Early identification of dyslexia and word reading development: Using cognitive and linguistic assessments to predict reading skill growth in early elementary school

Thesis submitted by

Kirk Vanacore

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Child Study and Human Development

Tufts University

May, 2018

Advisor:

Dr. Calvin Gidney

Abstract

Theories of and research on the etiology of dyslexia consistently point to two primary deficits of phonology and automaticity. Previous research has found that performance on early assessments of phonological awareness and automaticity are predictive of reading abilities in early elementary school. This study used secondary data from 161 students who were followed prior to beginning elementary school through the end of second grade. These students were assessed in each grade using a variety of cognitive and linguistic measures. Multi-level models were estimated to assess the growth rates of word identification and to compare growth rates for students who performed poorly on assessments of phonological awareness and automaticity before beginning elementary school. Further analyses were conducted to assess where the availability of curriculum that incorporated explicit phonics instruction moderated the relationship between performance on assessments of phonological awareness and automaticity before beginning elementary school and growth rates of word identification. Students who performed in the bottom quartile of a normed sample on measures of phonological awareness or phonological awareness and automaticity upon or prior to entering kindergarten had significantly lower rates of growth on word identification than their typically developing peers during the scope of the study. The availability of curriculum that incorporated explicit phonics instruction was not predictive of significant differences in the word identification growth rates for students who performed poorly on assessments of phonological awareness and automaticity before beginning elementary school. These findings bolster support for the predictive nature of phonological awareness and automaticity and indicated that deficits in these areas can be detected before a child begins formal education.

Acknowledgments

My deepest gratitude to my wife Kim and my son Lucca,

who tolerated my attempts to cloister myself in order to

complete this project.

Thanks to my committee, Dr. Calvin Gidney, Dr. Sara Johnson,

and Dr. Ola Ozernov-Palchik, as well as Dr. Martha Pott for guidance

throughout this process.

Table of Contents

Introduction

Reading is a cornerstone in American education and a prerequisite for participation in United States society, both economically and politically. Proficient literacy skills are associated with employment in professional occupations and substantially decreases one's likelihood of requiring public assistance (Kutner et al., 2007). In 2015, 64 percent of 8^{th}-grade students and 66 percent of 4^{th}-grade students were not reading proficiently, based upon the National Assessments of Educational Progress (NAEP), and 2015 was not an anomaly (National Center for Education Statistics, 2016). Although long-term trends show that overall reading scores have improved since 1970, they have remained relatively stagnant since 2008. Some academics and policy makers have cautioned against drawing dire conclusions from these statistics, citing NAEP's excessively high standards for proficiency (Ravich, 2014; Snow & Matthews, 2016), but there is consensus that improving the nation's quality of literacy instruction and increasing reading proficiency is essential to foster academic excellence and economic growth (Lonigan, Schatschneider & Westberg, 2008; National Reading Panel, 2000).

Nationally, early literacy intervention programs have been established as a way to bolster literacy rates throughout the country (Daryl, 2017), but the efficacy of these programs is contingent on early and accurate identification of students at risk for reading failure. Studies estimate that up to 20% of the general population is affected by the reading related learning disability commonly known as dyslexia (Eliot & Grigorenko, 2014; Shaywitz, 2003). Dyslexic students enter elementary schools behind their peers in literacy related skills (Schaars, Segers, & Verhoeven, L., 2017), and these disparities can persist throughout elementary education (Danache et al., 2014). Research has demonstrated that risk for dyslexia is already evident in kindergarten, but many studies use family history of dyslexia to gauge risk (Nagamine et al., 2009; Pennington & Lefly, 2001). Although family history is an indicator of potential disorder, it is not a precise diagnostic measurement. Furthermore, schools' methods for identifying dyslexic students are often belated and inaccurate (Denton et al., 2011; Elliot & Grigorenko, 2014).

Local and national governments have considered screening children for dyslexia in early elementary schools in order to ensure the neediest students receive timely and intensive literacy interventions (Decoding Dyslexia Massachusetts, 2017; Rose, 2009). For these policies to be effective, screening practices would need to be systematic and accurate. In the following study, I examine how risk identification before entering kindergarten is predictive of reading development in the first three years of elementary school. I also consider how specific characteristics of curriculum may or may not change the relation between literacy risk and development.

Early Literacy Development

Developing reading-related skills begins long before formal education. Early oral language skills proceed and are predictive of later literacy skills (Bartl-Pokorny et al., 2013; Harlaar et al, 2008). As children begin to the apply their oral language skills to written language, they must connect the sounds that make up words (phonemes) to the written representations of these sounds (graphemes) and then combine them into words (Owens, 2016). Snow and Mathews (2016) refer to this as one of the "constrained skills" of reading because they are limited in scope and should be mastered at a young age. By third grade, children need to have mastered these "constrained skills" – including connecting the 26 letters with 44 phonemes as well as common spelling rules and over 100 sight words – before going on to develop the "unconstrained skills" of vocabulary and comprehension that allow them to become functional readers (p. 58).

The initial constrained skill, which is a prerequisite for all other literacy related skills, is alphabet knowledge – the ability to name the letters of the alphabet and to know the sounds associated with each letter. A meta-analysis of 299 studies related to early literacy development conducted by the *National Early Childhood Panel* found that alphabet knowledge is more predictive of phonic decoding and fluency than oral language skills and lexical knowledge (Lonigan et al., 2008). Furthermore, Lerner and Lonigan (2016) found evidence for a bidirectional relationship between letter name knowledge and phonological awareness: the more letter-name knowledge a child has when entering preschool, the faster their phonological awareness grew, and high levels of phonological awareness when entering preschool predicted faster growth in letter-naming ability.

As children progress through preschool and into early elementary school, they must work towards connecting their knowledge of the alphabet with their understanding of phonemes to build a basis of orthographic information to understand the pattern of letters that form words (Owens, 2016). Explicit instruction of word decoding skills through phonology and orthography is referred to as phonics. The *National Early Childhood Panel* meta-analysis found that phonological skills were significant predictors in explaining the variance in reading outcomes, including fluency and comprehension, and that decoding was highly correlated with reading comprehension and spelling (Lonigan et al., 2008). Students with the learning disability dyslexia struggle specifically with the phonological skills that are foundational to decoding and therefore fundamental to reading (Elliot & Grigorenko, 2014). In order to ensure that these students have access to instruction that improves phonological and decoding skills, we must first identify who these students are.

Definitions and Subtypes of Dyslexia

Many students struggle to master these fundamental skills of reading. For some children, attaining literacy proficiency is hindered by a learning disability often referred to as dyslexia. Dyslexia is a neurological disorder; multiple studies have shown lower activation during reading tasks in areas of the brain associated with phonological and semantic processing, as well as working memory (Beneventi, Tønnessen, Ersland, & Hugdahl, 2010; Christodoulo et al., 2014; Katzir & Pare-Blagoev, 2006). Although neurological studies support a biological etiology for dyslexia, current understanding of the neurological foundations of dyslexia are not robust enough to allow for the development of biological diagnostic assessments (Rutter et al., 2006) and precise methods of the use of cognitive and behavioral assessments for dyslexia are not universally accepted (Elliot & Grigorenko, 2014).

There is ample debate in the literature over specific definitions of dyslexia that allow researchers and clinicians to differentiate who is and who is not dyslexic. Many researchers and clinicians define dyslexia as an unexpected inability to read, in so far as there is a discrepancy between general intelligence and reading ability (Christodoulo et al., 2014, Shaywitz, 2003; Swanson, 2014; Wolf, 2007). Using this criterion, students with high or average intelligence and weak reading related abilities are diagnosed with dyslexia, whereas struggling readers with lower IQs are considered "garden variety poor readers" (Elliot & Grigorenko, 2014, p. 18). Alternatively, some researchers argue that this distinction between struggling readers with low and high IQ is arbitrary and not clinically helpful (Fletcher, Lyon, Fuchs, & Barnes, 2007; Hoskyn & Sawnson, 2000) and that there is little evidence that measure of intelligence predicts reading related skills or response to literacy interventions (Gresham & Vellutino, 2010; Vellutino et al., 2006).

Despite this vigorous debate over the specific definition and etiology of dyslexia, studies consistently find that there are two distinct deficits associated with this disorder: phonological decoding and memory recall automaticity (Groot et al., 2017; Miller et al., 2006; Ozernov-Palchik et al., 2016; Powel et al. 2007; Wolf et al., 2002). The phonological deficit affects a reader's ability to connect graphemes (letters) with phonemes (sounds) in an efficient manner required to decode words (Stanovich, 1988). Phonological awareness is generally measured through tests of elision (ability to manipulate phonemes within a word), blending (ability to combing phonemes to make a word), and phonological working memory (ability to repeat unfamiliar groups of phonemes) (Wagner, Torgesen & Rashotte, 1999). As children grow older, researchers and clinicians also test students' ability to read and phonetically pronounce pseudowords as a measure of proficiency with the grapheme-phoneme connection (Torgesen, Wagner, & Rashotte, 2012). Rapid automatized naming measures the efficient retrieval of information from memory and is assessed using a timed test of identifying letters, numbers, colors, or objects (Wolf & Bowers, 1999).

Phonological deficit. There is broad support for the "phonological deficit hypothesis," which posits that fundamental difficultly experienced by those with dyslexia is an issue of connecting letter patterns to sounds (Stanovich, 1988; Vellutino et al., 2004; Eliott & Grigorenko, 20014). Varvara et al. (2014) found significant differences between dyslexic and typical school age students in phonological fluency, blending, and memory tasks. Farquharson et al. (2014) found that dyslexic children consistently underperformed on assessments of reading words and pseudowords when compared with typical children and children with a specific language impairment. Furthermore, dyslexic participants were less accurate at reading lists that contained words and pseudowords that were phonologically and orthographically alike and more accurate when the words were phonologically and orthographically different. This finding supports the theory that dyslexia primarily disrupts processes of phonological decoding that allow for typical children to distinguish between words even when they are spelt and sound similarly.

Phonological working memory, the ability to remember and recall patterns of phenomes within short periods of time, is considered a “critical building block for reading development” (Gathercole & Baddeley, 1993, pg 259). Dyslexic readers tend to have a specific phonological working memory deficit as opposed to a more global working memory deficit, which is associated with students with ADHD (Maehler & Schuchardt 2016). In a study of dyslexic children’s executive working memory process, Beneventi et al. (2010) asked dyslexic and typical children to identify the first or last “letter-sound” (phoneme) of the one or two-syllable word that the picture presented. Dyslexic children were significantly slower and less accurate in identifying the first or last phoneme of the one or two-syllable word that the picture presented. They specifically showed deficiencies on working memory tasks of identifying phenomes of previously presented pictures.

Automaticity deficit. The “double deficit hypothesis” builds upon the phonological deficit perspective of dyslexia by adding in the possibility of a deficit in rapid automatized naming (Wolf & Bowers, 1999). Slow automatized naming constitutes a combination of inefficiencies in cognitive processes – including attentional, perceptual, and memory retrieval processes – that result in an inability to rapidly connect orthographic codes (letter patterns) with the appropriate phonological (sound) and semantic (meaning) outputs (Wolf, 2000). Although automaticity has been identified as a distinct process from phonological decoding, these deficits are still interrelated (Powell et al., 2007). One possible explanation of this interrelation is that slow rapid naming interferes with an ability to rapidly identify orthographic codes and subsequently hinders a reader's capacity to efficiently identify words (Bowers et al., 1999; Georgiou et al., 2009).

Multiple studies have shown that some dyslexic children and adults display deficits in both phonological and automaticity (double-deficit), whereas others display a deficit in only one area: phonological or automaticity (Groot et al., 2017; Miller et al., 2006; Ozernov-Palchik et al., 2016; Powel et al. 2007; Wolf et al., 2002). The existence of dyslexic students with automaticity deficits and without phonological deficits suggests that automatized naming processes are distinct from phonological processes and each provides a unique contribution to the reading ability. The importance of automaticity may also be substantiated by evidence that individuals with double deficits experience more significant reading difficulties than those with single deficits (Miller et al., 2006).

Both of these deficits, phonological and automaticity, can be distinguished during diagnoses of dyslexia (Groot et al., 2017; Miller et al., 2006; Ozernov-Palchik et al., 2016; Powel et al. 2007; Wolf et al., 2002). Researchers and clinicians will, at times, diagnose students with either phonological deficit, automaticity deficit, or a double deficit (Eliott & Grigorenko, 20014). Yet, questions remain about how measures of phonological awareness and automaticity can be used for the screening of dyslexia at an early age.

Early Identification of Risk for Dyslexia

Early classification of students as dyslexic is a difficult endeavor, but identifying young students who will need additional support as they learn to read is essential. To determine students' levels of risk for reading difficulties many public schools use an approach called response to intervention (RTI), in which a teacher provides levels (referred to as Tiers) of instruction to different students based upon their needs (O'Donnell & Miller, 2011). In this model, the entire class may receive basic instruction (Tier 1), and those identified as not responding to this instruction receive subsequent instruction (Tier 2). Students who require further support may move on to even more intensive instruction (Tier 3) and may potentially be recommended for special education services. Although this strategy might seem to be an appropriate way to identify a student's level of risk and differentiate instruction based upon their needs, it has not been shown to be an effective method for accurately predicting risk of reading difficulty. A review of multiple studies found a broad range of efficacy in the RTI system in identifying students who experienced future reading difficulty, with false positives up to 60% and false negatives up to 50% (Elliot & Grigorenko, 2014). Another study found no difference in performance on reading related measures between students at Tier 2 and Tier 3 (Denton et al., 2011). The lack of measureable differences between these risk groups and the low efficacy of response to intervention in identifying risks suggests that many students are not receiving the level of instruction necessary for addressing their reading difficulties.

Academics often evaluate early risk for dyslexia by identifying children who have a family history of dyslexia, because dyslexia is highly heritable (Dandache et al. 2014; Nagamine et al., 2009; Pennington & Lefly, 2001). Although this standard may be a useful measure of risk in research settings, it is not practical for screening in education because many parents may not know whether their family members have dyslexia. Furthermore, coming from a family with a history of dyslexia predisposes a person to inherit traits associated with dyslexia, but it does not determine whether that person is dyslexic. Furthermore, Pennington and Lefly (2001) found that only 34% of preschoolers with high family risk for dyslexia could be diagnosed with a reading disorder in second grade. They also found that the strongest predictor for reading disability diagnosis in Grade 2 was a student's alphabet knowledge in preschool, not family history. Thompson et al. (2015) found that family history of dyslexia did predict dyslexia at an "acceptable clinical level," but that dyslexia could be predicted with satisfactory accuracy by assessing alphabet knowledge, rapid automatized naming, and phonological awareness before entering elementary school (p. 976). These findings suggests that although family history of dyslexia may be useful for researchers to identify risk in their samples, effective screenings should focus on measuring literacy related skills over heredity.

Studies that attempt to forecast dyslexia based upon pre-literacy skills have consistently found that phonological awareness and rapid automatized naming are the strongest predictors of dyslexia (Carroll et al., 2015; Pennington & Lefty 2001; Thompson et al., 2015). Many of these studies use logistic regression that predicts the probability of a dichotomous outcome where a student is either diagnosed or not diagnosed with dyslexia (Carroll et al., 2015; Thompson et al., 2015). But Pennington and Lefty (2001) found that some students who were both identified as high risk for dyslexia in kindergarten and not formality diagnosed for dyslexia by second grade were still significantly outperformed on reading measures by their peers who were never identified as at risk. This finding indicates the dichotomous outcome based on diagnosis may be insufficient when examining risk for dyslexia.

