11-7-2015

Using Curriculum-Based Measurement in the Assessment of Reading Disabilities

Jessica Menard
University of Windsor

Using Curriculum-Based Measurement in the Assessment of Reading Disabilities

by
Jessica Menard

A Dissertation
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada

Using Curriculum-Based Measurement in the Assessment of Reading Disabilities

by

Jessica Menard

APPROVED BY:

_____

Dr. John Kirby
Queen's University

_____

Dr. Daniel Mennill
Biological Sciences

_____

Dr. Erin Picard
Psychology

_____

Dr. Dennis Jackson
Psychology

_____

Dr. Carlin J. Miller
Psychology

_____

Dr. Glenn Rideout, Chair of Defense
Associate Dean, Graduate Studies and Research

8 April 2015

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this dissertation that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my dissertation does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my dissertation and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my dissertation, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this dissertation has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

The present investigation looked at students' reading achievement within the context of the Peer Assisted Learning Strategies (PALS) intervention. It consisted of three separate studies, all of which are related to reading achievement and intervention during the early years of school. The purpose of Study One was to determine whether students who are identified with reading disabilities via psychological assessment report make improvements over the school year subsequent to the implementation of this report. It was hypothesized that when teachers have access to psychological assessment reports, they will better understand their students' individual learning needs and that this will translate to improved scores in reading. This hypothesis was not supported; those students who underwent psychological assessment did not show significant improvement in their reading skills as compared to students who did not undergo psychological assessment. Study Two examined whether the reading skills of students who are considered low achievers in reading tend to regress to a greater extent during the traditional summer vacation, as compared to their high- and typically-peers, whether it takes the low achievers longer to recover from summer loss, and whether they show more shallow learning trajectories over the school year. The summer learning loss hypothesis was partially supported. In terms of summer learning loss, on a measure of word reading administered following the summer after Senior Kindergarten, the low achievers' scores remained stable over the summer, while the average and high achievers' scores increased. It is thought that the Grade 1 year marks an important time for the onset of summer learning loss as a phenomenon. Study Three assessed the role of language prosody as a predictor of reading outcomes within the PALS intervention. Language prosody was not

found to be a significant predictor of progress in PALS. The utility of curriculum-based measurements in the assessment of reading disabilities in a Canadian context is discussed.

ACKNOWLEDGEMENTS

Successful completion of this dissertation work would not have been possible without significant contributions by the following people:

- The Government of Ontario for the provision of the Ontario Graduate Scholarship as well as the Department of Psychology at the University of Windsor for providing funding for my work.
- My research supervisor, Dr. Carlin Miller. Thank you for all of your guidance and support over these past seven years. You've inspired my interest in research by modelling that passion yourself. Even when it waned (as I have come to believe is a natural part of the dissertation process), you helped me to rediscover my interest in this work through your support and encouragement.
- My contact at the school board and committee member, Dr. Erin Picard. Thank you for your help with accessing the data. Your insight and patience in working with me was an invaluable part of this experience. Thank you so much for your help with this work. Thank you as well to the school board staff who work tirelessly at teaching the students to read. And special thanks to the principals and vice-principals who let me come into their schools to collect data, the teachers who put up with me constantly running students back and forth to and from their classrooms, and to the students whose data comprise my thesis, especially those with whom I worked personally in the process of data collection.
- My stats consultant and committee member, Dr. Dennis Jackson. Thank you for the many statistical consulting meetings in which you engaged me in discussion on equal parts statistics and X-Files fanship, and for your flexibility in agreeing to step into an already-established committee.
- My external-internal committee member, Dr. Dan Mennill. Thank you for your strong scientific take on my work. Your insights and questions significantly strengthened the science of the work that I undertook. Thank you also for your careful attention to detail, and for helpfully pointing out my many typographical errors; this also significantly strengthened my document by ensuring that it was as free from errors as possible.
- My external examiner, Dr. John Kirby. Thank you for taking the time to visit Windsor to attend my defense in person, and for your insights before and during my defense meeting that will contribute considerably to the quality of my work. Thank you also for engaging me in a lively and thought-provoking discussion about the reading literature.
- Hands-down the BEST grad secretary the world has ever known, Barb Zakoor. Thank you for making that speech I "volunteered" to give at your retirement party super easy to write by legitimately being excellent at everything always.

- My husband, Avery McCarthy. Thank you for your nearly-unlimited patience with me during this process and for countless hours of tech support on Excel. Thanks also for "wife-ing" me, I s'pose.

- My parents (Mom, Dad, and Viv). Thank you for taking the time to attend my defense so that you could get a better sense of what I've spent the past seven years working to accomplish and for generally being supportive of my school-related life choices. Get ready to alert the media, Dad, the time is drawing nigh for that announcement in the Brockville newspaper!

- My labmate, Dragana. Thank you for taking time out of your life to do me a solid and deliver copies of my document all over campus. As well, thanks to Brie for being all up in the deposit process.

- My fellow residents, Louise, Kristen, and Ben. You guys have been beyond supportive, thus making internship that much better of a time! As well, my most appreciated supervisor, Emily Case. Thanks for your transcendently insightful narrative reframes of the dissertation process that helped me to focus on the positives.

- The Ro-Ho's, Amanda and Fred. Thanks for putting me up in your extra-comfy bed while I was in Windsor to defend, and for making the whole ordeal that much better of an experience by helping me celebrate it with Newfie screech rum and Sodastream diet cola afterwards.

- The friends that I've made since moving to Windsor. The peer support in this program is incredible, and you guys are no exception (obviously, or I wouldn't be thanking you). Windsor friends, you've helped me to have a grad school experience that was as free as possible from violations to its robustness, as well as rife with instances of feeling VALIDATED. Extra-special shout-outs go to my bestie MegaDuff, the Power Psychology Rangers (SaManDrea), FreManda Ro-Ho again, JRDN, Heather-My-Love, Small-ish Jen (which is a nickname I just made up), and Sandra (who is unfortunately nicknameless, but also no less important because of that). You guys are the best, and in the words of Ms. Kelly Clarkson (but actually Martin, Gottwald, & Kelly, 2008), "my life would suck without you."

References

Martin, M., Gottwald, L., & Kelly, C. (2008). My life would suck without you. Recorded by Kelly Clarkson. On *All I Ever Wanted* [CD]. New York City, NY: RCA Records.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF COMMON ABBREVIATIONS

Abbreviation	Definition
PALS	Peer Assisted Learning Strategies
CBMs	Curriculum-based measurements
LSF	Letter-Sound Fluency
WIF	Word Identification Fluency
PRF	Passage Reading Fluency
LD	Learning disabilities
RTI	Responsiveness-to-Intervention, or Response-to-Intervention
SD	Standard deviation
PEPS-C	Profiling Elements of Prosodic Systems – Child Version
WECDSB	Windsor-Essex Catholic District School Board

CHAPTER I

Literature Review

Goal and Purpose of the Present Dissertation Projects

The goal of this dissertation was to investigate several issues having to do with reading achievement and development in school-aged children. Three main research questions were addressed and it is hoped that these answers will lead to positive changes and further developments in reading instruction and achievement, both locally, and in the broader context of Canadian education. Aside from answering the three research questions directly, a secondary objective of the current work is to provide practical suggestions for helping struggling readers to catch up. This was one of the main motivations for conducting this dissertation research.

