewey, D. (2016). Prepartum and postpartum maternal depressive symptoms are related to children's brain structure in preschool. *Biological psychiatry*, 80(11), 859-868.
- Lee, H.-Y., Sie, Y.-S., Chen, S.-C., & Cheng, M.-C. (2015). The music perception performance of children with and without dyslexia in Taiwan. *Psychological reports*, 116(1), 13-22.
- Lefly, D. L., & Pennington, B. F. (2000). Reliability and validity of the adult reading history questionnaire. *Journal of learning disabilities*, 33(3), 286-296.
- Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron*, 72(6), 1080-1090.
- Leppanen, P. H., Hamalainen, J. A., Guttorm, T. K., Eklund, K. M., Salminen, H., Tanskanen, A., Torppa, M., Puolakanaho, A., Richardson, U., Pennala, R., & Lyytinen, H. (2012). Infant brain responses associated with reading-related skills before school and at school age. *leppanen, P. H., Hamalainen,*

- J. A., Guttorm, T. K., Eklund, K. M., Salminen, H., Tanskanen, A., Torppa, M., Puolakanaho, A., Richardson, U., Pennala, R., & Lyytinen, H. (2012). Infant brain responses associated with reading-related skills before school and at school age. *Neurophysiology clinical*, 42(1-2), 35-41.
- Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science*, 20(8), 1040-1048.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, 54(5), 358.
- Linkersdörfer, J., Jurcoane, A., Lindberg, S., Kaiser, J., Hasselhorn, M., Fiebach, C. J., & Lonnemann, J. (2014). The association between gray matter volume and reading proficiency: A longitudinal study of beginning readers. *Journal of cognitive neuroscience*.
- Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., & Fiebach, C. J. (2012). Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: an ALE meta-analysis. *PloS one*, 7(8), e43122.
- Livingstone, M. S., Rosen, G. D., Drislane, F. W., & Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the national Academy of Sciences*, 88, 7943-7947.
- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767-778.
- Lobier, M. A., Peyrin, C., Pichat, C., Le Bas, J.-F., & Valdois, S. (2014). Visual processing of multiple elements in the dyslexic brain: evidence for a superior parietal dysfunction. *Frontiers in human neuroscience*, 8, 479.
- London, J. (2012). *Hearing in time: Psychological aspects of musical meter*: Oxford University Press.
- Lopez-Barroso, D., Catani, M., Ripolles, P., Dell'Acqua, F., Rodriguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc Natl Acad Sci U S A*, 110(32), 13168-13173.
- Lovett, M. W., Frijters, J. C., Wolf, M., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2017). Early intervention for children at risk for reading disabilities: The impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology*, 109(7), 889.
- Lovett, Maureen W., Karen A. Steinbach, and Jan C. Frijters. "Remediating the core deficits of developmental reading disability: A double-deficit perspective." *Journal of learning disabilities* 33, no. 4 (2000): 334-358.
- Lubienski, C., & Lubienski, S. T. (2006). *Charter, private, public schools and academic achievement: New evidence from NAEP mathematics data* (Vol. 16): New York: National Center for the Study of Privatization in Education, Teachers College, Columbia University.
- Luby, J., Belden, A., Botteron, K., Marrus, N., Harms, M. P., Babb, C., . . . Barch, D. (2013). The effects of poverty on childhood brain development: the

- mediating effect of caregiving and stressful life events. *JAMA pediatrics*, 167(12), 1135-1142.
- Ludwig, K. U., Schumacher, J., Schulte-Korne, G., Konig, I. R., Warnke, A., Plume, E., . . . Hoffmann, P. (2008). Investigation of the DCDC2 intron 2 deletion/compound short tandem repeat polymorphism in a large German dyslexia sample. *Psychiatr Genet*, 18(6), 310-312.
- Lui, J. H., Hansen, D. V., & Kriegstein, A. R. (2011). Development and evolution of the human neocortex. *Cell*, 146(1), 18-36.
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of dyslexia*, 53(1), 1-14.
- Maassen, B., Groenen, P., Crul, T., Assman-Hulsmans, C., & Gabreëls, F. (2001). Identification and discrimination of voicing and place-of-articulation in developmental dyslexia. *Clinical linguistics & Phonetics*, 15(4), 319-339.
- Mackey, A. P., Finn, A. S., Leonard, J. A., Jacoby-Senghor, D. S., West, M. R., Gabrieli, C. F., & Gabrieli, J. D. (2015). Neuroanatomical correlates of the income-achievement gap. *Psychological science*, 26(6), 925-933.
- Mackey, A. P., Whitaker, K. J., & Bunge, S. A. (2012). Experience-dependent plasticity in white matter microstructure: reasoning training alters structural connectivity. *Frontiers in neuroanatomy*, 6.
- Magidson, J., & Vermunt, J. (2002). Latent class models for clustering: A comparison with K-means. *Canadian journal of marketing research*, 20(1), 36-43.
- Magidson, J., & Vermunt, J. K. (2001). Latent Class Factor and Cluster Models, Bi-Plots, and Related Graphical Displays. *Sociological methodology*, 31(1), 223-264.
- Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A Meta-analysis of Functional Neuroimaging Studies of Dyslexia. *Annals of the New York Academy of Sciences*, 1145(1), 237-259.
- Makris, N., Kennedy, D. N., McInerney, S., Sorensen, A. G., Wang, R., Caviness, V. S., & Pandya, D. N. (2005). Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cerebral cortex*, 15(6), 854-869.
- Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Petersen, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of experimental child psychology*, 66(2), 211-235.
- Manis, F. R., Seidenberg, M. S., & Doi, L. M. (1999). See Dick RAN: Rapid naming and the longitudinal prediction of reading subskills in first and second graders. *Scientific Studies of reading*, 3(2), 129-157.
- Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of development dyslexia. *Cognition*, 58(2), 157-195.
- Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of learning disabilities*, 17(10), 592-599.