Alternatively, using a student's growth of reading development may allow for a more complete understanding of a student's reading ability as it presents a student's reading growth trajectory. Clemens et al. (2012) demonstrated that the of using a single word reading measure to monitor students' growth in early elementary school was beneficial because growth on single word reading assessments was more strongly associated with end of the year reading outcomes than other literacy related measures. Considering this finding, along with the limitations of dichotomous outcomes for dyslexia outlined above, establishing how well preliteracy measures, such as phonological awareness and rapid automatized naming, to predict single word reading growth may provide insight into how these measures may be used in screening for dyslexia.

The Debate Over Literacy Curriculum

Literacy curriculum has been a surprisingly contentious topic, not just in the United States, but across English speaking countries. England, Australia, and the United States have all experienced intense political debates over how to teach children to read (Elliot & Grigorenk, 2014; Snyder, 2008). The controversy lies in a philosophical debate over traditional and progressive pedagogy. Traditional reading pedagogy requires overt instruction of the rules of written language (i.e., phonics). Advocates of progressive education often argue against explicit phonics instruction because of its "authoritarian" and "anti-democratic" method of transmitting rules and principles of reading from teacher to student (Anderson, 2000, as cited in Elliot & Grigorenk, 2014, p. 124). As an alternative, the progressive method of teaching reading – referred to as whole language – relies on the assumption that reading is learned through exposure to text just as oral language is acquired through exposure to speech (Goodman, 1992).

This equivalency between written and oral language is contested by research from the fields of linguistics, psychology, and neuroscience. Steven Pinker (1994), preeminent linguistic and cognitive psychologist, explains:

> For although language is an instinct, written language is not. Writing was invented a small number of times in history, and alphabetic writing, where one character corresponds to one sound, seems to have been invented only once. Most societies have lacked written language, and those that have it inherited it or borrowed it from one of the inventors (1994, pg. 189).

Reading is not an innate process because written language is a human creation, not a predestined product of being human. Wolf and Gottwald (2016) note that "we were not born to read;" there is nothing in our genetic code that causes us to read naturally (p. 7). The perspective that literacy is an invented form of communication bolsters the idea that it must be explicitly taught and that an understanding of the mechanics of written language – phonemes connecting to graphemes and graphemes used in an orthographical system – will assist in attaining literacy proficiency.

The National Reading Panel, commissioned by the United States Congress to review research concerning the causes of and possible solutions to low reading proficiency rates, recommended using both phonics instruction and whole language methods of teaching reading (National Reading Panel, 2000). The panel concluded that phonics instruction assisted students in developing skills to decode words, but did not address the issue of fluency. Instead, they recommended repeated reading as an "evidence-based practice" for increasing fluency (National Reading Panel, 2000). Unlike other interventions, which use explicit instruction of reading skills, repeated reading assumes the students will develop these skills by reading and rereading texts out loud. But support for this method of fluency instruction in academic literature is equivocal. Studies have shown that large effects on fluency are limited to texts that have been read repeatedly throughout the intervention and that repeated reading does not have significantly larger effects than continuously reading different texs (Therrien, 2004; O'Connor et al., 2007; Wexler et al. 2010) In an evaluation of the quality of published research supporting these interventions, Chard and colleagues (2009) concluded that repeated reading was not supported by enough quality research to be considered an evidence-based practice.

While examining the differing manifestations of dysfluent reading, Berninger and colleagues (2001) reasoned that for many struggling readers "simply practicing reading may not be sufficient because a faulty brain mechanism for creating direct associations between stimuli and responses interferes with the creation of automatic connections" (p. 388). Wolf and Katzir-Cohen (2001) argue that fluency development depends on more than simply repeated exposure to texts. Fluency involves accuracy and automaticity of "sublexical processes, lexical processes, and their integration in single-word reading and connected text" (p. 219). From this perspective, fluent reading is not attained purely by encountering words; it requires a "breadth and depth" of linguistic knowledge, including knowledge of language's sublexical and morphological structure, as well as word polysemy (Tannenbaum, Torgesen & Wagner, 2006). From this perspective, teaching phonics in conjunction with vocabulary, practiced reading, and comprehension strategies should improve fluency.

Attempting to predict word reading growth requires accounting for the pedagogical methods to which children are exposed. Compton et al. (2010) points out that school level variables, such as teaching methods and teacher experience, may confound attempts to use literacy related measures for dyslexia screening. Therefore school level variables, such as curriculum, should be taken into account when identifying potential predictors of reading development.

The Current Study

Considering debates over how to best identify students who need specific attention for literacy instruction, I sought to identify factors that may be predictive of literacy development. The following analyses were designed to examine the development of single word reading and how this development may differ for students with distinct deficit profiles. Because tracking development of word reading skills is often confounded by environmental factors, such as the curriculum used in each classroom (Compton et al., 2010), I examined whether a school's reported use of a curriculum that incorporates explicit phonics instruction was associated with differences in rates of development for word reading. Through the following analysis, I addressed the following research questions:

1. What is the average growth rate in single word reading over the first three years of school?
2. Does growth of single word reading ability during the first three years of school differ for students who are identified as at risk for dyslexia in kindergarten/preschool when compared to typical students?
3. Does attending a school that uses explicit phonics instruction moderate the relationship between risk for dyslexia and growth of single word reading ability in the first three years of school?
4. Does growth of single word reading ability during the first three years of school differ for students who are identified as at risk for specific dyslexia deficits profiles in kindergarten/preschool when compared to typical students?
5. Does attending a school that uses explicit phonics instruction moderate the relationship between risk for specific dyslexia deficits profiles and growth of single word reading ability in the first three years of school?

Method

Participants

The current study is a secondary analysis of data collected by the Gabrieli Lab at MIT. The Gabrieli Lab recruited participants from 20 schools in Massachusetts. Permission and informed consent letters were sent to the parents of kindergarten and Pre-K children attending these 20 schools. Children who provided verbal assent, and whose parents provided written informed consent, were included in the pre-literacy skills assessment phase of the study. In total, 1,433 English-speaking children were tested at during Pre-K or upon entering of kindergarten (referred to throughout this document as Pre-K). A subset of 189 children were followed longitudinally and were tested at the end of Grade 1 and Grade 2. Children with Year 1 IQ scores below 80 and/or who did not speak fluent English, and/or who were born pre-term were excluded from the longitudinal analysis. The subset of children included in the longitudinal study was selected to maintain a subsample composition representative of the larger sample regarding gender, age, ethnicity/race, and school type.

The current study focuses on 161 children from 18 schools who were part of the longitudinal study. Twenty-eight participants from the longitudinal study were excluded from this analysis due to lack of school level data. Age of participants at Year 1 ranged from 58 months to 77 months, with mean of 67.09 and a standard devotion of 3.96. The analytic sample consisted of 70% Predicting Risk for Dyslexia

The findings of this study are constant with previous research on deficits and prediction of dyslexia. Thompson et al. (2015) tracked students from age three and half to age eight to establish when dyslexia was detectable with accuracy and what variables to utilize in detection. They found that across ages the most constantly reliable were measures of phonological ability, automaticity, and letter knowledge. Carroll et al. (2015) also found that phonological awareness and rapid automatized naming were among the strongest predictors of poor reading, along with alphabet knowledge and short term verbal memory.

Notably, both Thompson et al. (2015) and Carroll et al. (2015) used logistic regression to determine the odds of a student's diagnosis of dyslexia in later grades based on performance on various measure at an early age. These models require dichotomous outcomes, a determination of whether someone is or is not dyslexic based on a specific standard. In both studies the standard for being considered dyslexic was being 1.5 standard deviations below the group which was previously determined to be low risk on literacy outcomes measures.

One unique advantage of the present study is that it focuses on growth trajectory as opposed to a purely binary outcome. Many researchers have argued against the use of simple percentile cut offs on a specific measure for diagnosing learning disabilities because this method has been shown to be ineffective at establishing who may need and respond to interventions (Elliot & Grigorenko, 2014; Pendleton, 2006). In a longitudinal study spanning from the beginning of kindergarten through the end of second grade, Pennington and Lefty (200) found that the students who were identified as high risk for dyslexia in kindergarten but were not diagnosed for dyslexia in second grade, still scored significantly less on reading measures than those who were never identified as at risk. This finding suggests that the cut off method of outcomes is deficient for identifying all children who may need additional intensive reading instruction than the typical student. Furthermore, it suggests that dyslexia is comprised of a multivariate continuum, wherein students can have varying intensities of several different deficits, all which effect reading abilities (Pennington, 2006; Crisp et al., 2011). Although this study began by using percentile cut offs to establish risk groups, multiple assessments of different abilities were used to determine those categories and the outcome focused on in this study was a developmental trajectory, not a potentially arbitrary standard for achievement.

The main findings of this study – that low performance on assessments of phonology, or on assessments of phonology and automaticity together, prior to a child receiving formal reading instruction is predictive of lower word reading growth rates – suggests that these measures could be used as screening tools for dyslexia. There is an important distinction between screening and diagnosis, as screening is meant to flag children who have a high probability of risk, whereas diagnosis is meant to provide a definitive categorization conclusion about whether or not a child has dyslexia (Helland et al. 2011). Alternatively, screening tools can provide parents and educators information about how to effectively allocate resources for helping children become proficient readers .White students, 21% Black or African American students, 6% multiracial, 3% Native American, and 1% Asian, with 14% of students who identified as Hispanic.

Procedures

Cognitive and literacy related assessments were administered by trained research assistants and speech-language pathology students on a one-to-one basis. Pre-literacy skills measures were taken at either the end of preschool or the beginning of kindergarten (Pre-K). Literacy skills assessments were administered at the end of Grade 1 and the end of Grade 2 (Year 3).

Measures

Longitudinal outcome measures (Pre-K, Grade 1 and Grade 2)

Word identification (Word ID)**.** The Woodcock Reading Mastery Tests – Revised/Normative Update (WRMT-R/NU; Woodcock, 1998) was administered. The Word ID subtest assesses single word reading skills. The child reads aloud single words of increasing difficulty. The raw Word ID score, which is the number of words that the student read during the subtest, was used in the following analyses.

Age. Age in months at time of Word ID assessments. Age was low end centered.

Phonological awareness (PA). 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) Non-Word Repetition: the child repeats a nonsense word. The mean of the standard scores of Elision and Blending and Non-Word Repetition scores was used to calculate the Phonological Awareness 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) assesses automaticity of memory retrieval. 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. Standard scores are based on an age normed *t*-distribution.

Pre-K measures

Risk for dyslexia. Risk for dyslexia was determined by having a standard score in the bottom quartile compared to the norming sample for RAN scores and/or PA. Participants with either scores – RAN or PA – in the bottom quartile were identified as being at risk for dyslexia. Risk for dyslexia is coded as a binary variable: having no risk was coded as zero and having risk was coded as one.

Risk for automaticity deficit (RAN risk). Automaticity risk was determined by having a standard score in the bottom quartile compared to the norming sample for RAN scores. Participants with scores in the bottom quartile for RAN were categorized as at risk for a rapid naming deficit. For use in regression models, RAN risk was coded as a binary variable (1 for students identified as having RAN risk, 0 for students who were not identified as having RAN risk).

Risk for phonological deficit (PA risk). Risk for phonological deficit was determined by having a standard score in the bottom quartile compared to the norming sample for PA scores. Participants with scores in the bottom quartile for PA were categorized as at risk for a rapid naming deficit. For use in regression models, RAN risk was coded as a binary variable (1 for students identified as having PA risk, 0 for students who were not identified as having PA risk).

Non-verbal IQ (IQ). Kaufman Brief Intelligence Test, Second Edition Matrices subtest, was used to measure nonverbal intelligence (KBIT-2; Kaufman & Kaufman, 2004). This 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). IQ was used in this analysis as a covariate, to account for the known relation between intelligence and literacy skill development. For this analysis, IQ was grand mean centered to assist interpretation of model intercepts and to analyze between group variation.

School level measure.

School curriculum (SC). Based upon teachers' responses to surveys, correspondence with school and district administrators, and information available on school and district websites, the curriculum(s) used by each school for literacy instruction were determined. To determine whether each school used curriculum that incorporates phonics instruction, I used peer reviewed articles that describe each curriculum and evaluations provided by the United States Institute for Educational Sciences "What Works Clearinghouse" (Institute for Educational Sciences, n.d.).

Analysis Plan

The purpose of these analyses was to understand how early identified risk profiles may be predictive of single word reading development and how this association may be different based upon whether a school used a curriculum that incorporates explicit instructing of phonics. To address these questions, I used RStudio's (Version 1.0.143) *lmer* package to estimate multilevel linear regression models. I used Maximum Likelihood (ML) estimation, which allows for the estimation of both between and within subject variation (Sniders & Bosker, 2012).

Missing Data. Missing data may have been caused by school absences or students leaving the school. These occurrences do not occur at random; they may be the effects of instable home environments, illness, or other factors that can be related to academic and cognitive outcomes. Therefore, the missing data in the data set used for these analyses were "not missing at completely random" and removing these incomplete cases could be inadvertently influence model estimation by removing students who may be at higher risk for dyslexia (Rubin, 1976; Enders, 2010). To replace missing data with reasonable estimates, I used multiple imputation through the *mice* package in R Studio as recommended by Snijders & Bosker (2012) for multilevel analysis. Multiple imputation uses regressions to repeatedly predict missing values (Enders, 2010). These predicted values are used in the analyses multiple times and then aggregated to produce an output for the analysis. Confidence intervals are adjusted in the outputs based upon the confidence intervals of the imputed values. To impute the missing data I used the PA measures (elision, blending, and working memory) and RAN measures for each time point, IQ, dyslexia risk profile, and school curriculum variables.

Intra-Class Correlation. To see how much of the overall variation (between- and within-school variation) in Word ID could be attributed to between-school variation, I estimated models for Word ID of each grade time point individually without any predictors and a random effect for the intercept, which models the variation of Word ID for each individual student. Inclusion of the random effect of the intercept allows for the estimation of how the mean Word ID scores vary between students. I used the variances from these models to calculate the Intra-Class Correlation (ICC) for each grade (Pre-K, Grade 1, Grade 2), a measure of the proportion of variance of the outcome variable between schools (τ_{V00}) relative to the overall variance of the outcome (τ_{V00} + σ^2). These calculations quantified how much of the variance among the Word ID scores can be attributed to the individual schools.