Organization of the Literature Review

The present discussion is a review of the relevant literature pertaining to the development of students' ability to read, in addition to current controversies within the literature on reading achievement, reading disabilities, and reading intervention. This discussion will begin by reviewing the process by which typically-developing readers acquire this skill and will then move on to discuss reading disabilities and the problem areas that are regularly seen in students whose reading is disordered. Next, discussions of reading instruction in general and interventions for students with reading problems will be presented. Finally, a brief overview will be given of the three studies that were conducted. It should be noted that there may be some redundancy in the information presented in this chapter and subsequent chapters. This is because each chapter is intended to be a stand-alone piece for eventual submission for publication.

Typical Acquisition of Reading Skills

Bottom-Up vs. Top-Down Processing

Knowing how to read well is an undeniably important skill set to possess, yet reading is also a very complicated process to master. It is assumed that reading achievement is dependent on a student's level of proficiency in areas such as phonological processing, phonological awareness, reading fluency, vocabulary, short-term memory, and comprehension, but it is less clear how all of these parts fit together in order to allow the student to read connected text proficiently and to use the information that is read to foster further learning. In the past, the thinking on how students acquire normal reading skills has divided the process into top-down and bottom-up conceptualizations. In a bottom-up approach, the student first takes in the orthographic information that is on the page, analyzing each component part and putting each letter, word, sentence, and so on together, in an effort to arrive at the larger picture; comprehension is built from the component parts (Otto, 1982). Orthography refers to the correspondence between phonemes (spoken sounds) and graphemes (their written representations) that is specific to individual languages. Orthographic information has to do with this grapheme-phoneme, or letter-sound, correspondence. Processing proceeds from the bottom up in that the reader begins by perceiving the smallest elements first, eventually arriving at the larger picture. In a top-down approach to reading, making sense of what is read is guided by the student's prior store of background knowledge. In other words, students make use of the contextual aspects of what they are reading in order to guide them through a given passage, or they start with an idea of what might be written on the page and then use that broad idea to fill in the blanks, using prior knowledge and

memory to make sense of what is written (Kintsch, 2005). Students begin reading by working from the bigger picture and then progress downwards to a more basic level of analysis in order read a given passage. It is likely that a combination of top-down and bottom-up processes are at play in reading, as is proposed by both the dual-route and connectionist models.

The Dual-Route Model of Reading Acquisition

According to a review by Bjaalid, Hoien, and Lundberg (1997), dual-route models of reading acquisition assume that there are two routes by which printed words can be recognized and subsequently understood: the direct route and the indirect route. The direct route, which is also called the visual-orthographic or lexical route, operates by reading words through activation of direct connections between the visual forms of words and their meanings. Ehri (2005) would call this "sight word reading," which refers to the ability to read familiar words by accessing them in memory and automatically associating one's orthographic representation of the word with its meaning without utilizing decoding strategies. Ehri (2005) posits that this is the most efficient, unobtrusive way to read words in connected text. The ability to access words via the direct route is the result of extensive exposure to and practice with written text. In other words, the direct route operates via top-down processing. According to Jobard, Crivello, and Tzourio-Mazoyer's meta-analytic review (2003) of neuroimaging studies, reading that occurs along this route is thought to arise from the co-activation of the Visual Form Word Area (VFWA) of the brain, situated in the left occipitotemporal region, and various semantic areas. This means that a combination of brain activation that is aimed at recognizing the orthographic

patterns associated with whole words, and activation that is aimed at attaching meaning to these words, is required to read via the direct route.

The indirect route, also called the phonological or nonlexical route, involves mediation by the phonological processing system. Letters are sequentially translated into sounds by the application of phonological rules (Bjaalid et al., 1997), and as such, the words are sounded out in an effort to eventually arrive at the meaning of the whole word. Accessing words via the indirect route involves bottom-up processing. According to Jobard and colleagues (2003), the grapho-phonological conversion that is required to read via the indirect route relies on activation in left-hemisphere brain structures, situated in the area commonly referred to as the perisylvian region.

The letter-sound translation associated with indirect-route reading involves more processing than the largely automatic sight word recognition process, so reading via the indirect route is generally slower and reserved for words with which the reader has little or no familiarity. The indirect route is also more heavily relied upon in beginning readers. According to Bjaalid and colleagues (1997), the two routes are somewhat dependent on each other, in that some words can be read via either route, or using a combination of both routes, but the distinction between lexical (direct) and sublexical (indirect) processing remains. Additionally, the phonological recoding ability that is involved in reading via the indirect route is assumed to be a prerequisite for establishing the automatic orthographic connections which underlie efficient direct-route reading.

Connectionist Models of Reading Acquisition

Connectionist models, on the other hand, emphasize a single, richly interconnected system, as opposed to two separate routes, for the recognition and

understanding of all types of words, including words that have been previously encountered (i.e., those accessed via the direct route in the dual-route theory) and real words that have not been seen before in addition to non-words (those which are accessed via the indirect route). Furthermore, connectionist models do not posit the eventual bypass of letter-sound translations for purely direct-route access; instead, phonological decoding is an important part of word identification at all levels of reading development (Bjaalid et al., 1997). Therefore, connectionist models of reading acquisition do not distinguish between two separate routes of processing, but instead posit that reading acquisition occurs as a result of an intimately connected and inseparable combination of bottom-up and top-down processing, viewing the process of learning to read as more of a continuum rather than as a dichotomy.

A Combined Framework for Reading Skills Acquisition

Given the criticisms that dual-route models are largely based on observations of adults with acquired dyslexia following brain injury, and connectionist models are largely based on empirical evidence from already-skilled readers, Bjaalid and colleagues (1997) go on to propose a combined framework for reading acquisition that involves elements of both dual-route- and connectionist-type models and accounts for and explains a broader range of reading acquisition-related concerns. In their combined framework, the three processing systems at play in the dual-route and connectionist theories (i.e., the orthographic, the phonological, and the semantic processors) are united in an integrated system that involves bidirectional associations between these three systems. As well, Bjaalid and colleagues' (1997) model also proposes the presence of two additional systems: the visual and articulatory processors. The visual processor operates during the

initial stage of text processing, and its function is to produce clear, unmasked visual images. The articulatory processor produces speech through neuromuscular activities based on activation of the word's articulatory code, reflecting the close associations between semantic, phonological, and articulatory knowledge, which appear to be activated even during silent reading and thinking (Bjaalid et al., 1997). Hence, Bjaalid and colleagues' (1997) combined framework is based on multiple pathways, most of which run bidirectionally between the five separate processors discussed. Bjaalid and colleagues (1997) recognize that, while their combined framework can account for a number of important issues in reading acquisition, it is not an entirely refutable or testable model. It is presented here as simply one possible illustration of the complexity of reading skills acquisition.