- Marino, C., Meng, H., Mascheretti, S., Rusconi, M., Cope, N., Giorda, R., . . . Gruen, J. R. (2012). DCDC2 genetic variants and susceptibility to developmental dyslexia. *Psychiatric genetics*, 22(1), 25.
- Marino, C., Scifo, P., Della Rosa, P. A., Mascheretti, S., Facoetti, A., Lorusso, M. L., . . . Perani, D. (2014). The DCDC2/intron 2 deletion and white matter disorganization: focus on developmental dyslexia. *Cortex*, 57, 227-243.
- 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.
- 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.
- Martino, J., Brogna, C., Robles, S. G., Vergani, F., & Duffau, H. (2010). Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data. *Cortex*, 46(5), 691-699.
- Martino, J., Hamer, P. C. D. W., Berger, M. S., Lawton, M. T., Arnold, C. M., de Lucas, E. M., & Duffau, H. (2013). Analysis of the subcomponents and cortical terminations of the perisylvian superior longitudinal fasciculus: a fiber dissection and DTI tractography study. *Brain structure and function*, 218(1), 105-121.
- Mascheretti, S., Bureau, A., Battaglia, M., Simone, D., Quadrelli, E., Croteau, J., & Marino, C. (2013). An assessment of gene-by-environment interactions in developmental dyslexia-related phenotypes. *Genes, Brain and Behavior*, 12(1), 47-55.
- Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E., . . . Brandeis, D. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biological psychiatry*, 66(4), 341-348.
- McArthur, G., Atkinson, C., & Ellis, D. (2009). Atypical brain responses to sounds in children with specific language and reading impairments. *Developmental science*, 12(5), 768-783.
- McBride-Chang, C. (1999). The ABCs of the ABCs: the development of letter name and letter sound knowledge. *Merrill-Palmer Quarterly*, 45, 2.
- Messaoud-Galusi, S., Hazan, V., & Rosen, S. (2011). Investigating speech perception in children with dyslexia: Is there evidence of a consistent deficit in individuals? *Journal of Speech, Language, and hearing research*, 54(6), 1682-1701.
- Meyler, A., Keller, T. A., Cherkassky, V. L., Donghoon, L., Hoeft, F., Whitefield-Gabrieli, S., . . . Just, M. A. (2007). Brain activation during sentence comprehension among good and poor readers. *Cerebral cortex*, 17, 2780-2787.
- Milne, R. D., Syngienotis, A., Jackson, G., & Corballis, M. C. (2002). Mixed lateralization of phonological assembly in developmental dyslexia. *Neurocase*, 8(3), 209-209.

- Misra, M., Katzir, T., Wolf, M., & Poldrack, R. A. (2004). Neural systems for rapid automatized naming in skilled readers: Unraveling the RAN-reading relationship. *Scientific studies of reading*, 8(3), 241-256.
- Misra, P., Goswami, A., & Pandav, C. S. (2004). A study of the perception, communication and coverage of pulse polio immunization programme in a Delhi slum. *Indian J Public health*, 48(4), 216-217.
- Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding? *Journal of experimental child psychology*, 64(2), 199-231.
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: a meta-analysis of print exposure from infancy to early adulthood. *Psychological bulletin*, 137(2), 267.
- Moll, K., Loff, A., & Snowling, M. J. (2013). Cognitive endophenotypes of dyslexia. *Scientific studies of reading*, 17(6), 385-397.
- Monzalvo, K., Fluss, J., Billard, C., Dehaene, S., & Dehaene-Lambertz, G. (2012). Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status. *Neuroimage*, 61(1), 258-274.
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant behavior and development*, 16(4), 495-500.
- Moreno, S., Friesen, D., & Bialystok, E. (2011). Effect of musical training on promoting preliteracy skills: preliminary causal evidence. *Music perception*, 29, 165-172.
- Mori, S., Crain, B. J., Chacko, V. P., & Van Zijl, P. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of neurology*, 45(2), 265-269.
- Moritz, C., Sasha, Y., Papadelis, G., Thomson, J., & Wolf, M. (2012). Links between early rhythm skills, musical training, and phonological awareness. *Reading and writing*.
- Morris, R. D., Lovett, M. W., Wolf, M., Sevcik, R. A., Steinbach, K. A., Frijters, J. C., & Shapiro, M. B. (2012). Multiple-component remediation for developmental reading disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *Journal of learning disabilities*, 45(2), 99-127.
- Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., . . . Shaywitz, B. A. (1998). Subtypes of reading disability: variability around a phonological core. *Journal of educational psychology*, 90(3), 347.
- Morrison, F. J., & Manis, F. R. (1982). Cognitive processes and reading disability: A critique and proposal *Verbal processes in children* (pp. 59-93): Springer.
- Mount, C. W., & Monje, M. (2017). Wrapped to adapt: experience-dependent myelination. *Neuron*, 95(4), 743-756.
- 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.
- Myers, C. A., Vandermosten, M., Farris, E. A., Hancock, R., Gimenez, P., Black, J. M. . . . Hoeft, F (2014). Structural changes in white matter are uniquely

- related to children's reading development. *Psychological science*, 25(10), 1870-1883.
- Natalie, H., & Noble, K. G. (2014). Socioeconomic status and structural brain development. *Frontiers in neuroscience*, 8, 276.
- Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading research quarterly*, 32(2), 154-167.
- 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.
- Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-analysis of executive components of working memory. *Cerebral cortex*, 23(2), 264-282.
- Needle, J. L., Fawcett, A. J., & Nicolson, R. I. (2006). Balance and dyslexia: An investigation of adults' abilities. *European journal of cognitive psychology*, 18(6), 909-936.
- Newbury, D., Paracchini, S., Scerri, T., Winchester, L., Addis, L., Richardson, A. J., . . . Monaco, A. (2011). Investigation of dyslexia and SLI risk variants in reading-and language-impaired subjects. *Behavior genetics*, 41(1), 90-104.
- Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: reuniting the developmental disorders? *Trends in neurosciences*, 30(4), 135-141.
- Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in neurosciences*, 24(9), 508-511.
- Niogi, S. N., & McCandliss, B. D. (2006). Left lateralized white matter microstructure accounts for individual differences in reading ability and disability. *Neuropsychologia*, 44(11), 2178-2188.
- Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., . . . Libiger, O. (2015). Family income, parental education and brain structure in children and adolescents. *Nature neuroscience*, 18(5), 773-778.
- Noble, K. G., Houston, S. M., Kan, E., & Sowell, E. R. (2012). Neural correlates of socioeconomic status in the developing human brain. *Developmental science*, 15(4), 516-527.
- Noble, K. G., Korgaonkar, M. S., Grieve, S. M., & Brickman, A. M. (2013). Higher education is an age-independent predictor of white matter integrity and cognitive control in late adolescence. *Developmental science*, 16(5), 653-664.
- Noble, K. G., & McCandliss, B. D. (2005). Reading development and impairment: behavioral, social, and neurobiological factors. *Journal of developmental & behavioral pediatrics*, 26(5), 370-378.
- Noble, K. G., Tottenham, N., & Casey, B. (2005). Neuroscience perspectives on disparities in school readiness and cognitive achievement. *The future of children*, 15(1), 71-89.
- Noble, K. G., Wolmetz, M. E., Ochs, L. G., Farah, M. J., & McCandliss, B. D. (2006). Brain-behavior relationships in reading acquisition are modulated by socioeconomic factors. *Developmental science*, 9(6), 642-654.