Research question one. To estimate an average trajectory of word reading development over the first three years of school, I regressed Word ID on age at time of assessment and age at time of assessment squared. Age at time of assessment squared was added to assess whether word reading development was nonlinear. I included a random effect for the intercept, which models the variation of Word ID between individuals at the first time point. Each variable – age and age squared – was added sequentially to evaluate whether their addition improved model fit. To assess changes in model fit, I calculated the likelihood ratio test, which evaluated whether the change in deviance (-2LL) is statistically significant based on a chi square distribution (χ^2) when new components (i.e., predictors or random effects) are added to the models. Variables that improve model fit were included in the model. After establishing which variables were to be included, I sequentially added random effects for each variable, and evaluated whether their addition improved model fit based upon the likelihood ratio test. These random effects model the variation of the coefficient between students and schools. The full proposed model for research question one is represented by the following equation (prior to testing whether all variables and random effects significantly reduced the deviance in the model):

(1) $Word\ ID = \pi_{0ij} + \pi_{1ij}(Age_{ti}) + \pi_{2ij}(Age_{ti})^2 + e_{tij}$

$$\pi_{0ij} = B_{000} + r_{0ij}$$

$$B_{000} = \gamma_{000} + u_{0ij}$$

$$\pi_{1ij} = B_{100} + r_{1ij}$$

$$B_{100} = \gamma_{100} + u_{1ij}$$

$$\pi_{2ij} = B_{200} + r_{2ij}$$

$$B_{200} = \gamma_{200} + u_{2ij}$$

Research question two. To assess whether risk for dyslexia at kindergarten/preschool was predictive of students' Word ID development in the first three years of school, I added variables for dyslexia risk and IQ to the model specified for research question one. Dyslexia risk and IQ were time invariant attributes of each student, so they were added to the level that models individual students (referred to as Level 2). IQ was included as a covariate to model the potential association between a student's non-verbal intelligence and Word ID. Each variable – dyslexia risk and IQ – was added sequentially to evaluate whether their addition improves model fit. To assess changes in model fit, I calculated the likelihood ratio test, which evaluates whether the change in deviance (-2LL) is statistically significant based on a chi square distribution (χ^2) when new components (i.e., predictors or random effects) are added to the models. Variables that improve model fit were included in the model. After establishing which variables are to be included, I sequentially added random effects for each variable, and evaluated whether their addition improves model fit based upon the likelihood ratio test. The full proposed model for research question two is represented by the following equation (prior to testing whether all variables and random effects significantly reduced the deviance in the model):

(2) $Word\ ID = \pi_{0ij} + \pi_{1ij}(Age_{ti}) + \pi_{2ij}(Age_{ti})^2 + e_{tij}$

$$\pi_{0ij} = \beta_{000} + \beta_{010}(Dyslexia\ Risk_i) + \beta_{020}(IQ_i) + r_{0ij}$$

$$\beta_{000} = \gamma_{000} + u_{0ij}$$

$$\beta_{010} = \gamma_{010} + u_{0ij}$$

$$\beta_{020} = \gamma_{020} + u_{0ij}$$

$$\pi_{1i} = \beta_{100} + \beta_{110}(Dyslexia\ Risk_i) + \beta_{120}(IQ_i) + r_{1ij}$$

$$\beta_{100} = \gamma_{100} + u_{1ij}$$

$$\beta_{110} = \gamma_{110} + u_{1ij}$$

$$\beta_{120} = \gamma_{120} + u_{1ij}$$

$$\pi_{2i} = \beta_{200} + \beta_{210}(Dysle\boxed{x}a\ Risk_i) + \beta_{220}(IQ_i) + r_{2ij}$$

$$\beta_{200} = \gamma_{200} + u_{2ij}$$

$$\beta_{210} = \gamma_{210} + u_{2ij}$$

$$\beta_{220} = \gamma_{220} + u_{2ij}$$

Research question three. To assess whether attending a school that uses explicit phonics instruction moderated the relationship between deficit risk profile and the development of Word ID in the first three years of school, I added a variable for curriculum to the model specified for research question two. Curriculum was an attribute of the schools the individual students attend, so it was added to the level that models each school (referred to as Level 3). The variable for curriculum will be added to evaluate whether their addition improves model fit. The full proposed model for research question three is represented by the following equation (prior to testing whether all variables and random effects significantly reduced the deviance in the model):

(3) $Word\ ID = \pi_{0ij} + \pi_{1ij}(Age_{ti}) + \pi_{2ij}(Age_{ti})^2 + e_{tij}$

$$\pi_{0ij} = \beta_{000} + \beta_{010}(Dysleixa\ Risk_i) + \beta_{020}(IQ_i) + r_{0ij}$$

$$\beta_{000} = \gamma_{000} + u_{0ij}$$

$$\gamma_{000} = \delta_{000} + \delta_{001}(Curriculum_j) + v_{0ij}$$

$$\beta_{010} = \gamma_{010} + u_{0ij}$$

$$\gamma_{010} = \delta_{010} + \delta_{011}(Curriculum_j) + v_{0ij}$$

$$\beta_{020} = \gamma_{020} + u_{0ij}$$

$$\gamma_{020} = \delta_{020} + \delta_{021}(Curriculum_j) + v_{0ij}$$

$$\pi_{1i} = \beta_{100} + \beta_{110}(Dysleixa\ Risk_i) + \beta_{120}(IQ_i) + r_{1ij}$$

$$\beta_{100} = \gamma_{100} + u_{1ij}$$

$$\gamma_{100} = \delta_{100} + \delta_{101}(Curriculum_j) + v_{1ij}$$

$$\beta_{110} = \gamma_{110} + u_{1ij}$$

$$\gamma_{110} = \delta_{110} + \delta_{111}(Curriculum_j) + v_{1ij}$$

$$\beta_{120} = \gamma_{120} + u_{1ij}$$

$$\gamma_{120} = \delta_{120} + \delta_{121}(Curriculum_j) + v_{1ij}$$

$$\pi_{2i} = \beta_{200} + \beta_{210}(Dysleixa\ Risk_i) + \beta_{220}(IQ_i) + r_{2ij}$$

$$\beta_{200} = \gamma_{200} + u_{2ij}$$

$$\gamma_{200} = \delta_{200} + \delta_{201}(Curriculum_j) + v_{2ij}$$

$$\beta_{210} = \gamma_{210} + u_{2ij}$$

$$\gamma_{210} = \delta_{210} + \delta_{211}(Curriculum_j) + v_{2ij}$$

$$\beta_{220} = \gamma_{220} + u_{2ij}$$

$$\gamma_{220} = \delta_{220} + \delta_{221}(Curriculum_j) + v_{2ij}$$

Research question four. To assess whether dyslexia risk profile at kindergarten/ preschool is predictive of students' Word ID development in the first three years of school, I added variables for each dyslexia risk profile and IQ to the model specified for research question one. Dyslexia risk profile consisted of two binary variables, PA risk and RAN risk. To account for whether having both PA and RAN risk (double deficit) was associated with a different Word ID development rate, added PA risk and RAN risk, and their interaction, to the prediction of the linear slope for age and the quadratic slope for age. Dyslexia risk profile and IQ are modeled as time invariant attributes of each student, so they were added to the level that models individual students (referred to as Level 2). IQ was included as a covariate to model the potential association between a student's non-verbal intelligence and Word ID. Each variable – IQ, PA risk, RAN risk, and their respective interactions – were added sequentially to evaluate whether their addition improves model fit. The full proposed model for research question four is represented by the following equation (prior to testing whether all variables and random effects significantly reduced the deviance in the model):

(4) $Word\ ID = \pi_{0ij} + \pi_{1ij}(Age_{ti}) + \pi_{2ij}(Age_{ti})^2 + e_{tij}$

$$\pi_{0ij} = \beta_{000} + \beta_{010}(PA\ Deficit_i) + \beta_{020}(RAN\ Deficit_i) + \beta_{030}(PA\ Deficit_i * RAN\ Deficit_i) + \beta_{040}(IQ_i) + r_{0ij}$$

$$\beta_{000} = \gamma_{000} + u_{0ij}$$

$$\beta_{010} = \gamma_{010} + u_{0ij}$$

$$\beta_{020} = \gamma_{020} + u_{0ij}$$

$$\beta_{030} = \gamma_{030} + u_{0ij}$$

$$\beta_{040} = \gamma_{040} + u_{0ij}$$

$$\pi_{1i} = \beta_{100} + \beta_{110}(PA\ Deficit_i) + \beta_{120}(RAN\ Deficit_i)$$

$$+ \beta_{130}(PA\ Deficit_i * RAN\ Deficit_i) + \beta_{140}(IQ_i) + r_{1ij}$$

$$\beta_{100} = \gamma_{100} + u_{1ij}$$

$$\beta_{110} = \gamma_{110} + u_{1ij}$$

$$\beta_{120} = \gamma_{120} + u_{1ij}$$

$$\beta_{130} = \gamma_{130} + u_{1ij}$$

$$\beta_{140} = \gamma_{140} + u_{1ij}$$

$$\pi_{2i} = \beta_{200} + \beta_{210}(PA\ Deficit_i) + \beta_{220}(RAN\ Deficit_i)$$

$$+ \beta_{230}(PA\ Deficit_i * RAN\ Deficit_i) + \beta_{240}(IQ_i) + r_{2ij}$$

$$\beta_{200} = \gamma_{200} + u_{2ij}$$

$$\beta_{210} = \gamma_{210} + u_{2ij}$$

$$\beta_{220} = \gamma_{220} + u_{2ij}$$

$$\beta_{230} = \gamma_{230} + u_{2ij}$$

$$\beta_{240} = \gamma_{240} + u_{2ij}$$

Research question five. To assess whether attending a school that uses explicit phonics instruction moderated the relationship between deficit risk profile and the development of Word ID in the first three years of school, I added a variable for curriculum to the model specified for research question two. Curriculum is an attribute of the schools the individual students attend, so they was added the level that models each school (referred to as Level 3). Only variables that improved model fit were included in the model. The full proposed model for research question five is represented by the following equation (prior to testing whether all variables and random effects significantly reduced the deviance in the model):

(5) $Word\ ID = \pi_{0ij} + \pi_{1ij}(Age_{ti}) + \pi_{2ij}(Age_{ti})^2 + e_{tij}$

$$\pi_{0ij} = \beta_{000} + \beta_{010}(PA\ Deficit_i) + \beta_{020}(RAN\ Deficit_i)$$

$$+ \beta_{030}(PA\ Deficit_i * RAN\ Deficit_i) + \beta_{040}(IQ_i) + r_{0ij}$$

$$\beta_{000} = \gamma_{000} + u_{0ij}$$

$$\gamma_{000} = \delta_{000} + \delta_{001}(Curriculum_j) + v_{0ij}$$

$$\beta_{010} = \gamma_{010} + u_{0ij}$$

$$\gamma_{010} = \delta_{010} + \delta_{011}(Curriculum_j) + v_{1ij}$$

$$\beta_{020} = \gamma_{020} + u_{0ij}$$

$$\gamma_{020} = \delta_{020} + \delta_{021}(Curriculum_j) + v_{2ij}$$

$$\beta_{030} = \gamma_{030} + u_{0ij}$$

$$\gamma_{030} = \delta_{030} + \delta_{031}(Curriculum_j) + v_{3ij}$$

$$\beta_{040} = \gamma_{040} + u_{0ij}$$

$$\gamma_{040} = \delta_{040} + \delta_{041}(Curriculum_j) + v_{0ij}$$

$$\pi_{1i} = \beta_{100} + \beta_{110}(PA\ Deficit_i) + \beta_{120}(RAN\ Deficit_i)$$

$$+ \beta_{130}(PA\ Deficit_i * RAN\ Deficit_i) + \beta_{140}(IQ_i) + r_{1ij}$$

$$\beta_{100} = \gamma_{100} + u_{1ij}$$

$$\gamma_{100} = \delta_{100} + \delta_{101}(Curriculum_j) + v_{1ij}$$

$$\beta_{110} = \gamma_{110} + u_{1ij}$$

$$\gamma_{110} = \delta_{110} + \delta_{111}(Curriculum_j) + v_{1ij}$$

$$\beta_{120} = \gamma_{120} + u_{1ij}$$

$$\gamma_{120} = \delta_{120} + \delta_{121}(Curriculum_j) + v_{1ij}$$

$$\beta_{130} = \gamma_{130} + u_{1ij}$$

$$\gamma_{130} = \delta_{130} + \delta_{131}(Curriculum_j) + v_{1ij}$$

$$\beta_{140} = \gamma_{140} + u_{1ij}$$

$$\gamma_{140} = \delta_{140} + \delta_{141}(Curriculum_j) + v_{1ij}$$

$$\pi_{2i} = \beta_{200} + \beta_{210}(PA\ Deficit_i) + \beta_{220}(RAN\ Deficit_i)$$

$$+ \beta_{230}(PA\ Deficit_i * RAN\ Deficit_i) + \beta_{240}(IQ_i) + r_{2ij}$$

$$\beta_{200} = \gamma_{200} + u_{2ij}$$

$$\gamma_{200} = \delta_{200} + \delta_{201}(Curriculum_j) + v_{2ij}$$

$$\beta_{210} = \gamma_{210} + u_{2ij}$$

$$\gamma_{210} = \delta_{210} + \delta_{211}(Curriculum_j) + v_{2ij}$$

$$\beta_{220} = \gamma_{220} + u_{2ij}$$

$$\gamma_{220} = \delta_{220} + \delta_{221}(Curriculum_j) + v_{2ij}$$

$$\beta_{230} = \gamma_{230} + u_{2ij}$$

$$\gamma_{230} = \delta_{230} + \delta_{231}(Curriculum_j) + v_{2ij}$$

$$\beta_{240} = \gamma_{240} + \gamma_{241}(Curriculum_j) + u_{2ij}$$

$$\gamma_{240} = \delta_{240} + \delta_{241}(Curriculum_j) + v_{2ij}$$

Results

Preliminary Analysis

Descriptive Statistics. I analyzed descriptive statistics to assess the normality of distributions of each continuous variable and to identify possible outliers. Table 1 presents the descriptive statistics for the continuous variables used in both the imputation of missing data and in the models estimated to address this study's research questions. Word ID during Pre-K had a positive skew and was leptokurtic because most of the students had scores of zero. The distributions of the Grade 1 and Grade 2 Word ID scores were bimodal (as displayed in figure 1), which suggests that there may be subgroups within the sample population. IQ was normally distributed with a slight positive skew. Ages in each grade were normally distributed, with age in Pre-K being slightly platykurtic and age in Grade 2 having a slight positive skew.

Most of the variables used exclusively for the imputation of missing data were normally distributed. RAN in Pre-K had a slight negative skew and was leptokurtic. RAN in Grade 1 and Grade 2 were normally distributed. Blending in Grade 1 had a slight positive skew, but was normal in Pre-K and Grade 2. PA working memory was normally distributed within all grades. PA elision in all grades had negative skews, especially in Grade 2.

Outliers. There was one case with z-score for IQ above three. Because Word ID and age are longitudinal measures and input into the model as one variable each, I calculated z-scores for each variable combined across time. There were no z-scores above three for either Word ID or Age. To identify potential multivariate outliers, Mahalanobis distance was calculated for the continuous variables (IQ, Age, and Word ID). Eighteen observations (from twelve participants) were identified as multivariate outliers. These observations were all from Grade 1 and Grade 2. There was an even distribution of deficit profiles and IQs within this group. These cases tended to either have notably high or low Word ID scores. I ran analyses with and without these 12 participants and the pattern of the results was the same with regard to interpretation (i.e., statistical significance). Accordingly, the results presented here include all participants

Correlations. Table 2 provides Pearson's correlations among continuous variables, including the continuous variables used for the imputation of missing data. The correlations were calculated among the variables at each time point (Pre-K, Grade 2, and Grade 1). Correlations greater than or equal to .50 are considered large, between .50 and .30 are considered moderate, and between .30 and .10 are considered small (Field, 2013). As expected the outcome variable, Word ID, was significantly correlated with most of the PA variables at most of the time points. Word ID in Pre-K had small to moderate significant correlations with PA blending and elision in Pre-K, Grade 1, and Grade 2, as well as PA working memory in Grade 2. Word ID in Grade 1 and Grade 2 had significant small or moderate correlations with PA blending, working memory, and elision in Pre-K, Grade 1, and Grade 2. Word ID in Pre-K was moderately correlated with RAN only in Pre-K, but not in Grade 1 or Grade 2. Word ID in Grade 1 had a significant moderate correlation with RAN in Pre-K, Grade 1, and Grade 2. Word ID in Grade 2 had a moderate significant correlation with RAN in Pre-K, and small significant correlations with RAN in Grade 1 and Grade 2. This pattern suggests that the RAN measure taken in Pre-K (as done in this study to create the Dyslexia Deficit Profiles) may be more predictive of Word ID than RAN measures taken in later years.