Reading Comprehension

According to theories such as the dual-route, connectionist, and combined framework models, the acquisition of reading ability is a complex process, even for typically-developing readers. Acquisition of word recognition alone is dependent on the capacity of many different and seemingly independent processing systems to work in tandem with each other. If one assumes that reading is a largely bottom-up process, then word recognition is assumed to occur relatively independently of comprehension, whereas if one assumes that reading involves largely top-down processes, then comprehension becomes an important factor in guiding lower-order systems. Based on the evidence previously presented, it is thought that reading involves a combination of both approaches. Therefore, a discussion of reading comprehension is warranted, both from the standpoint of basic reading skills acquisition in a top-down conceptualization,

and also from the standpoint of it being a subsequent step in a bottom-up conceptualization of reading acquisition, after the basic skills involved in reading words have been mastered.

Reading comprehension is complex, and it involves multiple processes, including decoding ability, word reading ability, attention, memory, and vocabulary knowledge (Perfetti, Landi, & Oakhill, 2005). According to Perfetti and colleagues' review (2005), comprehension occurs as the reader builds a mental representation of a text message, and the comprehension processes that give rise to this representation occur at multiple levels across units of language, including the word level (lexical processes), sentence level (syntactic processes), and the more over-arching text level. Across these levels, lower-order processes such as word identification interact with the reader's background knowledge to produce a mental model of the text that the reader can comprehend. It is in combining lower- and higher-level reading abilities that text comprehension is achieved.

Furthermore, according to Perfetti and colleagues (2005), the lower-level reading skills involved in reading comprehension include: (1) decoding and word identification ability; and (2) phonological awareness and processing. As previously discussed, it is necessary that the reader has adequately mastered these concepts in order to read words at all, before becoming able to comprehend what he or she is reading. Included among the mid-level factors at play in reading comprehension, referred to by Perfetti and colleagues (2005) as the "linguistic-conceptual machinery for comprehension" (p. 237) are: (1) syntactic processing, which is the understanding of the grammar of one's native language; (2) working memory systems, which aid the reader in remembering the words within a sentence, retrieving information from the preceding text, parsing the sentence,

and relegating other processes which involve cognitive resources; (3) and the ability to use words to build a conceptual understanding of the text, which is largely reliant on vocabulary and a well-developed store of background knowledge. Included among the higher-level factors discussed by Perfetti and colleagues (2005) are: (1) the ability to make inferences, since text is usually not completely explicit and often requires that leaps be made in order to bridge elements in the text or otherwise support the coherence necessary for comprehension; (2) comprehension monitoring, which refers to the reader's ability to monitor one's his or her own comprehension of the text so that he or she will know when a failure of comprehension has occurred (e.g., an apparent inconsistency) and re-reading is required; and (3) sensitivity to story structure, which refers to the fact that different types of text (based on genre, linguistic style, or layout) can present novel problems that are solved only by experience with reading various text formats. All in all, skilled reading comprehension results when the aforementioned component skills are combined and the reader can impose meaning onto connected text and use this information to aid in further learning.

A Developmental Model of Reading Acquisition

Aside from theories which strive to explain the minute details of the acquisition of reading skills, there are also developmental theories for which the developers aimed to provide more of an overview of how students typically learn to read over time. One example is Chall's (1983) five-step model of the development of reading ability, which proposes that reading occurs in discrete stages that are qualitatively distinct (Kaplan & Walpole, 2005). Chall (1983) posits that reading begins with the acquisition of language skills in the pre-reading stage (stage zero). Then, in the beginning stages of true reading

development (ages 6 to 7), the child must learn the letter-sound relationships in order to decode printed words. Learning the correspondences between letters and sounds is a process which requires increasing proficiency in phonological processing and phonological awareness. Then in the second stage (ages 7 to 8), decoding fluency is gained through practice. The third stage of reading development (ages 9 to 13) marks the transition from "learning to read" to "reading to learn." The child begins to acquire a store of background information and a growing vocabulary through further practice in reading and by reading a wide variety of materials, and in doing so, the child is acquiring new information, thoughts, and ideas through reading. In stage four (ages 14 to 18), this knowledge is compared and evaluated. Different viewpoints and multiple interpretations of the text can be considered in stage four reading. Finally, stage five reading (ages 18 and up) involves the synthesis of information and the formation of advanced-level hypothetical thinking; reading in this stage is constructive, in that the reader can construct knowledge from the text. Snider and Tarver (1987) specify that each stage is dependent on the mastery of the previous one.

According to a more recent review, the National Reading Panel (2000) described a number of areas thought to be involved in successful reading. The first area is phonemic awareness, or the ability to focus on and manipulate the smallest sounds in spoken words (for example, awareness of which words rhyme with each other, knowing that /smile/ without the /s/ sound is /mile/, or knowing that /ship/ is made up of three separate phonemes, /sh/ /i/ /p/). Next, students learn the alphabetic code, learning which letters usually correspond to which phonemes, as well as learning how to apply this knowledge in their reading. This leads to the ability to read new words and to recognize previously

read words. Next, with practice, students develop fluency in reading, which means that they can read more quickly and effortlessly than beginning readers. Finally, comes comprehension, or understanding, of connected text, which is influenced by many factors including vocabulary acquisition and executive skills like working memory, monitoring, and problem-solving.

Reading Problems and Disordered Reading

Definition and Characteristics of Reading Disabilities

While most readers progress along a normal course as outlined above, this is not always true, as in the case of Specific Learning Disabilities (SLD) in reading. In general, the term "learning disability" refers to a category of disorders that affect the acquisition, organization, retention, understanding, or use of verbal or nonverbal information in individuals who otherwise possess at least the average abilities essential for thinking and/or reasoning (Learning Disabilities Association of Canada, 2002). According to the most recent version of the Diagnostic and Statistical Manual (DSM-5, 2013), the essential feature of a Specific Learning Disorder is "difficulties learning and using academic skills… despite the provision of interventions that target those difficulties" (p. 66). It is then specified whether the difficulties in learning and applying academic skills manifest primarily in impairment in reading, written expression, or mathematics, leaving open the opportunity for documenting comorbidity across impairment areas. The severity of the impairment(s) is also specified. Also important to note is the qualifier that "the learning difficulties are not better accounted for by intellectual disabilities, uncorrected visual or auditory acuity, other mental or neurological disorders, psychosocial adversity, lack of proficiency in the language of academic instruction, or inadequate academic instruction"

(p. 67). This qualification is also noted in the definition of a Learning Disability used by the Learning Disabilities Association of Ontario (2011), which states that "these disorders result from impairments in one or more psychological processes related to learning, in combination with otherwise average abilities essential for thinking and reasoning. Learning disabilities are specific, not global, impairments and as such are distinct from intellectual disabilities." A learning disability results from genetic and/or neurological factors, and may interfere with the acquisition and use of one or more of four factors: oral language, reading, written language, or mathematics (Learning Disabilities Association of Canada, 2002), as well as difficulties with organizational skills, social perception, and social interaction (Learning Disabilities Association of Ontario, 2011). The fact that learning disabilities are genetic or neurologically based does not mean that reading and other academic problems cannot be remediated; in fact, it has been shown that research-based practices can improve reading outcomes for children who are in remedial programs, as many students with reading disabilities tend to be (Blachman et al., 2004). Although, according to the Learning Disabilities Association of Canada (2002) and the Learning Disabilities Association of Ontario (2011), a learning disability is a lifelong condition for most people, it can be accommodated to the point where a diagnosis would not be made once strategies for accommodating problems are learned, and given an appropriate match between demands of the environment and the individual's characteristics.