- Nolan, F., & Asu, E. L. (2009). The pairwise variability index and coexisting rhythms in language. *Phonetica*, 66(1-2), 64-77.
- Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: a metaphor? Philosophical transactions of the royal society B, 369(1658), 20130396.
- Noordenbos, M., Segers, E., Serniclaes, W., Mitterer, H., & Verhoeven, L. (2012). Allophonic mode of speech perception in Dutch children at risk for dyslexia: A longitudinal study. *Research in developmental disabilities*, 33(5), 1469-1483.
- Norton, E. S., Beach, S. D., & Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Current opinion in neurobiology*, 30, 73-78.
- Norton, E. S., Black, J. M., Stanley, L. M., Tanaka, H., Gabrieli, J. D., Sawyer, C., & Hoeft, F. (2014). Functional neuroanatomical evidence for the double-deficit hypothesis of developmental dyslexia. *Neuropsychologia*, 61, 235-246.
- Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual review of psychology*, 63, 427-452.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling*, 14(4), 535-569.
- O'Brien, B. A., Wolf, M., & Lovett, M. W. (2012). A taxometric investigation of developmental dyslexia subtypes. *Dyslexia*, 18(1), 16-39.
- 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.
- 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.
- Olson, R. K., Keenan, J. M., Byrne, B., & Samuelsson, S. (2014). Why do children differ in their development of reading and related skills? *Scientific studies of reading*, 18(1), 38-54.
- 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.
- Orton, S. T. (1925). Word-blindness in school children. *Archives of neurology & psychiatry*, 14(5), 581-615.
- Ou, X., Andres, A., Cleves, M. A., Pivik, R., Snow, J. H., Ding, Z., & Badger, T. M. (2014). Sex-specific association between infant diet and white matter integrity in 8-y-old children. *Pediatric research*, 76(6), 535-543.
- Overy, K. (2003). Dyslexia and music. From timing deficits to musical intervention. *Annals of the New York Academy of Sciences*, 999, 497-505.
- Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, 25(1), 46-59.
- Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D., & Gaab, N. (2017). Longitudinal stability of pre-reading skill profiles

- of kindergarten children: implications for early screening and theories of reading. *Developmental science*, 20(5).
- Ozernov-Palchik, O., Yu, X., Wang, Y., & Gaab, N. (2016). Lessons to be learned: how a comprehensive neurobiological framework of atypical reading development can inform educational practice. *Current opinion in behavioral sciences*, 10, 45-58.
- Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the ‘dyslexia paradox’: reading brain and behavior for early markers of developmental dyslexia. *Wiley Interdisciplinary Reviews: Cognitive Science*, 7(2), 156-176.
- Pammer, K. (2014). Brain Mechanisms and Reading Remediation: More Questions Than Answers. *Scientifica*, 2014.
- Papadopoulos, T. C., Georgiou, G. K., & Kendeou, P. (2009). Investigating the double-deficit hypothesis in Greek: Findings from a longitudinal study. *Journal of learning disabilities*, 42(6), 528-547.
- Papanicolaou, A. C., Simos, P. G., Breier, J. I., Fletcher, J. M., Foorman, B. R., Francis, D., . . . Davis, R. N. (2003). Brain mechanisms for reading in children with and without dyslexia: a review of studies of normal development and plasticity. *Dev Neuropsychology*, 24(2-3), 593-612.
- Paracchini, S., Thomas, A., Castro, S., Lai, C., Paramasivam, M., Wang, Y., . . . Monaco, A. P. (2006). The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Human molecular genetics*, 15(10), 1659-1666.
- Paris, S. G., & Oka, E. R. (1989). Strategies for comprehending text and coping with reading difficulties. *Learning disability quarterly*, 12(1), 32-42.
- Patel, A. D. (2008a). *Music, language, and the brain*: Oxford university press.
- Patel, A. D. (2008b). Science & music: talk of the tone. *Nature*, 453(7196), 726-727.
- Paulesu, E., Danelli, L., & Berlingeri, M. (2014). Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Frontiers in human neuroscience*, 8, 830.
- Pavlikis, A. E., Noble, K., Pavlikis, S. G., Ali, N., & Frank, Y. (2015). Brain imaging and electrophysiology biomarkers: is there a role in poverty and education outcome research? *Pediatric neurology*, 52(4), 383-388.
- Pennington, B. F., & Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annual review of psychology*, 60, 283-306.
- Pennington, B. F., Gilger, J. W., Olson, R. K., & DeFries, J. C. (1992). The External Validity of Age-Versus IQ-Discrepancy Definitions of Reading Disability Lessons From a Twin Study. *Journal of learning disabilities*, 25(9), 562-573.
- Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child development*, 72(3), 816-833.
- Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., . . . Willcutt, E. G. (2012). Individual prediction of dyslexia by single versus multiple deficit models. *Journal of abnormal psychology*, 121(1), 212.

- Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly* (1982-), 283-319.
- Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. (2011). Human voice recognition depends on language ability. *Science*, 333(6042), 595.
- Perrachione, T. K., Del Tufo, S. N., Winter, R., Murtagh, J., Cyr, A., Chang, P., ... & Gabrieli, J. D. (2016). Dysfunction of rapid neural adaptation in dyslexia. *Neuron*, 92(6), 1383-1397.
- Peschansky, V. J., Burbridge, T. J., Volz, A. J., Fiondella, C., Wissner-Gross, Z., Galaburda, A. M., . . . Rosen, G. D. (2010). The effect of variation in expression of the candidate dyslexia susceptibility gene homolog Kiaa0319 on neuronal migration and dendritic morphology in the rat. *Cereb Cortex*, 20(4), 884-897.
- Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet*, 379(9830), 1997-2007.
- Peterson, R. L., & Pennington, B. F. (2015). Developmental Dyslexia. *Annual review of clinical psychology*, 11, 283-307.
- Petrill, S. A., Deater-Deckard, K., Thompson, L. A., De Thorne, L. S., & Schatschneider, C. (2006). Reading Skills in Early Readers Genetic and Shared Environmental Influences. *Journal of Learning Disabilities*, 39(1), 48-55.
- Peynircioglu, Z. F., Durgunoglu, A. Y., & Úney-Küsefog˘lu, B. (2002). Phonological awareness and musical aptitude. *Journal of Research in reading*, 25(1), 68-80.
- Phillips, B. M., Clancy-Menchetti, J., & Lonigan, C. J. (2008). Successful phonological awareness instruction with preschool children: Lessons from the classroom. *Topics in early childhood special education*, 28(1), 3-17.
- Piai, V., Anderson, K. L., Lin, J. J., Dewar, C., Parvizi, J., Dronkers, N. F., & Knight, R. T. (2016). Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. *Proceedings of the National Academy of Sciences*, 201603312.
- Pierpaoli, C., & Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic resonance in Medicine*, 36(6), 893-906.
- Pinel, P., Fauchereau, F., Moreno, A., Barbot, A., Lathrop, M., Zelenika, D., . . . Dehaene, S. (2012). Genetic variants of FOXP2 and KIAA0319/TTRAP/THEM2 locus are associated with altered brain activation in distinct language-related regions. *Journal of neuroscience*, 32(3), 817-825.
- Piras, F., Cherubini, A., Caltagirone, C., & Spalletta, G. (2011). Education mediates microstructural changes in bilateral hippocampus. *Human brain mapping*, 32(2), 282-289.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage*, 10, 15-35.

- Poulsen, M. (in press). The challenge of early identification of later reading difficulties. Perspectives on language and literacy.
- Povel, D.J., & Essens, P. (1985). Perception of temporal patterns. *Music perception: An interdisciplinary journal*, 2(4), 411-440.
- Powers, S. J., Wang, Y., Beach, S. D., Sideridis, G. D., & Gaab, N. (2016). Examining the relationship between home literacy environment and neural correlates of phonological processing in beginning readers with and without a familial risk for dyslexia: an fMRI study. *Annals of dyslexia*, 1-24.
- Preston, J. L., Molfese, P. J., Frost, S. J., Mencl, W. E., Fulbright, R. K., Hoeft, F., . . . Pugh, K. R. (2016). Print-Speech Convergence Predicts Future Reading Outcomes in Early Readers. *Psychological science*, 27(1), 75-84.
- Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *The journal of neuroscience*, 35(7), 3276-3284.
- Price, C. J. (1998). The functional anatomy of word comprehension and production. *Trends in cognitive sciences*, 2(8), 281-288.
- Price, C. J. D., J. T. . (2011). The interactive account of the ventral occipitotemporal contributions to reading. *Trends in cognitive sciences*, 15(6), 246-253.
- Pugh, K. R., Landi, N., Preston, J. L., Mencl, W. E., Austin, A. C., Sibley, D., . . . Constable, R. T. (2013). The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain and language*, 125(2), 173-183.
- Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., . . . Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability(developmental dyslexia). *Mental retardation and developmental disabilities research reviews*, 6(3), 207-213.
- Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., . . . Shaywitz, B. A. (2001). Neurobiological studies of reading and reading disability. *Journal of commun disorders*, 34(6), 479-492.
- Pugh, K. R., Shaywitz, B.A., Shaywitz, S.E., Constable, R.T., Skudlarski< P., Fulbright, R.K. (1996). cerebral organization of component processes in reading. *Brain*, 119(4), 1221-1238.
- Raberg, T., & Wimmer, H. (2003). On the automaticity/cerebellar deficit hypothesis of dyslexia: balancing and continuous rapid naming in dyslexic and ADHD children. *Neuropsychologia*, 41(11), 1493-1497.
- Raizada, R. D., & Kishiyama, M. M. (2010). Effects of socioeconomic status on brain development, and how cognitive neuroscience may contribute to levelling the playing field. *Frontiers in human neuroscience*, 4(3).
- Raizada, R. D., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *Neuroimage*, 40(3), 1392-1401.
- Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current opinion in neurobiology*, 13(2), 212-218.

- Ramus, F., Pidgeon, E., & Frith, U. (2003). The relationship between motor control and phonology in dyslexic children. *Journal of child psychology and psychiatry*, 44(5), 712-722.
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, 126(4), 841-865.
- Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *The quarterly journal of experimental psychology (Hove)*, 61(1), 129-141.
- Raschle, N., Zuk, J., Ortiz-Mantilla, S., Sliva, D. D., Franceschi, A., Grant, P. E., . . . Gaab, N. (2012). Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines. *Annals of the New York Academy of Sciences*, 1252(1), 43-50.
- Raschle, N. M., Becker, B. L. C., Smith, S., Fehlbauer, L. V., Wang, Y., & Gaab, N. (2015). Investigating the influences of language delay and/or familial risk for dyslexia on brain structure in 5-year-olds. *Cerebral cortex*, 27(1), 764-776.
- Raschle, N. M., Chang, M., & Gaab, N. (2011). Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage*, 57(3), 742-749.
- Raschle, N. M., Stering, P. L., Meissner, S. N., & Gaab, N. (2013). Altered neuronal response during rapid auditory processing and its relation to phonological processing in prereading children at familial risk for dyslexia. *Cerebral cortex*, bht104.
- Raschle, N. M., Zuk, J., & Gaab, N. (2012). Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. *Proceedings of the national Academy of Sciences*, 109(6), 2156-2161.
- Rauschecker, A. M., Bowen, R. F., Perry, L. M., Kevan, A. M., Dougherty, R. F., & Wandell, B. A. (2011). Visual feature-tolerance in the reading network. *Neuron*, 71(5), 941-953.
- Rauschecker, A. M., Deutsch, G. K., Ben-Shachar, M., Schwartzman, A., Perry, L. M., & Dougherty, R. F. (2009). Reading impairment in a patient with missing arcuate fasciculus. *Neuropsychologia*, 47(1), 180-194.
- Rautenberg, I. (2015). The effects of musical training on the decoding skills of German-speaking primary school children. *Journal of research in reading*, 38(1), 1-17.
- Reardon, S., & Portilla, X. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, 2(3), 1-18.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither opportunity*, 91-116.
- Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human brain mapping*, 30(10), 3299-3308.