Frequencies and cross tabulations for categorical variables. Of the 161 participants, 80 were identified as being at risk for dyslexia. Of the participants identified as being at risk for dyslexia, 32 were identified as at risk for only a RAN deficit, 31 were identified as at risk for only a PA deficit, and 17 were identified as at risk for a Double Deficit. Table 3 presents the means of Word ID for participants identified as being at risk for dyslexia (the binary variable) and those who were not at each time point. There were significant differences between Word ID means of the group not identified as at risk for dyslexia and the group who were identified as at risk for dyslexia in Pre-K ($t(123.14) = 2.41, p = 0.02$) Grade 1 [$t(148.8) = 3.22, p < 0.01$], and Grade 2 [$t(142.58) = 3.27, p < 0.01$]. With regard to deficit profile group (the three-category variable), a One-Way ANOVA, presented in table 3, showed statistically significant differences in Word ID means between deficit profile groups in Grade 1 [$F(3, 147) = 3.22, p < 0.01$] and Grade 2 [$F(3, 141) = 3.27, p < 0.01$], but not in Pre-K [$F(3, 125) = 2.47, p = 0.07$]. Boxplots in Figure 1 show the variation of mean Word ID by Grade 2 grouped deficit risk profile.

Imputation. I created 50 imputed data sets using PA (elision, blending, and working memory) and RAN measures and for each time point as well as the variables used in the models estimated to address the research questions (IQ, dyslexia risk profile, and school curriculum variables). This number of data sets is more than twice the number of data sets recommended for robust analysis using imputation (Enders, 2010). These imputed data sets were used in the following analyses and the results reported below were produced by pooling results from the analyses of each imputed data set.

Intraclass-Correlation (ICC). To assess the amount of total variance in Word ID that could be related to differences in mean Word ID scores between schools, the variances from the null models for each grade (Pre-K, Grade 1, Grade 2) were used to calculate the intra-class correlation. The null models included an intercept, which estimates the overall mean in Word ID at that time point, and random effects for that intercept, which estimate the variance of the Word ID means for the individual students and the variance between the Word ID means for individual students and means for Word ID for the school.

The boxplots in Figure 2 shows the variation of Word ID within and between schools. In Pre-K, the variance between school means for Word ID and the overall mean for Word ID, $\tau_{v00,}$ was 1.54. The variance between the mean for all students and the individual student means for Word ID, σ^2, was 2.34. The intra-class correlation, accordingly, was .40, which indicated that 40% of the variation of Word ID in Pre-K was between schools (with the remaining 60% between students within schools). In Grade 1, the variance between school means for Word ID and the overall mean for Word ID, $\tau_{v00,}$ was 1.34. The variance between the mean for all students and the individual student means for Word ID, σ^2, was 3.99. The intra-class correlation, accordingly, was .25, which indicated that 25% of the variation of Word ID in Grade 1 was between schools. In Grade 2, the variance between school means for Word ID and the overall mean for Word ID, $\tau_{v00,}$ was .28. The variance between the mean for all students and the individual student means for Word ID, σ^2, was 1.49. The intra-class correlation, accordingly, was .11, which indicated that .11% of the variation of Word ID in Grade 2 was between schools.

Research Question One

To estimate the rate of change in Word ID between Pre-K and the end of Grade 2, I regressed Word ID on the students' age in months and the students' age in months squared (to account for the possibility that the rate of growth was not linear). I added each fixed and random effect into the model sequentially and only included each effect if its addition improved model fit. The addition of the following variables all individually improved model fit: the fixed effect for age ($\chi^2(1) = 322.81$, $p > .0001$), the fixed effect for age squared ($\chi^2(1) = 45.57$, $p > .001$), the random effect for age ($\chi^2(1) = 5.525$, $p > .05$), and the random effect for age squared ($\chi^2(1) = 5.47$, $p = 0.019$). These variables were retained in subsequent models estimated for this study.

The model estimated to answer research question one is summarized in Table 6. The intercept, which estimated the average Word ID score for students when they were 60 months of age, was -6.47 (γ_{000}= -6.47, SE = 0.56, p < .0001). Although a negative score on Word ID is not possible, this number can be interpreted to mean that most of the students were able to identify words until they were older than 60 months. The coefficients for age and age squared were 1.37 (γ_{100}= 1.37, SE = 0.57, p < .0001) and -0.01 (γ_{200}= -0.01, SE = 0.001, p < .0001) respectively. These coefficients were added together to produce the rate of change, such that, on average, students were able to identify 1.36 more words for the first month (61 months of age) and 1.37 for the next month (62 months of age) and so on. The small significant negative coefficient for age squared suggests that this rate of increase in Word ID is not linear and that it decreases slightly over time.

To tests the assumption that the residuals were normally distributed, I produced a histogram of the residuals and calculated the Shapiro-Wilk statistic. The histogram displayed a normal distribution of residuals. The Shapiro-Wilk statistic for these residuals showed that this distribution was not significantly different from a normal distribution (W = 0.995, p = 0.146). To tests the assumption that the residuals were evenly distributed regardless of the age at which Word ID was measured, I plotted the residuals against age. This plot displayed and even distribution of residuals when Word ID was assessed at over 80 months of age. But the model was consistently over-predicting Word ID when the measures were taken at close to 60 months of age and under-predicting Word ID for when the measures were taken between 70 and 80 months. This pattern may exist because at these younger ages most participants were unable to identify any words.

Research Question Two

To assess whether risk for dyslexia moderated the relationship between Word ID and age, over and above IQ, I regressed Word ID on age, age squared, IQ, and risk for dyslexia, with interaction coefficients for both IQ and risk for dyslexia with linear slope for age and quadratic slope for age. Random effects for the intercept, age, and age squared were included in the model. IQ and risk for dyslexia were added sequentially and only included in subsequent models if their addition improved model fit.

IQ did not significantly improve model fit when compared to the model for rate of change for Word ID estimated in research question one (χ^2(1) = 2.02, p = 0.155). The coefficient for IQ's prediction of the intercept – Word ID at 60 months – was not statistically significant ($\gamma_{01}0$= 0.01, SE = 0.06, p = .83). The coefficient IQ's interaction with age – IQ's change in the association between age and Word ID – was not statistically significant (γ_{110}= 0.01, SE = 0.006, p = .20). The coefficient IQ's interaction with age squared – IQ's change in the association between age squared and Word ID – was not statistically significant (γ_{210}= -0.0002, SE = 0.0001, p = .17). Therefore, the model did not indicate a statistically significant association between the rate of change for Word ID and IQ. Because it did not improve model fit nor did it significantly predict Word ID at 60 months or rate of change in Word ID, IQ was not incorporated into any other subsequent models.

The model estimated to answer research question two is summarized in Table 7. The addition of the binary dyslexia risk variable to the model did not significantly improve model fit when compared to the model for rate of change for Word ID estimated for research question one (χ^2(3) = 4.74, p = .192). The intercept, which estimated the average Word ID score at 60 months of age for students who were not identified as at risk for dyslexia, was -6.49 (γ_{000}= -6.37, SE =

0.807, $p < .0001$). The coefficient of binary dyslexia right predicting the intercept, which estimated the difference between the averages of Word ID score at 60 months of age for students identified as having and not having risk for dyslexia, was nonsignificant (γ_{010}= -.27, SE = 1.132, p = .81). Therefore, the model did not indicate a statistically significant difference between the average Word ID score at 60 months of age when comparing students who were identified as at risk for dyslexia and those who were not identified as at risk for dyslexia.

The coefficients for the linear rate of change based on age and age squared (the quadratic slope for change) were 1.43 (γ_{100}= 1.43, SE = 0.080, $p < .0001$) and -0.02 (γ_{200}= -0.02, SE = 0.002, $p < .0001$) respectively. These coefficients are estimates of the rate of change of Word ID for students who were not identified as having risk for dyslexia. The coefficient for age and the interaction with risk for dyslexia was nonsignificant (γ_{200}= 0.10, SE = 0.113, $p = .36$). The coefficient for age squared and the interaction with risk for dyslexia was nonsignificant (γ_{200}> 0.001, SE = 0.002, $p = .69$). These nonsignificant interaction coefficients indicated that the model did not identify a statistically significant difference between the rate of change of Word ID when comparing students who were identified as at risk for dyslexia and those who were not identified as at risk for dyslexia, when using the binary risk variable.

To test the assumption that the residuals were normally distributed, I produced a histogram of the residuals and calculated the Shapiro-Wilk statistic. The histogram displayed a normal distribution of residuals. The Shapiro-Wilk statistic for these residuals showed that this distribution was not significantly different from a normal distribution (W = 0.995, p = 0.153). When testing the assumption that the residuals were evenly distributed regardless of the age at which Word ID was measured, the pattern of residuals was similar to the patter of residuals from the model produced in research question one. To test whether the distribution of residuals differed based upon risk, I plotted the residuals by risk profile. The distribution of residuals was constant for both groups of students.

Research Question Three

To assess whether a school using a curriculum that incorporates explicit phonics instruction (phonics curriculum) moderated the relationship between risk for dyslexia and rate of change of Word ID, I regressed Word ID on age, age squared, risk for dyslexia, and curriculum, with an interaction variable for risk for dyslexia, curriculum, and age as well as interaction variables for dyslexia, curriculum, and age squared. Random effects for the intercept age and age squared were included in the model. IQ was not included in the model because it did not significantly improve model fit when compared to the model for rate of change for Word ID ability estimated for research question one ($\chi^2(1) = 2.02$, $p = 0.155$).

The model estimated to answer research question three is summarized in Table 8. The addition of the phonics curriculum variable did not significantly improve model fit when compared to the model for rate of change for Word ID ability estimated for research question one ($\chi^2(6) = -5.21$, $p < .05$). The intercept, which estimated the average Word ID score at 60 months of age for students who were not identified as having risk for dyslexia and did not have phonics curriculum available at their schools, was -7.71 (δ_{000}= -7.71, SE = 1.649, $p < .0001$). The coefficients for risk for dyslexia and phonics curriculum predicting the intercept were nonsignificant. The non-significance of these coefficient indicates that the model did not identify a statistically significant difference between the average Word ID score at 60 months of age for students at risk for dyslexia who did not have phonics curriculum in their schools (δ_{001}= 1.40, SE = 2.983, $p = .332$) or for students without risk for dyslexia who did have phonics curriculum in their schools (δ_{010}= 2.66, SE = 1.983, $p = .199$). The interaction between phonics curriculum and dyslexia risk predicting the intercept was not significant (δ_{010}= -2.09, SE = 2.63 p = .428). The non-significance of this interaction coefficient indicates that the model failed to find a statically significant difference between the average Word ID score at 60 months of age for students at risk for dyslexia who have phonics curriculum at their schools and students without phonics curriculum who did not have phonics curriculum at their schools.

The coefficients for age and age squared were 1.36 (δ_{100}= 1.36, SE = 0.152, p < .0001) and -0.32 (δ_{200}= -0.32, SE = 0.002, p < .0001) respectively. These are estimates of the rate of change of Word ID for students who were not identified as having risk for dyslexia and who did not have phonics curriculum in their schools. The coefficient for the interaction between for age and risk for dyslexia was not significant (δ_{110}= -0.04, SE = 0.147, p = .793). The non-significance of this coefficient indicates that the model did not identify significant differences between the linear rate of change of Word ID when comparing students who were and students who were not identified as being at risk for dyslexia, all of whom attended schools without phonics curriculum. The coefficient for the interaction between the linear slope for age and risk for dyslexia and phonics curriculum was not significant (δ_{111}= -0.32, SE = 0.256, p = .218). The non-significance of this coefficient indicates that the model did not identify significant differences between the linear rate of change of Word ID for students who did attend schools with phonics curriculum and were identified as being at risk for dyslexia and did not attend schools with phonics curriculum and were not identified as being at risk for dyslexia. The coefficient for the interaction between for age squared (quadratic slope for age) and risk for dyslexia was not significant (δ_{210}= -0.01, SE = 0.004, p = .208). The coefficient for the interaction between for age squared (quadratic slope for age) and phonics curriculum was not significant (δ_{201}= 0.003, SE = 0.004, p = .456). The coefficient for the interaction between for age squared and risk for dyslexia and phonics curriculum was also not significant (δ_{211}= 0.01, SE = 0.005, p = .172).

To tests the assumption that the residuals were normally distributed, I produced a histogram of the residuals and calculated the Shapiro-Wilk statistic. The histogram displayed a normal distribution of residuals. The Shapiro-Wilk statistic for these residuals showed that this distribution was not significantly different from a normal distribution (W = 0.997, p = 0.419). When testing the assumption that the residuals were evenly distributed regardless of the age at which Word ID was measured, the pattern of residuals was similar to the patter of residuals from the model produced in research question one. The distribution of residuals was consistent between for both students with and without risk. To test whether the distribution of residuals differed based upon school curriculum, I plotted the residuals by risk profile. The distribution of residuals was constant for both regardless of school curriculum.

Research Question Four

To assess whether dyslexia risk profile moderated the relationship between Word ID and age, I regressed Word ID on age, age squared, PA risk (binary), RAN risk (binary) and an interaction between PA risk and RAN risk (binary variable for double deficit risk). The models also included interaction coefficients for PA risk, RAN risk, the interaction between PA risk and RAN risk with age and age squared. Random effects for the intercept age and age squared were included in the model. IQ was not included in the model because it did not significantly improve model fit when compared to the model for rate of change for Word ID ability estimated for research question one ($\chi^2(1) = 2.02, p = 0.155$). Each deficit variable was added sequentially into the model. First PA risk was added, then RAN risk, and finally the interaction between PA risk and RAN risk, which represents the association those students with risk for double deficits (both

PA and RAN deficits). The models estimated to answer research question four are summarized in Table 9.

Addition of PA Risk. The addition of PA risk significantly improved model fit when compared to the model for rate of change for Word ID ability estimated for research question one (χ^2 (3) = 47.31, p < .0001). The intercept, which estimated the average Word ID score at 60 months of age for students who were not identified as having any risk for dyslexia, was -8.95 (γ_{000}= -8.95, SE = 0.818, p < .0001). The model failed to find a statistically significant difference between the average Word ID score at 60 months of age for students without PA risk and those with PA risk (γ_{010}= 1.14, SE = 1.299, p = .3815). The coefficients for age and age squared were 1.69 (γ_{100}= 1.69, SE = 0.080, p < .0001) and -0.02 (γ_{200}= -0.02, SE = 0.002, p < .0001) respectively. These coefficients were added together to produce the rate of change, such that, on average, for students who were not identified as having PA risk were able to identify 1.67 more words for the first month (61 months of age) and 1.65 for the next month (62 months of age) and so on.

The model estimated significant differences in the rate of growth for students with PA risk, when compared with students who were identified as not having PA risk. The coefficient for the interactions between age and PA risk was -0.34 (γ_{110}= -0.34, SE = 0.182, p = .0002). The coefficient for the interactions between age squared and risk for double deficit was 0.005 (γ_{210}= 0.005, SE = 0.004, p = .017). The significance of these coefficient indicates that students identified as at risk for PA deficit in Pre-K have, on average, a growth rate in Word ID development of 1.33 words per month.