It has been shown that approximately 3% of the Canadian school-aged population (ages 5-14) is affected by learning disabilities (Statistics Canada, 2006). Of those affected by learning disabilities in general, approximately 80% have their primary difficulty in

learning to read (Lyon, 1996). Reading disabilities are referred to interchangeably in the literature and by practitioners as Reading Disorders, Disorders of Reading, dyslexia, Specific Learning Disabilities, and/or disorders with impairment in Reading. According to the definition devised by the International Dyslexia Association (IDA, 2008), dyslexia refers to a cluster of symptoms which result in individuals having difficulties with specific language skills, particularly reading. These literacy-related difficulties often result in a deficit of at least two years' worth of learning for affected children, and in addition to problems with reading, students with dyslexia usually also experience difficulties with other language skills such as spelling, writing, and pronouncing words. As with learning disabilities in general, reading disabilities are associated with the presence of unexpected problems with reading and language that are not accounted for by lower-than-average intelligence, environmental obstacles to learning, or other disabilities.

The cited prevalence rates of reading disability are not equally applicable to all children. For example, boys are approximately twice as likely to be diagnosed with a reading disability (Flannery, Liederman, Daly, & Schultz, 2000). This phenomenon possibly reflects a referral bias, such that the prevalence rates may be more comparable across sexes than reported, but that boys are more likely to be referred for assessment and therefore they more often receive a diagnosis. Using international samples, Rutter and colleagues (2004) reported a similar 2:1 male:female ratio in the rates of diagnosis of reading disabilities. Flannery and colleagues (2000) found that this sex ratio holds regardless of racial differences (i.e., comparing African American and Caucasian children), suggesting that the sex ratio in the prevalence of reading disabilities is not moderated by the child's racial background.

Additionally, season-of-birth has been identified as another factor related to the presence of reading disabilities, with North American children born during the summer months demonstrating a higher rate of diagnosis than children born during other months of the year. There are two competing explanations for this phenomenon. The first is that children born during the summer are simply younger than most other children when they enter school, as most schools have a cutoff date for entry into Kindergarten that falls in either late August or early September. Therefore, the summer-born children are biologically less mature than their older fall-born peers and may be less prepared for the demands of formal education (Donfrancesco et al., 2010). The second explanation is that harmful environmental influences have aligned with critical periods of prenatal development (specifically, the second trimester of pregnancy). It is thought that this alignment may have a deleterious effect on the developing central nervous system of the fetus, thus predisposing the child for later difficulties in the acquisition of reading skills (Donfrancesco et al., 2010). According to Donfrancesco and colleagues (2010), these environmental influences include factors such as influenza and other viral infections that are more common during the winter months, drastic changes in temperature, and Vitamin D deficiencies that are more common during the winter months, especially in northern climates, where there is limited access to direct sunlight during this season.

Another variable that has been linked to reading is prosodic processing. Prosody is a linguistic term which describes the rhythmic and tonal aspects of speech, or the "music" of oral language (Hudson, Lane, & Pullen, 2005, p.704). In terms of the research, Holliman, Wood, and Sheehy (2010) found that receptive sensitivity to both speech rhythm and non-speech rhythm (i.e., musical or metrical) predicted a significant

amount of unique variance in reading attainment that was independent of the contributions to reading attainment made by variables such as age, vocabulary, phonological awareness, and working memory. Furthermore, studies have been conducted which link oral receptive prosody with the attainment of higher-order reading abilities such as text comprehension, as opposed to simple word reading. For example, Cohen, Douaire, and Elsabbagh (2001) investigated the influence of prosody and its written equivalent, punctuation, in text comprehension, finding that in a sample of typical adults altered prosody and punctuation affect performance in a similarly deleterious fashion which seriously impairs listening and text comprehension and subsequent word recognition. The implications of limited prosodic understanding for children who struggle with reading are likely to involve even more severe impairment. Finally, Miller and Schwanenflugel (2006) found that children who had quick and accurate oral reading skills had shorter and more adult-like pause structures, larger declines in pitch at the ends of declarative sentences, and larger rises in pitch at the ends of yes/no questions, providing further support to the link between prosody and reading achievement.

Additionally, reading disabilities often appear co-morbidly with other disorders. For example, Light and DeFries (1995) observed a significant covariance between reading and mathematics disabilities, which they attributed to a combination of genetic and shared-environmental influences. Additionally, individuals with reading disabilities are more likely to meet the criteria for Attention-Deficit/Hyperactivity Disorder (ADHD) than are individuals without a reading disability diagnosis (Willcutt & Pennington, 2000a). This association was stronger for symptoms of inattention than for symptoms of hyperactivity/impulsivity. Finally, the presence of a reading disability diagnosis is also

associated with psychiatric comorbidity. Willcutt and Pennington (2000b) found that individuals with reading disabilities exhibited higher rates of all internalizing and externalizing disorders than individuals without this diagnosis, but that the presence of a reading disability was not significantly associated with symptoms of aggression, delinquency, Oppositional Defiant Disorder, or Conduct Disorder after controlling for comorbid ADHD. Symptoms of depression and anxiety, however, remained significantly associated with the reading disability diagnosis after controlling for ADHD. Willcutt and Pennington (2000b) suggest that this finding could mean that reading disabilities are specifically associated with internalizing difficulties. It should also be noted that significant gender differences were found, in that the relation between reading disabilities and internalizing problems was largely restricted to girls, whereas the relation between reading disabilities and externalizing problems was stronger for boys. There is also evidence to suggest that children with reading disabilities are not at increased risk for internalizing problems. Miller, Hynd, and Miller (2005) found that children with dyslexia do not tend to display the symptoms associated with anxiety, depression, and somatization more often than children with typical reading achievement. Clearly, the evidence on this topic is equivocal and further research is needed to say with certainty which psychiatric comorbidities are commonly observed in children with reading disabilities. It is clear, however, that children who struggle with reading are at enhanced risk for problems that lie outside of the domain of academic achievement.

The difficulties faced by individuals with reading disabilities are far-ranging and impact multiple areas of life. In fact, reading problems in the early years and onward are shown to have significant short- and long-term outcomes for those affected. According to

a study which took a comprehensive look at ten different Statistics Canada datasets that range across the lifespan of Canadians (Learning Disabilities Association of Canada, 2007), poor short-term outcomes for individuals with reading disabilities and other learning disabilities include early school dropout, low educational attainment, and lower overall levels of literacy achievement as compared to students without learning disabilities. In the longer-term, poor outcomes for adults with learning disabilities include a lack of success at finding and keeping employment and related financial problems, and a three-fold increase in reported problems with physical, general, and mental health, including high levels of distress, depression, anxiety disorders, and suicidal thoughts. Given the greater likelihood of experiencing adverse life outcomes such as these, the critical importance of developing a solid foundation in reading and other academic domains is apparent.