- Richlan, F., Kronbichler, M., & Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies. *Human brain mapping, 34*(11), 3055-3065.
- Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage, 56*(3), 1735-1742.
- Richlan, F., Kronbichler, M., & Wimmer, H. (2012). Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies. *Hum Brain Mapp.*
- Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature neuroscience, 11*(4), 426.
- Robertson, C., & Salter, W. (1997). The phonological awareness test. *East Moline, IL: LinguiSystems.*
- Robertson, E. K., Joanisse, M. F., Desroches, A. S., & Ng, S. (2009). Categorical speech perception deficits distinguish language and reading impairments in children. *Developmental Science, 12*(5), 753-767.
- Rochelle, K. S., & Talcott, J. B. (2006). Impaired balance in developmental dyslexia? A meta-analysis of the contending evidence. *Journal of Child Psychology and Psychiatry, 47*(11), 1159-1166.
- Roelofs, A. (2005). From Popper to Lakatos: A case for cumulative computational modeling. *Twenty-first century psycholinguistics: Four cornerstones, 313*, 330.
- Rollins, N. K., Vachha, B., Srinivasan, P., Chia, J., Pickering, J., Hughes, C. W., & Gimi, B. (2009). Simple developmental dyslexia in children: alterations in diffusion-tensor metrics of white matter tracts at 3 T. *Radiology, 251*(3), 882-891.
- Romeo, R. R., Christodoulou, J. A., Halverson, K. K., Murtagh, J., Cyr, A. B., Schimmel, C., . . . Gabrieli, J. D. (2017). Socioeconomic Status and Reading Disability: Neuroanatomy and Plasticity in Response to Intervention. *Cerebral Cortex, 1*-16.
- Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics, 31*(3-4), 509-527.
- Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science, 323*(5916), 951-953.
- Sarubbo, S., De Benedictis, A., Maldonado, I. L., Basso, G., & Duffau, H. (2013). Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. *Brain structure and function, 218*(1), 21-37.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., . . . Abel, S. (2008). Ventral and dorsal pathways for language. *Proceedings of the national Academy of Sciences, 105*(46), 18035-18040.

- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the national Academy of Sciences*, 112(29), 8987-8992.
- Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., . . . Gabrieli, J. D. (2013). Tracking the roots of reading ability: white matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children. *Journal of neuroscience*, 33(33), 13251-13258.
- Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., . . . Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature neuroscience*, 19(9), 1250.
- Scarborough, H. (1990). Very early language deficits in dyslexic children. *Child development*, 61, 1728-1743.
- Scarborough, H. (1998). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of dyslexia*, 48(1), 115-136.
- Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. *Journal of educational psychology*, 81, 101-108.
- Scarborough, H. S. (1991). antecedents to reading disabilities: Preschool language development and literacy experiences of children from dyslexic families. *Reading disabilities*, 31-45.
- Scarborough, H. S., Dobrich, W., & Hager, M. (1991). Preschool literacy experience and later reading achievement. *Journal of learning disabilities*, 24(8), 508-511.
- Scerri, T. S., Darki, F., Newbury, D. F., Whitehouse, A. J., Peyrard-Janvid, M., Matsson, H., . . . Paracchini, S. (2012). The dyslexia candidate locus on 2p12 is associated with general cognitive ability and white matter structure. *PLoS One*, 7(11), e50321.
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, 26(1), 3-8.
- Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (2002). Relationship of Rapid Automatized Naming and Phonological Awareness in Early Reading Development Implications for the Double-Deficit Hypothesis. *Journal of learning disabilities*, 35(3), 245-256.
- Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004b). Kindergarten Prediction of Reading Skills: A Longitudinal Comparative Analysis. *Journal of educational psychology*, 96(2), 265-282.
- Schlaffke, L., Leemans, A., Schweizer, L. M., Ocklenburg, S., & Schmidt-Wilcke, T. (2017). Learning Morse Code Alters Microstructural Properties in the Inferior Longitudinal Fasciculus: A DTI Study. *Frontiers in Human Neuroscience*, 11(383).
- Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading. *Annual review of neuroscience*, 30, 475-503.

- Scholz, J., Klein, M. C., Behrens, T. E., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature neuroscience*, 12(11), 1370-1371.
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341-349.
- Schön, D., & Tillmann, B. (2015). Short-and long-term rhythmic interventions: perspectives for language rehabilitation. *Annals of the New York Academy of Sciences*, 1337(1), 32-39.
- Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., & Kronbichler, M. (2014). Resting-state and task-based functional brain connectivity in developmental dyslexia. *Cerebral cortex*, bhu184.
- Sénéchal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five year longitudinal study. *Child development*, 73(2), 445-460.
- Sénéchal, M., LeFevre, J.-A., Hudson, E., & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of educational psychology*, 88, 520-536.
- Senechal, M., leFevre, J. . (2002). Parental involvement in development of children's reading skills: a five year longitudinal study. *Child Development*, 73(2), 445-460.
- Senechel, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of child languags*, 24(1), 123-138.
- Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Demonet, J.-F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of speech, language, and hearing research*, 44(2), 384-399.
- Shankweiler, D., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of experimental psychology: Human learning and memory*, 5(6), 531.
- Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of aquisition. *Issues in education: Contributions from educational psychology*, 1, 1-57.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., . . . Lyon, G. R. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological psychiatry*, 52(2), 101-110.
- Shaywitz, B. A., Skudlarski, P., Holahan, J. M., Marchione, K. E., Constable, R. T., Fulbright, R. K., . . . Shaywitz, S. E. (2007). Age-related changes in reading systems of dyslexic children. *Annals of neurology*, 61(4), 363-370.
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., . . . Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics*, 104(6), 1351-1359.



- Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual review of psychology*, 59, 451-475.
- Shinoura, N., Midorikawa, A., Onodera, T., Tsukada, M., Yamada, R., Tabei, Y., . . . Yagi, K. (2013). Damage to the left ventral, arcuate fasciculus and superior longitudinal fasciculus-related pathways induces deficits in object naming, phonological language function and writing, respectively. *International Journal of Neuroscience*, 123(7), 494-502.
- Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of learning disabilities*, 22(8), 469-478.
- Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of learning disabilities*, 25(10), 618-629.
- Siegel, L. S. (2006). Perspectives on dyslexia. *Paediatr Child Health*, 11(9), 581-587.
- Siegel, L. S., & Linder, B. A. (1984). Short-term memory processes in children with reading and arithmetic learning disabilities. *Developmental Psychology*, 20(2), 200.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research*, 75(3), 417-453.
- Skoe, E., Krizman, J., & Kraus, N. (2013). The impoverished brain: disparities in maternal education affect the neural response to sound. *Journal of neuroscience*, 33(44), 17221-17231.
- Skottun, B. C. (2000). The magnocellular deficit theory of dyslexia: the evidence from contrast sensitivity. *Vision research*, 40(1), 111-127.
- Snowling, M. J. (2000). Language and literacy skills: who is at risk and why. *Speech and language impairments in children: Causes, characteristics, intervention and outcome*, 245-259.
- Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child development*, 74(2), 358-373.
- Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological bulletin*, 142(5), 498.
- Song, S.-K., Sun, S.-W., Ju, W.-K., Lin, S.-J., Cross, A. H., & Neufeld, A. H. (2003). Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *Neuroimage*, 20(3), 1714-1722.
- Song, S.-K., Sun, S.-W., Ramsbottom, M. J., Chang, C., Russell, J., & Cross, A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage*, 17(3), 1429-1436.
- Song, S.-K., Yoshino, J., Le, T. Q., Lin, S.-J., Sun, S.-W., Cross, A. H., & Armstrong, R. C. (2005). Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage*, 26(1), 132-140.
- Specht, K., Hugdahl, K., Ofte, S., Nygard, M., Bjornerud, A., Plante, E., & Helland, T. (2009). Brain activation on pre-reading tasks reveals at-risk status for

- dyslexia in 6-year-old children. *Scandinavian journal of psychology*, 50(1), 79-91.
- Spector, J. E. (2005). Instability of double-deficit subtypes among at-risk first grade readers. *Reading psychology*, 26(3), 285-312.
- Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and writing*, 28(4), 467-490.
- Spencer, R. M., & Ivry, R. B. (2013). Cerebellum and timing *Handbook of the cerebellum and cerebellar disorders* (pp. 1201-1219): Springer.
- Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of cognitive neuroscience*, 4(3), 232-243.
- Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of educational psychology*, 86(2), 221.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading research quarterly*, 360-407.
- Stanovich, K. E., & Siegel, L.S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of phonological-core variable-difference model. *Journal of educational psychology*, 86(1), 24-53.
- Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Relation between early reading acquisition and word decoding with and without context: A longitudinal study of first-grade children. *Journal of educational psychology*, 76(4), 668.
- Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of educational psychology*, 89(1), 114.
- Steady, L. M., Kirby, J. R., Parrila, R., & Compton, D. L. (2014). Classification of Double Deficit Groups Across Time: An Analysis of Group Stability From Kindergarten to Second Grade. *Scientific Studies of Reading*, 18(4), 255-273.
- Stein, J. (2014). Dyslexia: the role of vision and visual attention. *Current developmental disorders reports*, 1(4), 267-280.
- Stein, J. (2018). What is developmental dyslexia? *Brain sciences*, 8(2), 26.
- Stine, R. (1989). An introduction to bootstrap methods: Examples and ideas. *Sociological methods & research*, 18(2-3), 243-291.
- Stoodley, C. J. (2014). Distinct regions of the cerebellum show gray matter decreases in autism, ADHD, and developmental dyslexia. *Frontiers in systems neuroscience*, 8, 92.
- Stoodley, C. J., Harrison, E. P., & Stein, J. F. (2006). Implicit motor learning deficits in dyslexic adults. *Neuropsychologia*, 44(5), 795-798.
- Stoodley, C. J., & Stein, J. F. (2013). Cerebellar function in developmental dyslexia. *The cerebellum*, 12(2), 267-276.
- Stothard, S., Hulme, C., Clarke, P., Bamby, P., & Snowling, M. (2010). YARC York Assessment of Reading for Comprehension (Secondary).

- Strawn, J. R., Hamm, L., Fitzgerald, D. A., Fitzgerald, K. D., Monk, C. S., & Phan, K. L. (2015). Neurostructural abnormalities in pediatric anxiety disorders. *Journal of anxiety disorders*, 32, 81-88.
- Swaminathan, S., & Schellenberg, E. G. (2016). Music Training *Cognitive Training* (pp. 137-144): Springer.
- Szalkowski, C. E., Fiondella, C. G., Galaburda, A. M., Rosen, G. D., Loturco, J. J., & Fitch, R. H. (2012). Neocortical disruption and behavioral impairments in rats following in utero RNAi of candidate dyslexia risk gene Kiaa0319. *International Journal of Developmental Neuroscience*, 30(4), 293-302.
- Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *cortex*, 49(10), 2674-2688.
- Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *TRENDS in neurosciences*, 29(7), 382-390.
- Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the national Academy of Sciences of the United States of America*, 102(24), 8781-8785.
- Tanaka, H., Black, J. M., Hulme, C., Stanley, L. M., Kesler, S. R., Whitfield-Gabrieli, S., . . . Hoeft, F. (2011). The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychological science*, 22(11), 1442-1451.
- Tanaka, H. B., J.M., Hulme, C., Stanley, L. M., Kesler, S.R., Whitfield-Gabrieli, S...&Hoeft, F. (2011). The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychological science*, 22(11), 1442-1451.
- Tang, Y.-Y., Lu, Q., Geng, X., Stein, E. A., Yang, Y., & Posner, M. I. (2010). Short-term meditation induces white matter changes in the anterior cingulate. *Proceedings of the national Academy of Sciences*, 107(35), 15649-15652.
- Taylor, J., Roehrig, A., Hensler, B. S., Connor, C., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science*, 328(5977), 512-514.
- Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: an fMRI study. *Neuroreport*, 12(2), 299-307.
- Testolin, A., Stoianov, I., & Zorzi, M. (2017). Letter perception emerges from unsupervised deep learning and recycling of natural image features. *Nature human behaviour*, 1(9), 657.
- The Mathworks Inc. (2007). *MATLAB 7.5*. Natick, MA: The Mathworks Inc.
- Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., & Dehaene, S. (2014). Learning to read improves the structure of the arcuate fasciculus. *Cerebral cortex*, 24(4), 989-995.
- Thomason, M. E., Burrows, B. E., Gabrieli, J. D., & Glover, G. H. (2005). Breath holding reveals differences in fMRI BOLD signal in children and adults. *Neuroimage*, 25(3), 824-837.

- Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *Journal of Physiology - Paris*, 102(1-3), 120-129.
- Tierney, A., & Kraus, N. (2013a). The ability to move to a beat is linked to the consistency of neural responses to sound. *Journal of neuroscience*, 33(38), 14981-14988.
- Tierney, A., & Kraus, N. (2013b). Music training for the development of reading skills *Progress in brain research* (Vol. 207, pp. 209-241): Elsevier.
- Tierney, A. T., & Kraus, N. (2013c). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and language*, 124(3), 225-231.
- Tisdall, M. D., Hess, A. T., Reuter, M., Meintjes, E. M., Fischl, B., & van der Kouwe, A. J. (2012). Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magn Reson Med*, 68(2), 389-399.
- Tofighi, D., & Enders, C. K. (2008). *Identifying the correct number of classes in growth mixture models*.
- Tomalski, P., & Johnson, M. H. (2010). The effects of early adversity on the adult and developing brain. *Current opinion in psychiatry*, 23(3), 233-238.
- Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning disabilities research & practice*, 15(1), 55-64.
- Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. *Word recognition in beginning literacy*, 161-188.
- Torgesen, J. K., Rashotte, C. A., & Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333-355). Parkton, MD: York Press.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *TOWRE: Test of Word Reading Efficiency*. Austin, TX: PRO-ED, Inc.
- 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.
- Torppa, M., Tolvanen, A., Poikkeus, A.-M., Eklund, K., Lerkkanen, M.-K., Leskinen, E., & Lyytinen, H. (2007). Reading development subtypes and their early characteristics. *Annals of dyslexia*, 57(1), 3-32.
- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100(1), 73-99.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of cognitive neuroscience*, 21(10), 1934-1945.
- Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory phonology*, 4(1).

- Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nature neuroscience*, 6(7), 767-773.
- Tyszkka, J. M., Readhead, C., Bearer, E. L., Pautler, R. G., & Jacobs, R. E. (2006). Statistical diffusion tensor histology reveals regional dysmyelination effects in the shiverer mouse mutant. *Neuroimage*, 29(4), 1058-1065.
- Ullman, M. T., & Pullman, M. Y. (2015). A compensatory role for declarative memory in neurodevelopmental disorders. *Neuroscience & biobehavioral reviews*, 51, 205-222.
- Vaessen, A., Gerretsen, P., & Blomert, L. (2009). Naming problems do not reflect a second independent core deficit in dyslexia: Double deficits explored. *Journal of experimental child psychology*, 103(2), 202-221.
- Valdois, S., Peyrin, C., Lassus-Sangosse, D., Lallier, M., Demonet, J.-F., & Kandel, S. (2014). Dyslexia in a French-Spanish bilingual girl: behavioural and neural modulations following a visual attention span intervention. *Cortex*, 53, 120-145.
- van Bergen, E., de Jong, P. F., Maassen, B., & van der Leij, A. (2014). The effect of parents' literacy skills and children's preliteracy skills on the risk of dyslexia. *Journal of abnormal child psychology*, 42(7), 1187-1200.
- Van der Lely, H., Gardner, H., McClelland, A. G. R., & Froud, K. E. (2007). *Grammar and Phonology Screening Test:(GAPS):* DLDCN. com.
- Van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., . . . Martin, E. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage*, 47(4), 1940-1949.
- van der Weij, B., Pearce, M. T., & Honing, H. (2017). A probabilistic model of meter perception: Simulating enculturation. *Frontiers in psychology*, 8, 824.
- van Viersen, S., de Bree, E. H., Zee, M., Maassen, B., van der Leij, A., & de Jong, P. F. (2018). Pathways Into Literacy: The Role of Early Oral Language Abilities and Family Risk for Dyslexia. *Psychological science*.
- Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., & Ghesquière, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences*, 107(23), 10389-10394.
- Vandermosten, M., Boets, B., Poelmans, H., Sunaert, S., Wouters, J., & Ghesquiere, P. (2012). A tractography study in dyslexia: neuroanatomic correlates of orthographic, phonological and speech processing. *Brain*, 135(Pt 3), 935-948.
- Vandermosten, M., Boets, B., Wouters, J., & Ghesquiere, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience & Biobehavioral Reviews*, 36(6), 1532-1552.
- Vandermosten, M., Vanderauwera, J., Theys, C., De Vos, A., Vanvooren, S., Sunaert, S., . . . Ghesquière, P. (2015). A DTI tractography study in pre-readers at risk for dyslexia. *Developmental cognitive neuroscience*, 14, 8-15.
- Vaughn, S., Wexler, J., Leroux, A., Roberts, G., Denton, C., Barth, A., & Fletcher, J. (2012). Effects of intensive reading intervention for eighth-grade students

- with persistently inadequate response to intervention. *Journal of learning disabilities*, 45(6), 515-525.
- Vellutino, F. R., Fletcher, J. M., Snowling, M.J., & Scanlon, D. M. . (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of child psychology and psychiatry*, 45(1), 2-40.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of educational psychology*, 88(4), 601.
- Vermunt, J. K., & Magidson, J. (2013). Technical guide for Latent GOLD 5.0: Basic, advanced, and syntax. *Statistical Innovations Inc., Belmont, MA*.
- Vicari, S., Finzi, A., Menghini, D., Marotta, L., Baldi, S., & Petrosini, L. (2005). Do children with developmental dyslexia have an implicit learning deficit? *Journal of Neurology, neurosurgery & psychiatry*, 76(10), 1392-1397.
- Vicari, S., Marotta, L., Menghini, D., Molinari, M., & Petrosini, L. (2003). Implicit learning deficit in children with developmental dyslexia. *Neuropsychologia*, 41(1), 108-114.
- Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., . . . Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage*, 30(4), 1414-1432.
- Vogel, A. C., Church, J. A., Power, J. D., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2013). Functional network architecture of reading-related regions across development. *Brain and language*, 125(2), 231-243.
- Vukovic, R. K., & Siegel, L. S. (2006). The Double-Deficit Hypothesis A Comprehensive Analysis of the Evidence. *Journal of learning disabilities*, 39(1), 25-47.
- Waber, D. P., Weiler, M. D., Wolff, P. H., Bellinger, D., Marcus, D. J., Ariel, R., . . . Wypij, D. (2001). Processing of Rapid Auditory Stimuli in School-Age Children Referred for Evaluation of Learning Disorders. *Child development*, 72(1), 37-49.
- Wagner, A. D., Pare-Blagoiev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31(2), 329-338.
- Wagner, R. K., & Barker, T. A. (1994). The development of orthographic processing ability *The varieties of orthographic knowledge* (pp. 243-276): Springer.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *CTOPP: Comprehensive test of phonological processing*: Pro-ed.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., . . . Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental psychology*, 33(3), 468.