Addition of RAN risk. The addition of RAN risk did not significantly improve model fit when compared to the model for rate of change for Word ID ability with the PA risk variable (χ^2

(3) = 5.67, p < .05). The intercept, which estimated the average Word ID score at 60 months of age for students who were not identified as having any risk for dyslexia was -7.75 (γ_{000}= -7.75, SE = 0.975, p < .0001). The model failed to find statistically significant differences between the average Word ID score at 60 months of age for students without either PA or RAN risk and those with PA deficits (γ_{010}= 1.10, SE = 1.299, p = .3815) or those with RAN risk (γ_{020}= -0.33, SE = 1.363, p = .812). The coefficients for age and age squared were 1.62 (γ_{100}= 1.62, SE = 0.080, p < .0001) and -0.02 (γ_{200}= -0.20, SE = 0.002, p < .0001) respectively. These coefficients were added together to produce an estimated rate of change for students who were identified as having neither PA or RAN deficits, such that, on average, these students could identify 1.60 words more at 61 months of age. The model failed to find significant differences in the rate of growth for students with PA risk (γ_{110}= -0.20, SE = 0.118, p = 0.084) and RAN risk (γ_{110}= -0.15, SE = 0.119, p = 0.206) when compared with students who were identified as neither RAN nor PA risk.

Addition of PA and RAN risk interaction. The addition of dyslexia risk profile (interaction variable) did not significantly improve model fit when compared to the model for rate of change for Word ID (χ^2 (6) = -3.01, p > .05). The intercept, which estimated the average Word ID score at 60 months of age for students who were not identified as having any risk for dyslexia, was -7.76 (γ_{000}= -7.76, SE = 0.932, p < .0001). The interactions of the dyslexia profile variables with the intercept were nonsignificant. The model failed to find a statistically significant difference between the average Word ID score at 60 months of age for students at risk for PA deficits (γ_{010}= -0.47, SE = 1.562, p = .765), RAN deficits (γ_{020}= -1.141, SE = -1.588, p = .474), and double deficits (γ_{030}= 0.60, SE = 2.094, p = .775).

The coefficients for age and age squared were 1.43 (γ_{100}= 1.61, SE = 0.087, p < .0001) and -0.02 (γ_{200}= -0.02, SE = 0.002, p < .0001) respectively. These coefficients were added together to produce an estimated rate of change for students who were identified as having no risk for dyslexia, such that, on average, students were able to identify 1.59 words more every month. The model failed to estimate significant differences between rates of change in Word ID for students at risk for PA and RAN deficits when compared to students who were identified as having no risk. The coefficients for the interactions between age and risk for PA deficit were not significant (γ_{110}= -0.02, SE = 0.136, p = .903). The coefficients for the interactions between age and risk for RAN Deficit were not significant (γ_{120}= 0.04, SE = 0.136, p = .778). The coefficients for the interactions between age squared and risk for PA deficit (γ_{210}= -0.001, SE = 0.003, p = .373), risk for RAN Deficit (γ_{120}= -0.001, SE = 0.003, p = .667), were not significant.

Alternatively, the model estimated significant differences in the rate of growth for students at risk for double deficits, when compared with students who were identified as having no risk. The coefficient for the interactions between age and risk for double deficit was -0.55 (γ_{130}= -0.55, SE = 0.182, p = .003). The coefficient for the interactions between age squared and risk for double deficit was 0.01 (γ_{110}= 0.01, SE = 0.004, p = .007). These coefficients were added together to produce the rate of change, such that, on average, for students who were identified as at risk for double deficit were able to identify 0.54 more words for the first month (61 months of age) and 1.55 for the next month (62 months of age) and so on.

To tests the assumption that the residuals were normally distributed, I produced a histogram of the residuals and calculated the Shapiro-Wilk statistic. The histogram displayed a normal distribution of residuals. The Shapiro-Wilk statistic for these residuals showed that this distribution was not significantly different from a normal distribution (W = 0.995, p = 0.168). When testing the assumption that the residuals were evenly distributed regardless of the age at which Word ID was measured, the pattern of residuals was similar to the patter of residuals from the model produced for the previous research questions. To assess the distribution of residuals for each deficit risk group, I created boxplots of the residuals for those with no deficit risk, those with PA risk only, those with RAN risk only, and those with both PA and RAN risks. The distribution of residuals was constant for students with each deficit risk profile. But the distribution of residuals for those no deficits risk was greater compared to the distribution for each risk profile. This suggests that the model was over-predicting and under-predicting more for those without deficit risk.

Research Question Five

To assess whether a school using curriculum that incorporates explicit phonics instruction (phonics curriculum) moderated the relationship between dyslexia risk profile and rate of increase of Word ID from Pre-K to Grade 1, I regressed Word ID on age and age squared with an interaction variable for dyslexia risk profile, curriculum, and age as well as an interaction variable for dyslexia risk profile, curriculum, and age squared. Random effects for the intercept age and age squared were included in the model. IQ was not included in the model because it did not significantly improve model fit when compared to the model for rate of change for Word ID ability estimated for research question one ($\chi^2(1) = 2.02$, $p = 0.155$).

The model estimated to answer research question five is summarized in Table 10. The addition of Phonics Curriculum variables did not significantly improve model fit when compared to the model for rate of change for Word ID ability estimated for research question four ($\chi^2(9) = 4.657$, p <05). The intercept, which estimated the average Word ID score at 60 months of age for students who were not identified as having any risk for dyslexia and did not have phonics curriculum available at their schools, was -9.70 (δ_{000}= -9.70, SE = 1.770, $p < .0001$). The intercepts for each dyslexia risk profile were all nonsignificant. This finding indicates that the model failed to find a statistically significant difference in the average Word ID score at 60 months of age between students with PA risk (δ_{010}= 1.35, SE = 3.727, p = . 701), RAN risk (δ_{020}= 3.89, SE = 2.137, p = .155), or double deficit risk (δ_{030}= -2.80, SE = 5.168, p = .816), who did not go to schools with phonics curriculum, when compared to students who were not classified as having risk for dyslexia, who also did not go to schools with phonics with phonics curriculum.

The interactions between each dyslexia risk profile and phonics curriculum in the prediction of the intercept were all nonsignificant. This finding indicates that the model failed to find a statistically significant difference in the average Word ID score at 60 months of age between the students within each deficit profile group who went to schools with phonics curriculum and those in the same deficit profile group who did not. The model failed to find statistically significant differences between students with PA risk Word ID scores at 60 months of age who did and did not attend schools with phonics curriculum (δ_{011}= -0.97, SE = 4.173, p = . 816), the model failed to find statistically significant differences between students with RAN risk Word ID scores at 60 months of age who did and did not attend schools with phonics curriculum (δ_{021}= -5.77, SE = 3.517, p = .103), the model failed to find statistically significant differences between students with double deficit risk Word ID scores at 60 months of age who did and did not attend schools with phonics curriculum (δ_{031}= 4.65, SE = 5.740, p = .419).

The coefficients for age and age squared were 1.53 (δ_{100}= 1.53, SE = 0.172, $p < .0001$) and -0.02 (δ_{200}= -0.02, SE =0.003, $p < .0001$) respectively. These coefficients were added together to produce an estimated rate of change for students who were identified as having no risk for dyslexia, such that, on average, students could identify 1.51 words more every month. The model failed to estimate significant differences between the growth rates of each of groups students with deficit profile who did not have phonics curriculum in their schools and the students who were identified as not have risk for dyslexia. The coefficient for the interaction between age and PA risk (δ_{111}= 0.40, SE = 0.338, p = .917), RAN risk (δ_{112}= -0.04, SE = 1.244, p = .866), or double deficit risk (δ_{113}= -0.06, SE = 0.448, p = .890), were all non-significant. The coefficient for the interaction between age squared and PA risk (δ_{210}= -0.004, SE = 0.007, p = .555), RAN dyslexia risk profiles (δ_{220}= -0.001, SE = 0.005, p = .882), or double deficit dyslexia risk profiles (δ_{230}= 0.0002, SE = 0.008, p = . 976), were all non-significant.

The model failed to estimate statically significant differences between the rates of change of Word ID for students who did and did not attend schools with phonics curriculum. The coefficient for the interaction between age and phonics curriculum was not significant (δ_{101}= .23, SE = .214, p = .890). The coefficient for the interaction between age squared and phonics curriculum was not significant (δ_{201}= -.02, SE = 0.004, p = 0.274). The non-significance of this coefficient indicates that for students who were not identified as having risk for dyslexia, the model did not find statistically significant differences between their rates of growth in Word ID within their first years of school.

The model did not find statistically significant differences in rates between students with each deficit profile who had phonics curriculum and those students with each deficit profile without phonics curriculum. The coefficients for the interaction between age, PA deficit profile, and phonics curriculum were not significant (δ_{111}= 0.01, SE = 0.008, p = 0.315). The coefficients for the interaction between age, RAN risk, and phonics curriculum were not significant (δ_{121}= 0.005, SE = 0.006, p = 0.429). The coefficients for the interaction between age, double deficit profile, and phonics curriculum were not significant (δ_{131}= 0.01, SE = 0.010, p = 0.203). The coefficients for the interaction between age squared, PA risk, and phonics curriculum were not significant (δ_{211}= 0.01, SE = 0.008, p = 0.315). The coefficients for the interaction between age, RAN risk, and phonics curriculum were not significant (δ_{221}= 0.005, SE = 0.006, p = 0.429). The coefficients for the interaction between age, double deficit risk, and phonics curriculum were not significant (δ_{231}= 0.01, SE = 0.010, p = .203).

To test the assumption that the residuals were normally distributed, I produced a histogram of the residuals and calculated the Shapiro-Wilk statistic. The histogram displayed a normal distribution of residuals. The Shapiro-Wilk statistic for these residuals showed that this distribution was not significantly different from a normal distribution (W = 0.997, p = 0.422). When testing the assumption that the residuals were evenly distributed regardless of the age at which Word ID was measured, the pattern of residuals was similar to the patter of residuals from the model produced for the previous research questions. Consistent with the residuals analyzed for the model produced in for question four, this model was over-predicting and under-predicting more for those without deficit risk. The distribution of residuals was constant for both regardless of school curriculum.

Discussion

The purpose of these analyses was to evaluate a method for early identification of dyslexia by analyzing the developmental trajectory of students' word reading ability during their first three years of school. By dividing the students into subgroups based on their performance on cognitive and linguistic assessments upon or prior to entering kindergarten, I was able to assess whether students with different abilities entering elementary school had different trajectories in word reading development. These subgroups of students were created based on the double deficit hypothesis, which posits that dyslexia is characterized by a deficit in phonological awareness, automaticity, or both (Wolf et al., 2002; Miller et al., 2006; Powel et al. 2007; Ozernov-Palchik et al., 2016; Groot et al., 2017). Based upon this theory, students who performed in the bottom quartiles, compared to a normed sample, on assessments of phonological awareness and/or rapid automatized naming were identified as at risk for that respective defect. These subgroups had significantly different mean word identification abilities at various time points. Furthermore, two of the subgroups – those displaying risk for phonological deficit and those displaying risk for double deficits (risk for both automaticity and phonological deficits) – had significantly lower rates of growth on word identification than their typically developing peers during the scope of the study.

I also considered how instruction may be associated with different word identification developmental trajectories by modeling the effect of the curriculum available at the students' schools on growth of word identification. Researchers have indicated that accounting for the potential influence of instruction is essential to accurately predict reading development (Compton et al., 2010).

Because a substantial amount of research has found that explicit instruction of phonics is associated with better reading outcomes, especially in early elementary school (National Reading Panel, 2000), I classified each of the schools' curriculum as either incorporating explicit phonics instruction, or not. Notably, I did not directly measure the use of this curriculum at the schools, so the only effect modeled was the availability of phonics curriculum in the schools. Models that included this variable failed to find a significant effect of the presence of phonics curriculum on word identification growth for any of the risk subgroups or for the typical students. Accordingly, the results of similar analyses may be different if the actual implementation of the curriculum at the schools can be measured.

The models estimated for the analyses above showed that the between school variation in word identification for each school grew smaller and smaller throughout the first three years of elementary school. When the students were in kindergarten the between school variation was the greater than when the students were in first, and it was even smaller in second grade. This may seem counterintuitive, as students who spend more time in their respective schools should theoretically become more similar to one another than they are to their peers from other schools. There are two possibilities for this finding. First, students in the same schools did not necessarily have the same teachers and teachers may have greater influence over the instruction a student receives than the school does. Therefore, students form the same schools may have had very different educational experiences if they are not in the same classes. Second, students who are struggling readers may not be responding to the instruction of their teachers and therefore be falling behind their peers and increasing the within school variation in word identification. This hypothetical phenomenon may have been even more amplified in my models because the sample was specifically stratified to have more struggling readers.

There were significant differences between the word identification scores of students who were identified as at risk for dyslexia – having any of the risk profiles – and students who were identified as not at risk for dyslexia upon or prior to entering kindergarten, in first, and in second grade. But there were not significant differences between the rates of growth of these groups. This finding indicates that, on average, students who score in the bottom quartile on tests of phonological awareness and/or rapid automatized naming remain behind their peers, but they develop their word identification abilities at similar rates.

When evaluating the word identification development of students who scored in the bottom quartiles on phonological and automaticity assessments prior to or at the beginning of kindergarten, I found distinct patterns of word identification development. Students within various subgroups had specific patterns of word identification ability between kindergarten and the end of second grade. Both students who were identified as at risk for phonological deficits and those at risk for double deficits had significantly lower word identification scores, in varying grades, and had significantly lower rates of growth for word identification.

Students who were identified as at risk for phonological deficits, had lower word identification performance and growth than those who were not identified as at risk for phonological deficits. These at risk students had, on average, significantly different scores on word identification at the end of second grade, but not in Pre-K or first grade. Furthermore, these students who were identified as at risk for phonological deficits had significantly different rates of growth in word identification abilities between kindergarten and second grade when compared with the students who were identified as having no risk for phonological deficits. On average children with phonological deficits were increasing their word identification score by about four words every three months, whereas their peers without phonological deficits were increasing their scores at a rate of close to five words every three months. Put differently, children who were identified, prior to entering elementary school, as having risk for phonological deficits were learning to identify words at 80% of the rate of their typical peers.

Significant differences in word identification between students who scored in the bottom quartile on the rapid automatized naming assessment, when compared with a normed sample, and their typical-scoring peers were only found at the end of first grade. In contrast, their scores on average in pre-kindergarten and at the end of second grade, and their rates of development, were not significantly different from their typical-scoring peers. The difference in mean scores in each grade for these students and the typical student group never exceeded two words. Therefore, on average, this group of students were not lagging substantially behind their peers.

The students who were identified as at risk for a double deficit had, on average, substantially lower rates of growth in word identification during their first three years of school. These students were identifying, on average, just over two words more every two months on the word identification assessments while their typically-scoring peers were identifying just over three words every two months. In other words, student with risk for double deficits were developing word identification skills at two thirds the rate of their typical peers.