Assessment and Identification of Reading Disabilities

Aptitude-achievement discrepancy model. There is an ongoing debate regarding the optimal way to reliably and validly assess for and identify students with reading disabilities. According to Fletcher, Francis, Morris, and Lyon (2005) the aptitude-achievement method is the most widely utilized approach to identifying learning disabilities. This approach requires the measurement of aptitude, representing the individual's inherent potential for learning, and achievement, representing the actual accumulated learning of academic concepts by the individual. In implementing the aptitude-achievement discrepancy method of assessment and identification, the diagnosis of a reading disability is given if a discrepancy or gap of significant magnitude is observed between a student's scores on a test, or tests, of intelligence (aptitude) as

compared to his or her scores on measures of academic achievement. A discrepancy of greater than 1.5 standard deviations is considered to be severe enough to warrant a diagnosis in most cases (Sattler & Hoge, 2006), but other authors have also investigated the rate of learning disability prevalence using a 1.0 or 1.3 standard deviation cutoff (Proctor & Prevatt, 2003), finding, not surprisingly, that a smaller degree of discrepancy required for diagnosis leads to a higher hit rate.

This method is largely considered to be an invalid way to identify the presence of reading disabilities (Fletcher et al., 2004; Sattler & Hoge, 2006); however, it is the traditional model of diagnosing reading disabilities and a substantial literature is based upon it. One advantage of the discrepancy model is that special education services are provided only to those most likely to benefit from them, and thus a rationale is provided for dispensing services (Sattler & Hoge, 2006). A second advantage of the aptitude-achievement discrepancy model is that the achievement and the aptitude tests normally used in conducting a discrepancy assessment are known to have adequate reliability and validity, and the focus of the assessment is on the core area in which the student is experiencing difficulties (Sattler & Hoge, 2006). A third advantage is that the results of a discrepancy-based assessment are easily understood and conceptualized; an individual is diagnosed with a reading disability if the difference between his or her aptitude and his or her achievement in reading exceeds a predetermined cutoff point. In other words, the discrepancy model has what one might describe as "intuitive validity," in that the interpretation of the results obtained from this method of assessment make intuitive interpretational sense. A fourth advantage of the discrepancy model is that it is reflective of the actual definition of a learning disability, which states that it is a disorder that

affects the acquisition, organization, retention, understanding, or use of verbal or nonverbal information (which is suggestive of the achievement component), in individuals who otherwise possess at least the average abilities essential for thinking and/or reasoning (suggestive of the aptitude component; Learning Disabilities Association of Canada, 2002). While consideration of these advantages in choosing the appropriate method for assessing learning disabilities is important, evidence has amassed which suggests that the discrepancy model may not be the ideal method for the assessment of reading and other learning disabilities.

In its essence, the discrepancy approach involves the calculation of a difference score based on observed differences in scores on tests that are assumed to measure the constructs of aptitude and achievement. This difference score is used as a proxy for the true difference between these two latent constructs. So, in applying the discrepancy model, two questions arise: (1) Is this observed difference score a reliable and valid measure of the difference between the two latent unobservable constructs of interest? and (2) Is it an appropriate identifier of reading disabilities? (Fletcher et al., 2005). The first question is difficult to address, due to the fact that these constructs *are* unobservable and it is therefore difficult to determine whether tests of these constructs are really measuring what they are purported to be measuring (Sattler & Hoge, 2006). For example, Thompson, Detterman, and Plomin (1991) found that children's scores on intelligence tests and measures of school achievement are highly correlated. Given this finding, we must question whether two constructs which involve highly similar constructs (e.g., vocabulary, word identification, reading and language comprehension) can be adequately separated in order for valid comparisons to be made between them. It is also important to

note that there are a host of statistical problems associated with this comparison (e.g., scaling of items, comparability of norming groups, etc.; Sattler & Hoge, 2006).

As for the second question, which asks whether an aptitude-achievement discrepancy can be used to appropriately identify individuals with a learning disability in general, evidence that this approach is indeed not useful comes from Francis et al. (2005). These researchers tested the reliability of learning disability diagnoses that were made using the discrepancy method of classification. They found that by using an arbitrary cutoff point for how large the discrepancy must be in order to warrant this diagnosis, the reliability of the diagnosis did not hold over time. This is, in fact, true-to-form in the discrepancy approach in general; there is no universally accepted cut score (Fletcher, Lyon, Fuchs, & Barnes, 2007). This means that when groups are formed by imposing arbitrary cut-points, membership in the learning disabled group is unstable over time; those diagnosed with a learning disability at point A may not be diagnosed again at Point B, due merely to statistical change and not to practical or clinically significant improvement in achievement.

Furthermore, Stuebing and colleagues (2002) conducted a meta-analysis which found that negligible to small effect sizes were observed in terms of mean differences on measures of behaviour, achievement, and cognitive ability between students identified as having reading disabilities via the discrepancy approach and those students who were not identified. In other words, the discrepancy model was not able to adequately differentiate the two groups based on these variables, providing little evidence for the validity of this approach to the assessment of reading disabilities. If a discrepancy between intelligence and achievement does indeed form distinct and valid groupings, stability in the

classification of learning disabilities over time would be expected. Similarly, Sattler and Hoge (2006) stated that a disadvantage of the discrepancy model is that clinicians who are using the same discrepancy formula, but different tests to measure aptitude and achievement, may arrive at different classifications. It is also very important to consider the standard error of measurement of the tests that are chosen for comparison, as failure to do so could also lead to misclassification; however, doing so is not common practice.

Furthermore, Sattler and Hoge (2006) identified another disadvantage of the discrepancy model, which is that it may not take into account the child's absolute levels of performance. For example, given the finding that children of lower socio-economic status tend to have lower intelligence scores, these children are likely to be overlooked and therefore denied services, as their intelligence scores are not high enough to show the necessary discrepancy even though a learning or reading disability might indeed be present (Fuchs, Mock, Morgan, & Young, 2003). Similarly, a child whose Full Scale intelligence score is 150 and whose score on a test of reading achievement is 130 could technically be given a reading disability diagnosis based on the discrepancy model. However, giving this child the diagnosis would be inappropriate because his or her scores on measures of both aptitude and achievement are well above average, and thus this child may not be considered disabled. A final disadvantage of the discrepancy model is related to the timing of the distribution of services. The discrepancy model has been called a "wait-to-fail" model (Fuchs et al., 2003, p. 158) as it prevents children who need services and accommodations from receiving such considerations during the early years of schooling, given that these children must perform poorly for years before their achievement scores are sufficiently lower than their intelligence scores. Evidence such as

that presented by Francis and colleagues (2005), Sattler and Hoge (2006), and Fuchs and colleagues (2003) seriously calls into question the reliability, and therefore the validity, of learning and reading disability diagnoses that are made using the aptitude-achievement discrepancy model of classification.

Responsiveness-to-Intervention model. One alternative to the aptitude-achievement discrepancy model is the Responsiveness-to-Intervention (or Response-to-Intervention; RTI) approach to the classification of reading and learning disabilities. Using this approach, children who are struggling with learning to read are provided with more specialized instruction. If multiple waves of intervention are ineffective in improving performance, children are then classified with a reading disability. An RTI assessment involves ongoing monitoring of the progress made by a steadily dwindling group of children, wherein children are first given standard instruction in a regular classroom, and those who are not making adequate progress are given specialized and more intensive instruction. Progress is continually monitored and those children who still do not respond to a given number of waves of increasingly intensive instruction either qualify for special education services or at least for evaluations to determine the need for such services (Fuchs et al., 2003). The group of children becomes smaller and smaller because, as children begin to respond to the specialized interventions, they are taken out of the at-risk group and returned to less intensive instruction. According to the RTI model, following successive monitoring and intervention periods, only those students who truly do have reading disabilities will remain in the at-risk intervention group (Fletcher et al., 2007). Evidence for the effectiveness of the RTI approach in identifying children with reading disabilities comes from Vellutino, Scanlon, Small, and Fanuele

(2006), who found that either Kindergarten intervention alone, or Kindergarten intervention combined with intervention in Grade 1, are both useful in preventing early and long-term reading problems in most at-risk children, as well as in identifying those who are likely to experience continuing problems with reading.