- Wakana, S., Caprihan, A., Panzenboeck, M. M., Fallon, J. H., Perry, M., Gollub, R. L., . . . Mori, S. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*, 36(3), 630-644.
- Wandell, B. A., Rauschecker, A. M., & Yeatman, J. D. (2012). Learning to see words. *Annual review of psychology*, 63, 31-53.
- Wang, Y., Mauer, M. V., Raney, T., Peysakhovich, B., Becker, B. L., Sliva, D. D., & Gaab, N. (2016). Development of Tract-Specific White Matter Pathways During Early Reading Development in At-Risk Children and Typical Controls. *Cerebral cortex*, 1, 17.
- Wanzek, J., & Vaughn, S. (2007). based implications from extensive early reading interventions. *School psychology review*, 36(4), 541.
- Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of educational research*, 0034654313477212.
- Warmington, M., & Hulme, C. (2012). Phoneme awareness, visual-verbal paired-associate learning, and rapid automatized naming as predictors of individual differences in reading ability. *Scientific studies of reading*, 16(1), 45-62.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-WISC-IV*: Psychological Corporation.
- Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of research in reading*, 29(3), 288-303.
- Wheeler, B., Torchiano, M., & Torchiano, M. M. (2016). Package 'lmPerm'. *R Package Version*, 1.1-2.
- White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., & Ramus, F. (2006). The role of sensorimotor impairments in dyslexia: A multiple case study of dyslexic children. *Developmental science*, 9(3), 237-255.
- Wiederholt, J., & Bryant, B. (2012). *Gray oral reading test-(GORT-4)*. Austin, TX: Pro-ed.
- Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S.-E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and language*, 144, 26-34.
- Wilcke, A., Ligges, C., Burkhardt, J., Alexander, M., Wolf, C., Quente, E., . . . Kirsten, H. (2012). Imaging genetics of FOXP2 in dyslexia. *European journal of human genetics*, 20(2), 224-229.
- Wilson, A. M., & Lesaux, N. K. (2001). Persistence of phonological processing deficits in college students with dyslexia who have age-appropriate reading skills. *Journal of learn disabilities*, 34(5), 394-400.
- Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of educational psychology*, 92(4), 668.
- Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading research quarterly*, 123-141.
- Wolf, B. M., & Bowers, P. G. (1999). The double deficit hypothesis for the developmental dyslexias. *Journal of educational psychology*, 91, 1-24.

- Wolf, M. (2008). *Proust and the squid: the story and science of the reading brain*. Cambridge: Icon.
- Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading: A longitudinal study in average and impaired readers. *Child development*, 988-1000.
- Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of educational psychology*, 91(3), 415.
- Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of learning disabilities*, 33(4), 387-407.
- Wolf, M., & Denckla, M. B. (2005). *RAN/RAS: Rapid Automatized Naming and Rapid Alternating*. Austin, TX: PRO-ED, Inc.
- Wolf, M., Goldberg O'Rourke, A., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and writing: An interdisciplinary journal*, 15, 43-72.
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific studies of seading*, 5(3), 211-239.
- Wolff, P. H. (2002). Timing precision and rhythm in developmental dyslexia. *Reading and writing*, 15(1-2), 179-206.
- Wolff, P. H., Michel, G. F., Ovrut, M., & Drake, C. (1990). Rate and timing precision of motor coordination in developmental dyslexia. *Developmental psychology*, 26(3), 349.
- Wolff, U. (2010). Subgrouping of readers based on performance measures: a latent profile analysis. *Reading and writing*, 23(2), 209-238.
- Wong, F. C., Chandrasekaran, B., Garibaldi, K., & Wong, P. C. (2011). White matter anisotropy in the ventral language pathway predicts sound-to-word learning success. *Journal of neuroscience*, 31(24), 8780-8785.
- Wood, C. (2006). Metrical stress sensitivity in young children and its relationship to phonological awareness and reading. *Journal of research in reading*, 29(3), 270-287.
- Woodcock, R. (1998). Woodcock Reading Mastery Tests-Revised (WRMT-R/NU). 1998. *Circle Pines, Minnesota: American Guidance Service*, 5.
- Woodcock, R. (2011). Woodcock reading mastery test (WRMT-III): San Antonio: Pearson.
- Yamada, Y., Stevens, C., Dow, M., Harn, B. A., Chard, D. J., & Neville, H. J. (2011). Emergence of the neural network for reading in five-year-old beginning readers of different levels of pre-literacy abilities: An fMRI study. *Neuroimage*, 57(3), 704-713.
- Yap, R. L., & Leij, A. v. d. (1994). Testing the automatization deficit hypothesis of dyslexia via a dual-task paradigm. *Journal of learning disabilities*, 27(10), 660-665.
- 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.



- Yeatman, J. D., Dougherty, R. F., Ben-Shachar, M., & Wandell, B. A. (2012). Development of white matter and reading skills. *Proceedings of the national Academy of Sciences of the United States of America*, 109(44), E3045-3053.
- Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., & Feldman, H. M. (2012b). Tract profiles of white matter properties: automating fiber-tract quantification. *PLoS One*, 7(11), e49790.
- Yeatman, J. D., Rauschecker, A. M., & Wandell, B. A. (2013). Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections. *Brain and language*, 125, 146-155.
- Yu, X., Raney, T., Perdue, M. V., Zuk, J., Ozernov-Palchik, O., Becker, B. L., ... & Gaab, N. (2018). Emergence of the neural network underlying phonological processing from the prereading to the emergent reading stage: A longitudinal study. *Human brain mapping*, 39(5), 2047-2063.
- Zeffiro, T., & Eden, G. (2001). The cerebellum and dyslexia: perpetrator or innocent bystander?: Comment from Thomas Zeffiro and Guinevere Eden to Nicolson et al. *Trends in neurosciences*, 24(9), 512-513.
- Zhang, B., Dietrich, U. M., Geng, J.-G., Bicknell, R., Esko, J. D., & Wang, L. (2009). Repulsive axon guidance molecule Slit3 is a novel angiogenic factor. *Blood*, 114(19), 4300-4309.
- Zhang, Y., Whitfield-Gabrieli, S., Christodoulou, J. A., & Gabrieli, J. D. (2013). Atypical balance between occipital and fronto-parietal activation for visual shape extraction in dyslexia. *PloS one*, 8(6), e67331.
- Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the national Academy of Sciences*, 113(19), 5212-5217.
- Zoubrinetzky, R., Bielle, F., & Valdois, S. (2014). New insights on developmental dyslexia subtypes: heterogeneity of mixed reading profiles. *PloS one*, 9(6), e99337.