There were statistically significant differences in the mean word identification scores for the double deficit group when compared to their typically developing peers, such that the students who were identified as having no risk for dyslexia could identify, on average, almost seven more words than those identified as having risk for double deficit in second grade on the assessment. Notably this difference in mean word identification scores was reduced, on average, to just under four words at the end of second grade. This reduction in the differences between these groups' scores was also accounted by the growth rate model, as the quadratic coefficient for age was slightly less negative for the double deficit students when compared to the typical students. This finding indicates that the rate of change reduces more slowly over time for the double deficit students than the typical students. This pattern may indicate that, on average the double deficit students were catching up to their peers.

The finding that the subgroup of students at risk for a double deficit have, on average, the lowest word identification growth rate is constant with previous research which indicates that the double deficit is a quantitatively more severe deficit than the single deficits (Miller et al., 2006). Interestingly, the difference between those identified as having risk for phonological deficits and those with no deficits was not significant when those with double deficits were included in the model. This finding indicates that phonological awareness alone may not be predictive of word reading growth rates, when the assessments are taken at such a young age.

Notably, nonverbal intelligence was not predictive of word reading growth. This finding can be added to a growing number of studies that have failed to find that measure of intelligence has a predictive power for reading related skills (Ozernov-Palchik et al., 2016; Gresham & Vellutino, 2010; Vellutino et al., 2006). Furthermore, it supports the argument of researchers and clinicians who believe that using measure of intelligence in delineating definitions and diagnostic standards for dyslexia is arbitrary (Fletcher, Lyon, Fuchs, & Barnes, 2007; Hoskyn & Sawnson, 2000). An alternative argument is that, because the intelligence scores in preschool are not predictive of scores in elementary school (Deary et al., 2004), the failure to estimate significant coefficients for the IQ variables was possibility due to the lack of measure stability. Therefore, this finding may not eliminate the possibility of the importance of intelligence in reading development. Further research is required to fully understand intelligence's role in reading development and dyslexia.

Predicting Risk for Dyslexia

The findings of this study are constant with previous research on deficits and prediction of dyslexia. Thompson et al. (2015) tracked students from age three and half to age eight to establish when dyslexia was detectable with accuracy and what variables to utilize in detection. They found that across ages the most constantly reliable were measures of phonological ability, automaticity, and letter knowledge. Carroll et al. (2015) also found that phonological awareness and rapid automatized naming were among the strongest predictors of poor reading, along with alphabet knowledge and short term verbal memory.

Notably, both Thompson et al. (2015) and Carroll et al. (2015) used logistic regression to determine the odds of a student's diagnosis of dyslexia in later grades based on performance on various measure at an early age. These models require dichotomous outcomes, a determination of whether someone is or is not dyslexic based on a specific standard. In both studies the standard for being considered dyslexic was being 1.5 standard deviations below the group which was previously determined to be low risk on literacy outcomes measures.

One unique advantage of the present study is that it focuses on growth trajectory as opposed to a purely binary outcome. Many researchers have argued against the use of simple percentile cut offs on a specific measure for diagnosing learning disabilities because this method has been shown to be ineffective at establishing who may need and respond to interventions (Elliot & Grigorenko, 2014; Pendleton, 2006). In a longitudinal study spanning from the beginning of kindergarten through the end of second grade, Pennington and Lefty (200) found that the students who were identified as high risk for dyslexia in kindergarten but were not diagnosed for dyslexia in second grade, still scored significantly less on reading measures than those who were never identified as at risk. This finding suggests that the cut off method of outcomes is deficient for identifying all children who may need additional intensive reading instruction than the typical student. Furthermore, it suggests that dyslexia is comprised of a multivariate continuum, wherein students can have varying intensities of several different deficits, all which effect reading abilities (Pennington, 2006; Crisp et al., 2011). Although this study began by using percentile cut offs to establish risk groups, multiple assessments of different abilities were used to determine those categories and the outcome focused on in this study was a developmental trajectory, not a potentially arbitrary standard for achievement.

The main findings of this study – that low performance on assessments of phonology, or on assessments of phonology and automaticity together, prior to a child receiving formal reading instruction is predictive of lower word reading growth rates – suggests that these measures could be used as screening tools for dyslexia. There is an important distinction between screening and diagnosis, as screening is meant to flag children who have a high probability of risk, whereas diagnosis is meant to provide a definitive categorization conclusion about whether or not a child has dyslexia (Helland et al. 2011). Alternatively, screening tools can provide parents and educators information about how to effectively allocate resources for helping children become proficient readers.

Limitations and Future Research Directions

One of the main limitations of this analysis was the small number of data points, especially between kindergarten and first grade. Researchers who study early reading development note that literacy screenings can be the equivalent of “hitting a moving target” (pg. 131, Elliot & Grigorenko, 2014), collection of more data during these two years would have allowed a more precise estimate of the intercept, which represents the time when children transition from pre-literacy skills to actively reading words. These data would also increase the precision of the estimate of the word identification growth rate.

Another limitation for this study was the amount of missing data for the outcome variable. Imputation is an effective way of avoiding producing potentially spurious results by deleting cases, because these missing data points are often caused by factors relevant to academic achievement (Enders, 2010). But it is still ideal to analyze complete data sets.

Although, phonological awareness and automaticity have consistently been found as the most important variables in predicting dyslexia, there are other possible indicators of dyslexia that should be evaluated for their use as potential screening measures. Future research on early identification of risk for dyslexia could also consider variables such as motor skills, executive functioning, and auditory processing, which have been found to be predictive of dyslexia (Helland et al., 2011, Carroll et al., 2015).

Perhaps the most important limitation was the lack of data about instruction. Although I was able to model whether the presence of curriculum that incorporates explicit phonics instruction at a school was associated with a change in word identification growth rate, I was not able to model the use of that instruction in the classroom. Although there is ample research supporting the efficacy of phonics based instruction (National Reading Panel, 2000), simply having access to curriculum that incorporates phonics is not insufficient. One recent study found that many early elementary school teachers did not have an understanding of how to explicitly teach phonics to their children (Ehri & Flugman, 2017). After a year of professional development and mentoring in phonics instruction, these teachers improved both their instruction and reading outcomes for their students. In order to both understanding early elementary school students reading development and to understand the impact of curriculum on this development, future research should focus on studying how curriculum is implemented in the classroom. This will help us understand not just the content teachers are expected to teach, but also how that content is presented.

Conclusions and Implications

The goal of early identification of dyslexia is to ensure that parents and educators have accurate and useful information about the potential academic challenges their students may face, so that those parents and educators may make informed decisions about instruction and, when necessary, interventions. The analyses presented in this study supports previous work that assessments of phonology and automaticity are predictive of reading ability even when administered before a child is able to read a single word. Furthermore, these results show, that students who struggle with phonology, or both phonology and automaticity, are on different trajectories in terms of learning to read in early elementary school. As these students progress into third grade they will transition into a new pedagogical period where learning to read is supplanted by reading to learn (Snow & Mathews, 2016). The students who are behind are at a disadvantage not just in their language arts programs, but in their education as a whole. A lower growth rate in word reading means less access to the information that these students need to thrive academically.

Screening prior to entering elementary school offers the possibility of students having access to services that will allow them to become proficient readers. But early screening is only useful if it is interpreted properly. Those using these measures as screening tools must understand how to interpret the results for screening assessments for dyslexia, what dyslexia is, and how to provide appropriate instruction for those who are at risk for it. Educators must understand that early screening is not the same as diagnosis. Although the findings of this study and others show that phonological awareness and rapid automatized naming are predictive of ability reading outcomes and trajectories, these predictions are not deterministic. Simply because on average students have low word reading growth rates when they perform in the bottom quartile on measures of phonological awareness and automaticity, does not mean each identified student will have lower rates of growth. Screening tools can help identify students who are at risk so that educators can provide the necessary resources for these students, but these tools do not replace diagnostic measures that will confirm or negate the results of screening measures as the students grow older.

Beyond understanding how to interpret screening assessments, educators must understand what dyslexia is and how to provide effective instruction for those who are at risk for it. Although the lowest estimates of dyslexia's prevalence are around one in twenty, meaning that most classes in the Unites States will have at least one dyslexic student (Elliott & Grigorenko, 2014), many teachers can not define dyslexia. One survey of teachers entering the workforce in the United Kingdom and the United States found that the majority of respondees believed dyslexia is caused by a problem with the visual perception of words (Washburn et al., 2017). In another study, 42% of surveyed graduate and undergraduate education students believed that dyslexia was primarily characterized by a deficit in ability to comprehend text (Washburn et al., 2013). Providing an educator with the results of screening assessments indicating risk for dyslexia, without ensuring that educators understand the meaning of dyslexia, would be useless.

Finally, it is essential for teachers to understand the unique instructional needs of students who are at risk for dyslexia. As discussed above in the literature review, there is substantial evidence that sub-lexical instruction can effectively remediate reading difficulties for students with dyslexia (Tannenbaum et al., 2006; Berninger et al., 2001; Wolf & Katzir-Cohen, 2001).

After viewing multiple studies on the efficacy of various reading instructional programs, Torgesen (2000) found that programs focused on phonics instruction were most effective at helping students catch up to their peers. There is evidence that teachers do not understand the necessity for sub-lexical instruction (Ehri & Flugman, 2017). In order to ensure that the students who are identified as at risk for dyslexia receive the appropriate instruction, teachers must why and how understand this instruction is given.

References

Anderson, K (2000, June 18). The reading wars – A look at the report of the National Reading Panel, "Teaching Children to read," and the debate over what reading strategy works best in the classroom, *Los Angeles Times,* p. 5.

Bartl-Pokorny, K., Marschik, P., Sachse, S., Green, V., Zhang, D., Meer, L., Wolin, T. Einspieler, C. (2013). Tracking development from early speech-language acquisition to reading skills at age 13. *Developmental Neurorehabilitation, 16*(3), 188-195.

Beneventi, H., Tønnessen, F., Ersland, L., & Hugdahl, K. (2010). Executive working memory processes in dyslexia: Behavioral and fMRI evidence. *Scandinavian Journal of Psychology*, (51), 192-202.

Berninger, V. W., Abott, R. D., Billingsley, F., & Nagy, W. (2001). Processes underling time and fluency of reading. In M. Wolf, *Dyslexia, Fluency, and the Brain* (pp. 385-401). Timonium, Maryland: York Press.

Bowers, P. G., Sunseth, K., & Golden, J. (1999). The route between rapid naming and reading progress. *Scientific Studies of Reading,3*(1), 31-53. doi:10.1207/s1532799xssr0301_2

Bowers, P. G., Sunseth, K., & Golden, J. (1999). The Route Between Rapid Naming and Reading Progress. *Scientific Studies of Reading,3*(1), 31-53. doi:10.1207/s1532799xssr0301_2

Carroll, J. M., Solity, J., & Shapiro, L. R. (2015). Predicting dyslexia using prereading skills: The role of sensorimotor and cognitive abilities. Journal of Child Psychology and Psychiatry,57(6), 750-758. doi:10.1111/jcpp.12488

Christodoulo, J., Del Tufo, S., Lymberis, J., Saxler, P., Ghosh, S., Triantafyllou, C., Whitfield-Gabrieli, D.E., Gabrieli, J. (2014). Bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLoS ONE*, 9(7). Retrieved January 1, 2015, from www.polsone.org

Clemens, N. H., Shapiro, E. S., Wu, J., Taylor, A. B., & Caskie, G. L. (2012). Monitoring Early First-Grade Reading Progress. Journal of Learning Disabilities,47(3), 254-270. doi:10.1177/0022219412454455

Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A., . . . Crouch, R. C. (2010). Selecting at-risk first-grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. Journal of Educational Psychology, 102(2), 327-340. doi:10.1037/a0018448

Crisp, J., Howard, D., & Ralph, M. A. (2011). More evidence for a continuum between phonological and deep dyslexia: Novel data from three measures of direct orthography-to-phonology translation. *Aphasiology,25*(5), 615-641. doi:10.1080/02687038.2010.541470

Dandache, S., Wouters, J., Ghesquière, P. (2014). Development of reading and phonological skills of children at family risk for dyslexia: A longitudinal analysis from kindergarten to sixth grade. Dyslexia, 20 (4), 305-329 - 10.1002/dys.1482

Daryl, M.F. (2017). RTI in early intervention: promise and vulnerabilities. *Perspectives on Language and Literacy*, 43(4), 11-14.

Deary, I. J., Whiteman, M. C., Star, J. M., Whalley, L., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. Journal of Personality and Social Psychology, 86, 130–147.

Decoding Dyslexia Massachusetts (2017). MA dyslexia legislation. Retrieved January 23, 2018, from http://www.decodingdyslexiama.org/ma-dyslexia-legislation.html

Ehri, L. C., & Flugman, B. (2017). Mentoring teachers in systematic phonics instruction: Effectiveness of an intensive year-long program for kindergarten through 3rd grade teachers and their students. *Reading and Writing,31*(2), 425-456. doi:10.1007/s11145-017-9792-7

Elliott, J.G., & Grigorenko, E.L. (2014). *The dyslexia debate*. New York: Cambridge University Press.

Enders, C. K. (2010). *Applied missing data analysis*. NewYork: The Guilford Press.

Farquharson, K., Centanni, T., Franzluebbers, C., & Hogan, T. (2014). Phonological and lexical influences on phonological awareness in children with specific language impairment and dyslexia. *Frontiers in Psychology*, 5(838), 1-10.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M.A. (2007). *Learning Disabilities*. New York: Gilford.

Foorman, B. R., Chen, D.-T., Carlson, C., Moats, L., Francis, D. J., & Fletcher, J. M. (2003). The necessity of the alphabetic principle to phonemic awareness instruction. *Reading and Writing: An Interdisciplinary Journal*, 16, 289-324.

Gathercole, S. E., & Baddeley, A. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, Special Issue: Prediction of Reading and Spelling, 8(3), 259-272.

Goodman, K. S. (1992). Why whole language is today's agenda in education. *Language Arts, 69,* 354-363.

Gresham, F. M., & Vellutino, F. R. (2010). What is the role of intelligence in the identification of specific learning disabilities? Issues and clarifications. *Learning Disabilities Research & Practice,25*(4), 194-206. doi:10.1111/j.1540-5826.2010.00317.x

Harlaar, N., Hayiou-Thomas, M., Dale, P., & Plomin, R. (2008). Why do preschool language abilities correlate with later reading? A twin study. *Journal of Speech Language and Hearing Research,* 688-688.

Helland, T., Plante, E., & Hugdahl, K. (2011). Predicting Dyslexia at Age 11 from a Risk Index Questionnaire at Age 5. Dyslexia,17(3), 207-226. doi:10.1002/dys.432

Hoskyn, M., & Swanson, H.L. (2000). Cognitive processing for low achievers and children with learning disabilities: A selective meta-analytic review of the published literature. *School Psychology Review, 29,* 102-119.

Institute for Educational Sciences, *What Works Clearinghouse*., ies.ed.gov/ncee/wwc/FWW/Results?filters=%2CLiteracy.

Katzir, T., & Pare-Blagoev, J. (2006). Applying cognitive neuroscience research to education: The case of literacy. *Educational Psychologist*, 4(1), 53-74.

Kaufman, A.S., & Kaufman, N.L. (2004). Kaufman Brief Intelligence Test. *Wiley Online Library*.

Kutner, M. A., Greenberg, E., Jin, Y., Boyle, B., Hsu, Y., & Dunleavy, E. (2007). *Literacy in everyday life: results from the 2003 National Assessment of Adult Literacy* (United States, National Assessment of Adult Literacy). Washington, DC: National Center for Education Statistics.