Within the RTI approach to the classification of learning disabilities, there are two separate models, the more common of which is the problem-solving approach to RTI (Fuchs et al., 2003). The problem-solving approach to RTI takes into account the fact that no student characteristic (e.g., race, intelligence, socio-economic status) will determine whether or not a given intervention is effective. Instead, solutions to instructional and behavioural problems are deduced based on students' responsiveness to a four-stage process (Fuchs et al., 2003). At Stage 1, at-risk children are identified by teachers based on low reading achievement test scores and classroom performance. At Stage 2, the teacher consults with others about instructional modifications that will best meet the needs of the at-risk group. These modifications are then implemented, and their effects are monitored. At Stage 3, if the interventions are not successful in improving the reading achievement scores of some at-risk students, the school support team considers the causes of the problems seen in those who remain in the at-risk group, develops more intensive and targeted interventions, implements the newly-devised interventions, and continues to monitor any progress made by the at-risk group. In Stage 4, if the additional interventions still are not successful for some students in the at-risk group, these students will likely be assessed for eligibility for special education services (Sattler & Hoge, 2006), or more specifically, eligibility for a diagnosis of a reading disability. To summarize, within a problem-solving RTI framework, those students who show a positive response of

sufficient magnitude following intervention are removed from the intervention group and will subsequently receive standard classroom instruction. It is assumed that these students do not have reading disabilities. Those students who do not show the desired response at the conclusion of the intervention are the ones for whom a reading disability diagnosis is most likely to be appropriate.

The advantage to the problem-solving approach to RTI is that it is more sensitive to students' individual problems, in that the interventions can be tailored to meet the needs of a specific group of students. On the other hand, the intervention given to the students is not standardized, and therefore has not necessarily been previously proven efficacious. Furthermore, because the intervention to be given to the students is decided upon only by employees of a single school, the possibility of bias in instructional method and also bias that is based on preconceptions about the group of students who will receive the intervention cannot be ruled out. Finally, it is very difficult to conduct externally valid research on interventions that take a problem-solving approach as each intervention strategy is different, and valid comparisons cannot be made between two very different strategies implemented with diverse groups of students and/or teachers.

An alternative to the problem-solving model of RTI is the standard protocol model. In this approach, the same empirically-validated treatment is implemented for all students who present with deficits in a specific domain (e.g., a pre-validated intervention for phonemic awareness). An example of the implementation of standard protocol approach to RTI comes from Vellutino and colleagues (1996), who began by identifying poor readers at the beginning of Grade 1. These students were given a standardized intervention which targeted phonemic awareness, decoding, sight word reading, and

reading comprehension. Progress was monitored, and those who were still considered to be poor readers in Grade 2 were given additional intervention. The results of Vellutino and colleagues' (1996) study showed that two-thirds of the students originally classified as poor readers had demonstrated good or very good growth, while the remaining one-third remained in the low-achieving range despite having been given the same intervention as their peers. According to the general RTI model, the latter group consists of those students most likely to merit a learning disability diagnosis. To summarize, both the problem-solving model and the standard protocol model are RTI approaches, but the standard protocol model monitors all students' progress following the same intervention, whereas in the problem-solving model, the intervention is tailored specifically to each individual at-risk student's needs.

One advantage to the standard protocol approach to RTI is that it involves the standardized implementation of only one already-validated intervention strategy. Additionally, due to the influence of standardization of practices, the procedures for evaluating the effectiveness of the intervention strategy are usually more straightforward (Fuchs et al, 2003). Within a research context, a third advantage of the standard protocol procedure over the problem-solving method is that it is more conducive to study because valid external comparisons can be more readily made. One limitation of the standard protocol method is that one tutoring approach may not be suitable for all students with the same problem (Sattler & Hoge, 2006). Thus, it may not be that those students who remain low-achievers following intervention really merit a learning disability diagnosis; it may instead be that the specific intervention they were given was not as applicable to them as it may have been to others. Another limitation of the standard protocol approach

is that schools may not have access to the funds necessary to implement an expensive standardized method of intervention (Sattler & Hoge, 2006). It is important to note, however, that regularly-instituted special education protocols also tend to be very expensive, and so it is important for policy-makers to consider and strike a balance between cost-effectiveness and the implementation of valid reading interventions.

In addition to evaluating the validity and effectiveness of the specific types of RTI assessment and identification, it is also important to consider the overall benefits and downfalls of this approach in the assessment of learning disabilities. One advantage that the RTI model has over traditional psychometric approaches to the assessment of learning disabilities is that it is not a "wait-to-fail" model (Fletcher, Coulter, Reschley, & Vaughn, 2004), which means that students are not required to demonstrate years of underachievement before they become eligible for a reading disability assessment and subsequent diagnosis and placement for services (although it should be noted that a diagnosis is not required in order to qualify for special education services in the province of Ontario). Related to this point is the fact that within the RTI framework, the intervention is actually a part of the assessment process. Specifically, assessment in the RTI model is ongoing and progress is monitored often. Those students who do not respond within a given time frame are identified as needing more formalized special education services (Fletcher & Vaughn, 2009). This means that students who are simply low-achieving readers who need extra help *and* those who have a reading disability can benefit from an RTI approach. Relatedly, while the RTI method begins by targeting a greater number of students than does the discrepancy approach, the number of students who end up being diagnosed with learning disabilities may actually be reduced since

those who respond positively to early interventions do not need to be referred for further assessment (Sattler & Hoge, 2006). Another advantage of the RTI method is that it is not based on a measurement taken at a single time-point, but rather it is based on the ongoing performance of the student (Fletcher et al., 2007), thus reducing the effects of time-of-measurement error.

As with any relatively new method of assessment, there are disadvantages and unanswered questions associated with the RTI approach. For example, one criticism of this approach is that there is a paucity of validated measures designed to quantify responsiveness to intervention (Sattler & Hoge, 2006). In other words, the standard of practice for just how much improvement in the targeted skills is required for students to be considered to have "responded" to the treatment is currently underdeveloped. Sattler and Hoge (2006) posed a number of additional as-yet unanswered questions. For example, these authors questioned what appropriate interventions look like at each grade level, how to effectively monitor student progress, who manages RTI assessments and results, and the efficacy of RTI when there are concomitant problems, such as sensory deficits. An additional question which proponents of the RTI model have not yet addressed is: what do we do with the students who do not respond to even the most intensive of interventions? With previous models of learning disability identification, the procedure was to first identify the learning disability and then implement intervention. With the RTI model, intervention is built into the model and students' levels of responsiveness to it form the crux of identification procedures. Only once students have been put through the successively more intensive rounds of intervention are they labeled as having a learning disability. The question remains, however: what next? Educators and

administrators find themselves with the problem of not knowing what to do with those children who are identified as having learning disabilities following the RTI model, as they have already exhausted their intervention resources as a part of the identification process.

Cognitive deficits assessment model. There is an opposing camp to RTI whose proponents submit that the testing of cognitive processes is necessary in the determination of the appropriateness of a Learning Disability diagnosis. This group purports that the RTI approach is too simple and does not get at the basic psychological processes which underlie learning disabilities in reading and other areas of academic achievement. In other words, the cognitive deficits assessment models aims to do more than identify a student with a learning disability; rather, the emphasis is on identifying the deficits in functioning that are causing low achievement. According to Hale and colleagues (2004), using a RTI model without also using standardized instruments relies on inferences regarding the basic psychological processes, rather than "objective" measurement of these constructs. Hale and colleagues (2004) argued that the conceptual definition of learning disabilities implies a discrepancy between intact processes and those that are disordered, and that in order to measure these areas of integrity and deficit, it is necessary that well-validated, reliable, stable, and well-normed cognitive tests be a part of the assessment and identification strategy. Hale and colleagues (2004) pointed out that the removal of objective individual measurement of cognitive factors may increase the likelihood of classification errors, as poor academic achievement and failure to benefit from current instructional practices has often been linked to factors other than

learning disabilities, including experiencing low socio-economic status, being a person of racial or ethnic minority status, and having limited English proficiency.

A combined cognitive assessment/RTI framework. While both Fletcher (2006) and Hale and colleagues (2004) highlighted the need for a comprehensive evaluation of each individual student who may be eligible for a learning disability diagnosis, the nature of these evaluations appears to be quite different. From Fletcher's (2006) perspective, the comprehensive evaluation should address three issues: 1) that the student's response to general education instruction was below expectation, indicating the possibility of a disability; 2) the presence of low achievement scores across multiple academic domains; and 3) contextual factors and the presence of associated conditions that would better account for low academic achievement should be ruled out (e.g., intelligence tests and the assessment of adaptive functioning to rule out Intellectual Disability, assessments of language status, and assessment of behaviour problems that could interfere with the student's RTI experience). As mentioned previously, Hale and colleagues (2004) believe that well-validated, reliable, stable, and well-normed cognitive tests need to be a part of the assessment process in order to properly measure these areas of integrity and deficit in each individual student. Given this unresolved difference of opinion, the question remains as to which method is best suited for the task of defining, identifying, and remediating learning disabilities: a strictly RTI method or the method which requires comprehensive evaluation of basic psychological processes prior to classification.

There appears to be an emerging belief that a combination of the RTI and the cognitive assessment approaches will best serve students at-risk for learning disabilities (Hale et al., 2006). In Hale and colleagues' (2006) combined approach, a three-tiered RTI

model is implemented, wherein Tier 1 involves standard instruction for all students. If a student is identified as a non-responder in Tier 1, an individualized problem-solving approach would be undertaken at Tier 2, allowing the teacher and support staff to define and analyze the problem that the individual student is having, come up with individualized intervention strategies, and then develop a relevant monitoring process to measure the student's progress within Tier 2. If the student still fails to respond to Tier 2 intervention, then he or she moves on to Tier 3, which would consist of a standard evaluation of basic psychological processes. That the student meets the definition of a learning disability is assured only in Tier 3 and only if the comprehensive evaluation reveals that he or she has cognitive processing and achievement deficits that exist within the context of processing integrities (to be discussed). The combined RTI/cognitive processing assessment method of Hale and colleagues (2006) appears to combine the best features of both Learning Disability assessment models while also avoiding those problems which are inherent in a strictly ability-achievement discrepancy-based model.

Characteristics of Disordered Reading

In saying that it is necessary to assess cognitive processes when considering a reading disability diagnosis, it is important to know just what should be included in a comprehensive psychological assessment. Semrud-Clikeman (2005) provided a review of the neuropsychological processes that should be considered within the context of a reading disability assessment. Semrud-Clikeman (2005) and others proposed that impairments in a number of neuropsychological domains which fall outside of academic under-achievement form the core of the problems seen in students with reading disabilities.

The first domain which Semrud-Clikeman (2005) highlighted as being important in the determination of a reading disability is language ability, stating that language is a natural process of our brains, whereas reading is an acquired skill that needs to be overtly taught. As the process of learning to read is founded upon early language achievement, impairment in language skills should lead to significantly more difficulty in learning to read. The roles of phonological awareness, phonological processing, and vocabulary are three areas of language ability that have been previously identified as being important to consider when identifying students with reading disabilities.

Phonological awareness refers to an individual's possession of the knowledge that the speech stream consists of a sequence of sounds – specifically phonemes, the smallest units of sound (Yopp & Yopp, 2000). Phonological awareness requires attention, memory, and accurate phonemic perception and manipulation (McBride-Chang, 1995). Anthony and Lonigan (2004) proposed that a 2- to 7-year-old student's level of phonological awareness plays a key role in literacy development, as evidenced by the fact that children who are better at detecting rhymes or phonemes are quicker and more successful in learning to read. Pratt and Brady (1988) suggested that, for older as well as younger children, success at learning to read is directly related to the extent to which these children are aware of the phonological structure of spoken language and it is now well-documented that training children in phonological awareness tasks is associated with improved outcomes in reading and spelling (Bus & IJzendoorn, 1999; Ehri et al., 2001; Troia, 2004).

Phonological processing refers to the *use* of phonological information (i.e. language sounds) in processing written and oral language (Wagner & Torgesen, 1987).

Wagner and Torgesen (1987) proposed that phonological processing plays a causal role in learning to read, and in the reading difficulties experienced by individuals with reading disabilities. Considerable evidence has been amassed in support of this proposal (Bruck, 1990, 1992; Iversen & Tunmer, 1993; Wagner, Torgesen, Laughon, & Rashotte, 1993; Torgesen et al., 1999; Dally, 2006), although it is still unclear exactly why some children develop such deficits (Troia, 2004). Troia (2004) suggests that reading disabilities appear as a consequence of students' diminished capacity for phonological processing by affecting the way in which cognitive resources are allocated. Specifically, when greater-than-normal working memory resources must be allocated to decoding, this takes aware from resources which would be allocated to the understanding of words and passages in a reader with strong phonological processing skills, and thus reading comprehension deteriorates.

The role of vocabulary becomes important when considering the importance of a base of prior knowledge on reading achievement. Muter, Hulme, Snowling, and Stevenson (2004) found that a larger vocabulary was in fact a significant predictor of better reading comprehension in 5- and 6-year-old children, suggesting that some prior word knowledge is indeed an important factor in the process of reading comprehension. Similarly, Ouellette (2006) found that a typical Grade 4 student's depth of vocabulary (i.e. the size of oral vocabulary and the depth of semantic knowledge) directly predicted that individual's performance on tests of reading comprehension. Furthermore, early evidence to support the influence of vocabulary on reading comprehension comes from Beck, Perfetti, and McKeown (1982), who found that those participants who had

undergone vocabulary instruction performed significantly higher on tests of reading comprehension than did those who had not received training.