Lerner, Matthew D., and Christopher J. Lonigan. (2016) Bidirectional relations between phonological awareness and letter knowledge in preschool revisited: A growth curve analysis of the relation between two code-related skills. *Journal of Experimental Child Psychology*, vol. 144, 2016, pp. 166–183., doi:10.1016/j.jecp.2015.09.023.

Lonigan, C., Schatschneider, C., & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing and spelling. In *Developing early literacy: report of the National Early Literacy Panel* (pp. 55-79). Washington, D.C.: National Institute for Literacy.

Maehler, C., & Schuchardt, K. (2016). Working memory in children with specific learning disorders and/or attention deficits. *Learning and Individual Differences,49*, 341-347. doi:10.1016/j.lindif.2016.05.007

Miller, C., Miller, S., Bloom, J., Jones, L., Lindstrom, W., Craggs, J., Garcia-Barrera, M., Semrud-Clikeman, M., Gilger, J., Hynd, G. (2006). Testing the double-deficit hypothesis in an adult sample. *Annals of Dyslexia, 56*(1), 83-102.

Nagamine, M., Black, J., Mazaika, P., Tanaka, H., Stanley, L., Heitzmann, J., . . . Hoeft, F. (2009). Neural basis of phonological processing in kindergarten children at risk for dyslexia. *NeuroImage,47*. doi:10.1016/s1053-8119(09)70810-9

National Center for Education Statistics. "Reading Assessment." United States Department of Education. Web. 4 December 2016. <http://nces.ed.gov/nationsreportcard/reading/>.

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

O'Connor, R. E., White, A., & Swanson, H. L. (2007). Repeated reading versus continuous reading: Influences on reading fluency and comprehension. *Exceptional Children, 74*(1), 31-46. doi:10.1177/001440290707400102

O'Donnell, P. S., & Miller, D. N. (2011). Identifying students with specific learning disabilities: School psychologists' acceptability of the discrepancy model versus response to intervention. *Journal of Disability Policy Studies,22*(2), 83-94. doi:10.1177/1044207310395724

Owens, R. E. (2016). *Language development: an introduction*. Harlow: Pearson Education.

Ozernov-Palchik, O., Norton, E. S., Sideridis, G., Beach, S. D., Wolf, M., Gabrieli, J. D., & Gaab, N. (2016). Longitudinal stability of pre-reading skill profiles of kindergarten children: implications for early screening and theories of reading. *Developmental Science*, 20(5). doi:10.1111/desc.12471

Pennington, B. F., & Lefly, D. L. (2001). Early Reading Development in Children at Family Risk for Dyslexia. Child Development,72(3), 816-833. doi:10.1111/1467-8624.00317

Pennington, B.F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, 101, 385–413.

Pennington, B.F., & Lefly, D.L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, 72, 816–833.

Pinker, S. (1994). *The language instinct: the new science of language and mind*. London: Alle Lane.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology,98*(1), 46-68. doi:10.1016/j.jecp.2007.04.003

Ravitch, D. (2014). *Reign of error: the hoax of the privatization movement and the danger to Americas public schools*. New York: Vintage Books.

Rose, J. (2009). *Identifying and teaching children and young people with dyslexia and literacy difficulties. (The Rose Report)*. Nottingham, UK: DCSF Publications.

Rubin, D.B. (1976). Inference and missing data. *Biometicka*, 63, 581-592.

Rutter, M., Kim-Cohen, J., & Maughan, B. (2006). Continuities and discontinuities in psychology between childhood and adult life. *Journal of Child Psychology and Psychiatry,* 47, 276-295.

Schaars, M. M., Segers, E., & Verhoeven, L. (2017). Word Decoding Development during Phonics Instruction in Children at Risk for Dyslexia. *Dyslexia, 23*(2), 141-160. doi:10.1002/dys.1556

Shaywitz, S. (2003). *Overcoming dyslexia: A new and complete science-based program for reading problems at any level*. New York: A.A. Knopf.

Snow, C. E., & Matthews, T. J. (2016). Reading and language in the early grades. *The Future of Children,26*(2), 57-74. doi:10.1353/foc.2016.0012

Snyder, I. (2008). *The literacy wars: why teaching children to read and write is a battleground in Australia*. Crows Nest, NSW, Australia: Allen & Unwin.

Stanovich, K.E. (1994). Explaining the differences between the dyslexic and the garden variety reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, 21, 590-604.

Stothard, S., Hulme, C., Clarke, P., Barmby, P., & Snowling, M. (2010). YARC York Assessment of Reading for Comprehension (Secondary). London: G.L. Assessment.

Swanson, H. L., Harris, K. R., & Graham, S. (2014). *Handbook of learning disabilities*. New York: The Guilford Press.

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, *10*(4), 381-398.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education, 25*(4), 252-261.

Thompson, P. A., Hulme, C., Nash, H. M., Gooch, D., Hayiou-Thomas, E., & Snowling, M. J. (2015). Developmental dyslexia: Predicting individual risk. *Journal of Child Psychology and Psychiatry,56*(9), 976-987. doi:10.1111/jcpp.12412

Thompson, P. A., Hulme, C., Nash, H. M., Gooch, D., Hayiou-Thomas, E., & Snowling, M. J. (2015). Developmental dyslexia: Predicting individual risk. Journal of Child Psychology and Psychiatry,56(9), 976-987. doi:10.1111/jcpp.12412

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

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). Test of Word Reading Efficiency–Second Edition (TOWRE-2). Austin, TX: Pro-Ed.

Varvara, P., Varuzza, C., Sorrentino, A. C., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience,* 8(120). doi:10.3389/fnhum.2014.00120

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. doi:10.1046/j.0021-9630.2003.00305.x

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1999). CTOPP: Comprehensive Test of Phonological Processing. Austin, TX: PRO-ED.

Washburn, E. K., Binks-Cantrell, E. S., & Joshi, R. M. (2013). What Do Preservice Teachers from the USA and the UK Know about Dyslexia? *Dyslexia,20*(1), 1-18. doi:10.1002/dys.1459

Washburn, E., Mulcahy, C., Musante, G., & Malatesha Joshi, R. (2017). Novice Teachers' Knowledge of Reading-related Disabilities and Dyslexia. *Learning Disabilities: A Contemporary Journal,15*(2), 169-191.

Wexler, J., Vaughn, S., Roberts, G., & Denton, C. A. (2010). The efficacy of repeated reading and wide reading practice for high school students with severe reading disabilities. *Learning Disabilities Research & Practice*, 25, 2–10. doi: 10.1111/j.15 405826.2009.00296.x

Wiederholt, J. L., & Bryant, B. R. (2012a). *Gray oral reading test: Examiner's manual* (5th ed.). Austin, TX: Pro-Ed.

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

Wolf, M., & Gottwald, S. (2016). *Tales of literacy for the 21st century*. Oxford, United Kingdom: Oxford University Press.

Wolf, M., & Stoodley, C. (2007). *Proust and the squid: The story and science of the reading brain*. New York, NY: HarperCollins

Wolf, M., Bowers, P.G., & Briddle, K. (2000). Naming-speed processes for the developmental dyslexia. *Journal Educational Psychology,* 91, 415-438.

Wolf, M., Golberg O'Rouke, 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*.

Woodcock, R. (2011). Woodcock Reading Mastery Test (WRMT-III). San Antonio, TX: Pearson.

Table 1
Descriptive Statistics for Continuous Variables

Variable	N	Mean	SD	Range		SE	Skew	Kurtosis
				Min	Max			
IQ	161	99.01	9.72	80	131	0.77	0.59	0.12
Age Pre-K (months)	161	67.22	3.94	60	77	0.31	0.14	-0.84
Age Grade 1 (months)	151	86.84	4.21	78	100	0.34	0.32	-0.35
Age Grade 2 (months)	146	99.08	4.29	90	115	0.35	0.58	0.63
Word ID Pre-K	129	1.58	3.77	0	27	0.33	4.37	23.52
Word ID Grade 1	151	21.04	5.78	5	37	0.47	-0.07	-0.23
Word ID Grade 2	145	24.26	5.19	8	36	0.43	-0.35	0.20
Word ID All Grades	425	16.23	10.99	0	37	0.53	-0.37	-1.29
RAN Pre-K	93	106.11	14.33	54	135	1.49	-0.82	1.03
RAN Grade 1	151	105.34	12.37	79	136	1.01	-0.02	-0.38
RAN Grade 2	144	102.46	12.25	77	134	1.02	0.21	-0.52
PA Blending Pre-K	148	10.30	2.28	6	17	0.19	-0.05	-0.40
PA Blending Grade 1	151	11.16	2.10	8	18	0.17	0.89	0.30
PA Blending Grade 2	146	10.36	2.17	6	16	0.18	0.58	-0.27
PA Working Memory Pre-K	158	8.79	2.55	4	17	0.20	0.55	0.10
PA Working Memory Grade 1	151	8.99	1.97	4	15	0.16	0.26	-0.14
PA Working Memory Grade 2	145	9.03	2.17	4	15	0.18	0.23	-0.22
PA Elision Pre-K	150	9.85	2.18	6	16	0.18	0.42	-0.57
PA Elision Grade 1	151	11.85	2.94	7	19	0.24	0.38	-0.90
PA Elision Grade 2	146	11.27	2.88	5	16	0.24	-0.17	-1.16

Table 2:
Pearson's Correlations Among Continuous Variables

	Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	IQ	–														
2.	Word ID (Pre-K)	0.04	–													
3.	Word ID (First)	0.15	0.41*	–												
4.	Word ID (Second)	0.05	0.34*	0.85*	–											
5.	RAN (Pre-K)	0.00	0.43*	0.49*	0.41*	–										
6.	RAN (First)	0.13	0.02	0.33*	0.23*	0.54*	–									
7.	RAN (Second)	0.08	0.08	0.31*	0.22*	0.51*	0.66*	–								
8.	PA Blending (Pre-K)	0.19*	0.36*	0.42*	0.39*	0.18	0.05	0.02	–							
9.	PA Blending (First)	0.20*	0.25*	0.35*	0.38*	0.05	0.16	0.14	0.49*	–						
10.	PA Blending (Second)	0.18*	0.35*	0.31*	0.36*	0.11	0.08	0.06	0.47*	0.55*	–					
11.	PA Working Memory (Pre-K)	0.11	0.12	0.19*	0.19*	-0.11	-0.15	-0.08	0.44*	0.18*	0.30*	–				
12.	PA Working Memory (First)	0.16	0.03	0.39*	0.33*	0.13	0.13	0.19*	0.46*	0.32*	0.26*	0.32*	–			
13.	PA Working Memory (Second)	0.11	0.34*	0.40*	0.42*	0.28*	0.06	0.06	0.52*	0.42*	0.55*	0.38*	0.46*	–		
14.	PA Elision (Pre-K)	0.22*	0.34*	0.45*	0.41*	0.26*	0.07	0.07	0.62*	0.44*	0.38*	0.26*	0.37*	0.42*	–	
15.	PA Elision (First)	0.22*	0.20*	0.57*	0.57*	0.26*	0.24*	0.24*	0.42*	0.41*	0.40*	0.35*	0.34*	0.43*	0.44*	–
16.	PA Elision (Second)	0.18*	0.27*	0.45*	0.48*	0.22*	0.10	0.10	0.38*	0.37*	0.39*	0.23*	0.24*	0.36*	0.44*	0.66*

* $p > .05$

Table 3

T-tests of Word Identification by Dyslexia Risk Group

	No Risk (N = 96)		Risk (N = 86)		t-test		
Measures	M	SD	M	SD	statistic	df	p-value
Word ID (Pre-K)	2.33	4.12	0.77	3.19	2.41	123.14	0.02
Word ID (Grade 1	22.48	5.82	19.54	5.39	3.22	148.80	0.00
Word ID (Grade 2)	25.55	5.13	22.83	4.91	3.27	142.58	0.00

Table 4

ANOVA of Word Identification by Risk Profile Group

	DF	Sum of Squares	Mean Squared	eta squared	F-value	p-value
	Word ID (Pre-K)					
Between groups	3	101.6	146.30	0.06	2.47	0.07
Within groups	125	1717.8	24.40			
Total	128	1819.4				
	Word ID (Grade 1)					
Between groups	3	659	219.79	0.13	7.41	0.00
Within groups	147	4360	29.66			
Total	150	5019				
	Word ID (Grade 2)					
Between groups	3	439	146.30	0.11	6.00	0.00
Within groups	141	3441	24.40			
Total	144	3880				

Table 5

Cross Tabulations of mean Word Identification by Dyslexia Profile Group

		Word ID (Mean)		
Deficit Risk Group	N	Pre-K	Grade 1	Grade 2
No Deficit	81	2.33	22.48	25.55
RAN Deficit	32	1.54	21.30*	24.59
PA Deficit	31	0.19	19.86	22.00*
Double Deficit	17	0.50	15.69*	20.71*

* Post hoc Tukey test showed significant difference compared with No Deficit group mean [$F(3,125) = 6.00$ $p < .05$].

* Post hoc Tukey test showed difference compared with No Deficit group mean [$F(3,125) = 6.00$ $p < .05$].

Table 6

Estimates from model for rate of Word ID growth

Fixed Effects	Coefficient	S.E.	p-value
Intercept (γ_{00})	-6.47	0.563	0.00
Age (γ_{10})	1.37	0.057	0.00
Age squared (γ_{20})	-0.01	0.001	0.00
Random Effects	**Parameter**		
τ^2_{00}	0.29		
τ^2_{10}	0.01		
τ^2_{20}	3.62		
σ^2	3.35		
Deviance (-2LL)	1469.639		

Table 7

Estimates from model for dyslexia risk rate of Word ID growth

Fixed Effects	Coefficient	S.E.	p-value
Intercept (γ_{000})	-6.49	0.567	0.00
Risk for Dyslexia (γ_{010})	0.01	0.061	0.83
Age (γ_{100})	1.38	0.057	0.00
Age * Risk for Dyslexia (γ_{110})	0.01	0.006	0.20
Age Squared (γ_{200})	-0.01	0.001	0.00
Age Squared * Risk for Dyslexia (γ_{220})	0.00	0.000	0.17
Random Effects	**Parameter**		
τ^2_{00}	0.29		
τ^2_{10}	0.01		
τ^2_{20}	3.66		
σ^2	3.35		
Deviance (-2LL)	-1464.901		

Table 8

Estimates from model for dyslexia risk rate of Word ID growth with phonics curriculum

Fixed Effects	Coefficient	S.E.	p-value
Intercept (δ_{000})	-7.72	1.649	0.000
Curriculum (δ_{001})	0.09	2.230	0.039
Risk for Dyslexia (δ_{010})	2.66	1.983	0.199
Curriculum * Risk for Dyslexia (δ_{011})	-2.09	2.635	0.428
Age (δ_{100})	1.36	0.152	0.000
Age * Risk for Dyslexia (δ_{110})	0.19	0.215	0.384
Age * Curriculum (δ_{101})	0.12	0.188	0.521
Age * Risk for Dyslexia * Curriculum (δ_{011})	-0.32	0.256	0.218
Age Squared (δ_{200})	-0.01	0.003	0.000
Age Squared * Risk for Dyslexia (δ_{200})	0.00	0.004	0.456
Age Squared * Curriculum (δ_{201})	-0.01	0.004	0.208
Age Squared * Risk for Dyslexia * Curriculum (δ_{211})	0.01	0.005	0.172
Random Effects	**Parameter**		
Intercept τ^2_{U0}	12.30		
τ^2_{U1}	0.06		
τ^2_{U2}	0.00		
τ^2_{V00}	2.25		
τ^2_{V10}	0.02		
τ^2_{V20}	0.00		
σ^2	11.97		
Deviance (-2LL)	1439.85		

Table 9

Estimates from model for dyslexia risk profile rate of Word ID growth

	PA Risk Added			RAN Risk Added			Risk Interaction Added		
Fixed Effects	Coefficient	S.E.	p-value	Coefficient	S.E.	p-value	Coefficient	S.E.	p-value
Intercept (γ_{000})	-8.95	0.818	0.000	-7.75	0.975	0.000	-7.76	0.932	0.000
Risk for PA Deficit (γ_{010})	1.14	1.299	0.382	0.10	1.339	0.940	-0.47	1.562	0.765
Risk for RAN Deficit (γ_{020})				-0.33	1.363	0.812	-1.14	1.588	0.474
Risk for Double Deficit (γ_{030})							0.60	2.094	0.775
Age (γ_{100})	1.69	0.080	0.000	1.62	0.092	0.000	1.60	0.087	0.000
Age * Risk for PA Deficit (γ_{110})	-0.34	0.111	0.002	-0.20	0.118	0.084	-0.02	0.136	0.903
Age * Risk for RAN Deficit (γ_{120})				-0.15	0.119	0.206	0.04	0.136	0.778
Age * Risk for Double Deficit (γ_{130})							-0.55	0.182	0.003
Age Squared (γ_{200})	-0.02	0.002	0.000	-0.02	0.002	0.000	-0.02	0.002	0.000
Age Squared * Risk for PA Deficit (γ_{210})	0.01	0.002	0.017	0.00	0.002	0.305	0.00	0.003	0.373
Age Squared * Risk for RAN Deficit (γ_{220})				0.00	0.002	0.148	0.00	0.003	0.667
Age Squared * Risk for Double Deficit (γ_{230})							0.01	0.004	0.007
Random Effects	Parameter			Parameter			Parameter		
Intercept τ^2_{U0}	30.87			29.13			30.55		
τ^2_{U1}	0.12			0.12			0.12		
τ^2_{U2}	0.00			0.00			0.00		
τ^2_{V00}	2.63			4.03			4.09		
τ^2_{V10}	0.05			0.05			0.05		
τ^2_{V20}	0.00			0.00			0.00		
σ^2	5.87			6.86			6.61		
Deviance (-2LL)	1422.33			1428.00			1424.42		

Table 10

Estimates from model for dyslexia risk profile rate of Word ID growth with phonics curriculum

Fixed Effects	Coefficient	S.E.	p-value
Intercept (δ_{000})	-9.70	1.770	0.000
Curriculum (δ_{001})	1.43	3.727	0.701
Risk for PA Deficit (δ_{010})	1.35	2.137	0.537
Risk for RAN Deficit (δ_{020})	3.89	2.719	0.155
Risk for Double Deficit (δ_{030})	-2.80	5.168	0.589
Curriculum * Risk for PA Deficit (δ_{011})	-0.97	4.173	0.816
Curriculum *Risk for RAN Deficit (δ_{021})	-5.77	3.517	0.103
Curriculum * Risk for Double Deficit (δ_{031})	4.65	5.740	0.419
Age (δ_{100})	1.53	0.172	0.000
Age * Risk for PA Deficit (δ_{111})	0.04	0.338	0.917
Age * Risk for RAN Deficit (δ_{121})	-0.04	0.244	0.866
Age * Risk for Double Deficit (δ_{131})	-0.06	0.448	0.890
Age * Curriculum (δ_{101})	0.23	0.214	0.284
Age * Risk for PA Deficit * Curriculum (δ_{111})	-0.30	0.377	0.425
Age * Risk for RAN Deficit * Curriculum (δ_{121})	-0.11	0.314	0.729
Age * Risk for Double Deficit *Curriculum (δ_{131})	-0.64	0.499	0.202
Age Squared (δ_{200})	-0.02	0.003	0.000
Age Squared * Risk for PA Deficit (δ_{210})	0.00	0.007	0.555
Age Squared * Risk for RAN Deficit (δ_{220})	0.00	0.005	0.812
Age Squared * Risk for Double Deficit (δ_{230})	0.00	0.008	0.976
Age Squared * Curriculum (δ_{201})	0.00	0.004	0.274
Age Squared * Risk for PA Deficit * Curriculum (δ_{211})	0.01	0.008	0.315
Age Squared * Risk for RAN Deficit * Curriculum (δ_{221})	0.00	0.006	0.439
Age Squared * Risk for Double Deficit *Curriculum (δ_{231})	0.01	0.010	0.203

Random Effects	Parameter
Intercept τ^2_{U0}	13.46
τ^2_{U1}	0.08

τ^2_{U2}	0.00
τ^2_{V00}	2.69
τ^2_{V10}	0.04
τ^2_{V20}	0.00
σ^2	8.75
Deviance (-2LL)	1429.74

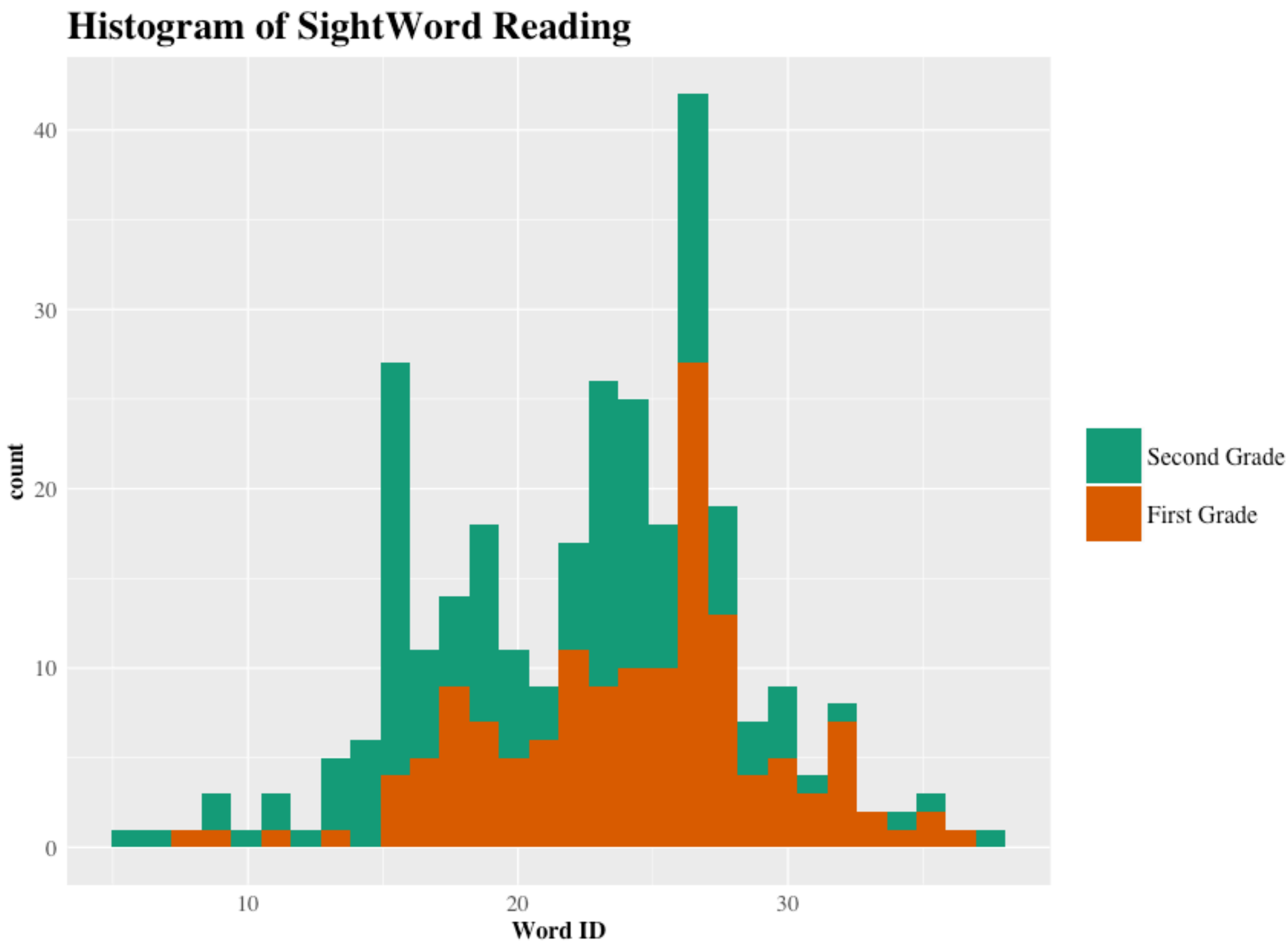

Figure 1: Histogram of Word ID scores in first and Grade 2

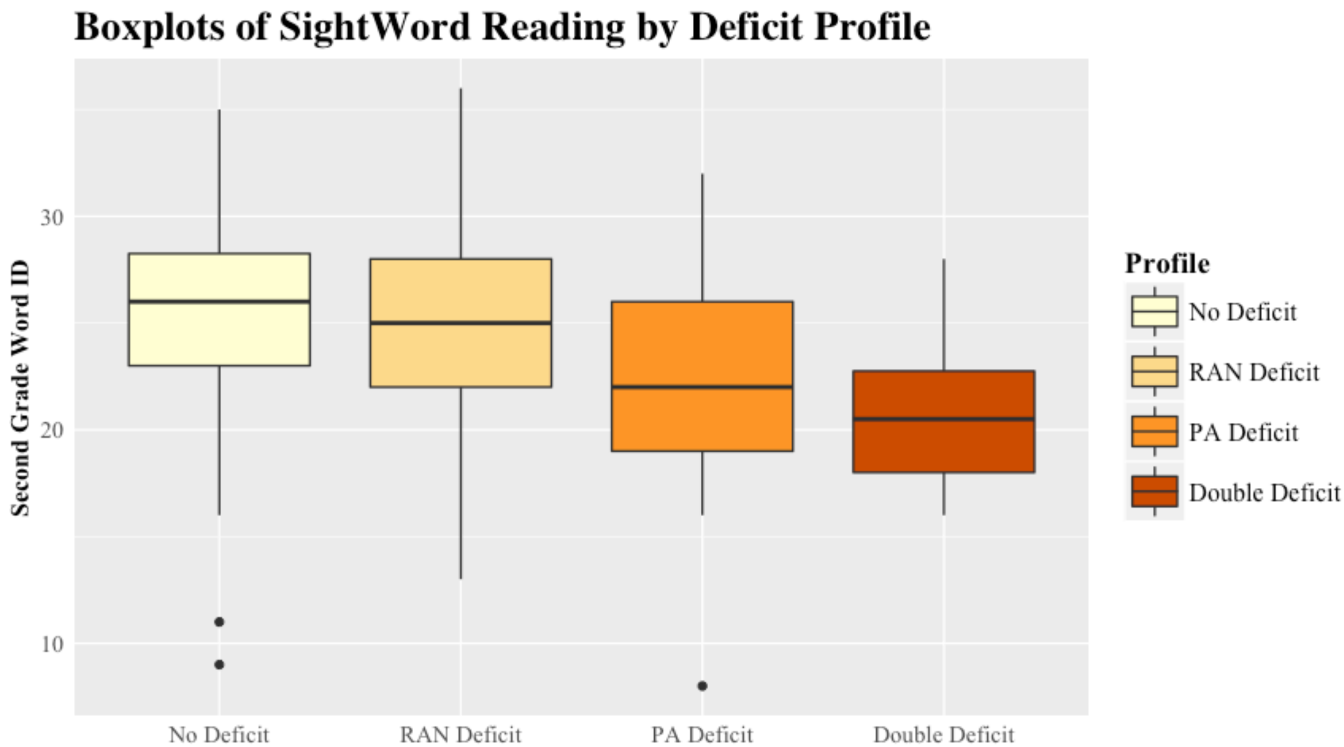

Figure 2: Boxplots of Grade 2 Word ID grouped by student and deficit profile

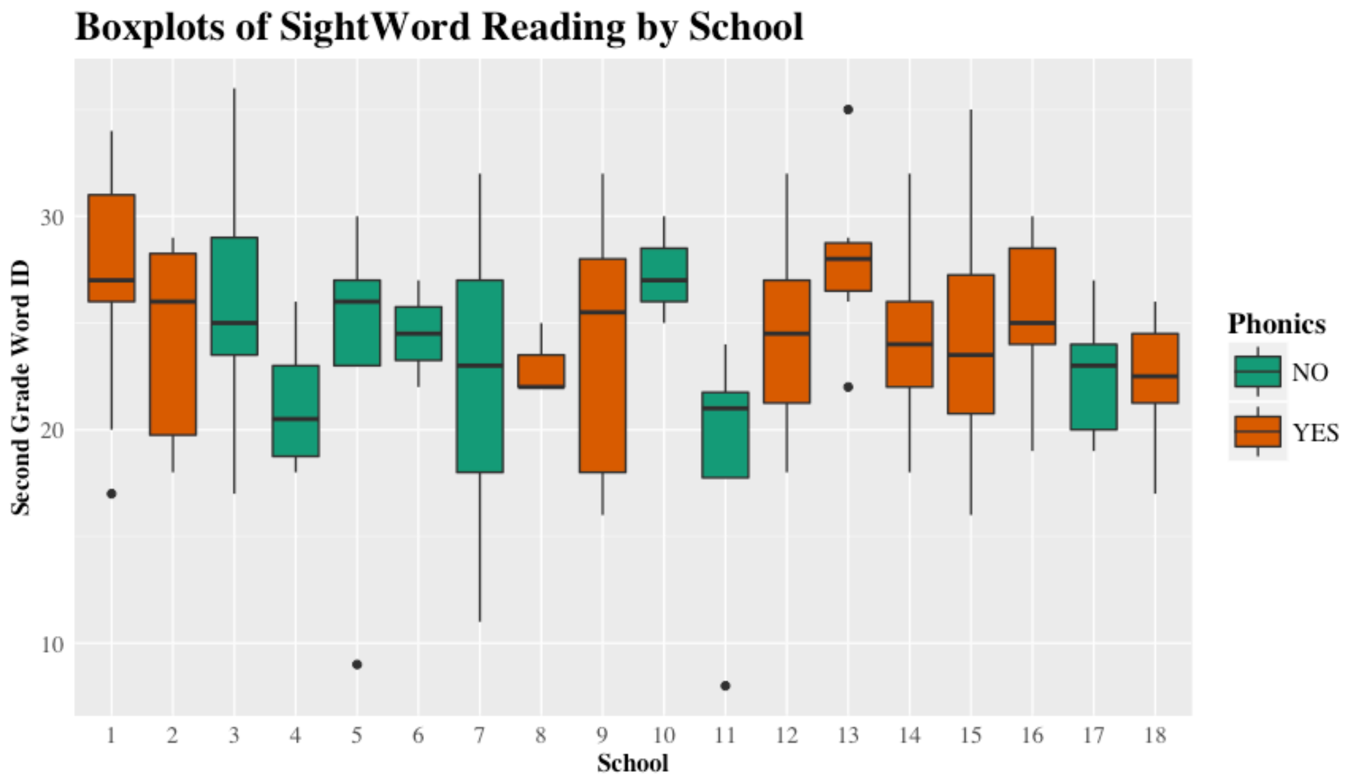

Figure 3: Boxplots of fluency grouped by school and school curriculum type