As previously mentioned within the context of phonological processing, the ability to decode words is another important consideration in students at-risk for reading disabilities. Semrud-Clikeman (2005) points out that the main difficulty experienced by students with reading disabilities, especially in the later years of reading instruction, relates not to simply being able to decode the words successfully, but to the rate at which the words are read, also known as reading fluency. This is consistent with Chall's (1983) model of reading, which posits that learning to read is a stage-like process wherein decoding ability is mastered first, but that the child must move beyond simple decoding to the ability to read fluently if he or she is going to be a successful reader. Furthermore, Semrud-Clikeman (2005) also stated that it is important to evaluate the child's speed of information processing outside of the confounding effects of decoding ability.

Another important neuropsychological process highlighted by Semrud-Clikeman (2005) as being impaired in students with reading disabilities is working memory. She states that "in order to decode words, the child's working memory must be functional and allow the child to retain a 'template' of the letters until the word is sounded out" (Semrud-Clikeman, 2005, p. 565). Working memory was also previously highlighted within the present discussion as being an important factor in the normal acquisition of reading ability. As mentioned previously, according to Troia (2004) working memory is important in the sense that reading problems arise as a consequence of less-than-ideal allocation of cognitive resources. Another way in which the working memory system is important is in allowing the child to access previously learned materials, in that "[i]f

difficulty is present at the outset, or at the working memory stage, the child will have difficulty recalling previously learned skills[…]and thus decoding will be slower and effortful" (Semrud-Clikeman, 2005, p. 565).

The final area that Semrud-Clikeman (2005) highlighted as being important for the learning process is executive functioning skills, in that proficient executive functions help a child to evaluate his or her performance and inhibit a response to irrelevant stimuli, thus increasing focus when reading. Additionally, the selection of what is important to encode is essential in learning to read, as is the ability to "hear" what one is reading and evaluate its correctness, as well as being able to self-correct mistakes. In sum, impaired reading appears to involve a breakdown in basic psychological processes, as well as a breakdown in the ability to coordinate these processes in order to facilitate higher-order reading skills such as fluency and comprehension.

Another domain outside of Semrud-Clikeman's review (2005) in which impairment is often found in students with reading disabilities is rapid serial naming. The ability to successfully engage in rapid naming tasks involves quick lexical access and the ability to rapidly translate information from the lexical store into verbal output. Raberger and Wimmer (2003) presented children with a digit and colour naming task wherein they had to name either a digit or the colour associated with a dot as quickly as they could. The dependent variable here was the time that it took for the children to name all 50 of the test items. Raberger and Wimmer (2003) found that those children in the reading disabled group performed much more poorly on this test of rapid naming than did children with ADHD and non-impaired controls. Similarly, Scarborough (1998) found that rapid naming was a particularly good predictor of reading outcome over time in

children who are designated as having reading disabilities, while Kirby, Georgiou, Martinussen, and Parrila (2010) promoted the idea of including measures of naming speed in efforts at early identification for those students at risk for reading disabilities, as naming speed can be measured prior to the onset of formal reading instruction. Based on evidence such as this, it appears as though impairments in rapid naming ability are also at the core of the deficits experienced by students with reading disabilities.

Related to the literature on rapid naming deficits in students with reading disabilities, Wolf and Bowers (1999) proposed a double-deficit hypothesis for children whose reading problems are especially severe. The double-deficit hypothesis posits that "phonological deficits and the processes underlying naming speed are separable sources of reading dysfunction, and their combined presence leads to profound reading impairment" (p. 416). In other words, students with reading disabilities may present with phonological deficits alone, rapid naming deficits alone, or both problems in combination, and those who present with the double deficit will experience the greatest amount of trouble in acquiring and applying reading skills. According to Wolf and Bowers (1999) knowledge regarding which of these three dyslexic subtypes a child falls under will have important implications for intervention. Specifically, they proposed that readers who present with solely phonological deficits should benefit from phonics-based interventions, but that readers with naming-speed deficits and both types of deficits (double-deficit readers) will not benefit as readily from phonics-based interventions and additional considerations will need to be made in planning remediation for thusly-affected children.

Up until this point, the present discussion has focused on the areas of deficit that are most often seen in students with reading disabilities. However, according to the perspective of Hale and colleagues (2004), these deficits must be seen within the context of processing integrities if the reading problems are to be considered a true reading disability that is not better accounted for by low achievement in reading due to inadequate education or a global cognitive impairment. Fletcher and colleagues (1994) presented the idea of the phonological limitation hypothesis, which suggests that difficulties in printed word decoding, phonological segmentation of spoken words, rapid naming, and verbal short-term memory form a coherent syndrome, and that students who have reading disabilities will have deficiencies that are specific to phonological awareness and related language measures, whereas students who are more generally poor readers will have broad-based cognitive deficiencies. Fletcher and colleagues (1994) measured nine areas of cognitive and linguistic ability: phoneme deletion, visual-spatial deletion, verbal short-term memory, nonverbal short-term memory, speech production, vocabulary/word finding, rapid naming, visual-motor abilities, and visual attention. They hypothesized that measures of phonological awareness would be strongly related to reading disability diagnosis, but measures of visual-spatial and visual-motor ability, nonverbal memory, and visual attention skills would be weakly related to reading disability diagnosis. Their hypothesis was supported, in that measures of phoneme deletion are clearly the most robust correlates of impaired reading in students with reading disabilities. Therefore, when in the presence of deficits in the phonological domains, integrities in the non-phonological domains that were assessed by Fletcher and colleagues (1994; specifically, visual-spatial and visual-motor ability, nonverbal memory,

and visual attention skills) should be suggestive of a reading disability, as opposed to a simple "garden-variety" reading impairment (Stanovich, 1988).

Reading Instruction

As a parallel to the bottom-up/top-down debate, there is a debate between advocates of two different approaches to reading instruction. Because reading instruction is intimately tied to the way in which the process of learning to read is conceptualized, it is assumed that whichever approach is taken by teachers and educators generally reflects students' learning needs. The first of these two approaches to instruction is known as the whole language approach to reading instruction. This approach assumes that top-down processing is more often employed in acquiring reading skills and that students must first grasp basic language concepts which will then guide them as they learn to read. The code-based, or phonics instruction, approach focuses on the teaching of phonics first, which deal with the basics of the letter-sound relationships and the ability to decode words outside of the influence of context. Within this framework, it is only once students have mastered the phonics-based concepts that they are introduced to more complex and higher-level reading concepts.

According to Foorman (1995), the classic form of the debate between whole language- and phonics-based asks the question: Is reading instruction more effective when it emphasizes whole-language learning or when it emphasizes the alphabetic code that relates letters and sounds? According to Hempenstall (1997), early work from Chall (1967) suggested that "systematic teaching of phonics tended to produce better word recognition, spelling, vocabulary and comprehension in all children, not only those from the at-risk groups (such as students of lesser intelligence, or those from lower socio-

economic backgrounds)" (p. 408). More recently, the National Reading Panel (2000) has also claimed that empirical evidence clearly favours explicit instruction in alphabetic coding, or phonics. On the other hand, a meta-analysis by Stahl, McKenna, and Pagnucco (1994) found that whole-language approaches also have a small positive effect on reading achievement, but there were too few studies available at the time to test whether this effect is statistically significant. As seems to be often the case in many areas of inquiry, there are advocates for a combined approach, wherein the theoretical frameworks that underlie both whole language and phonics-based teaching methods are brought together. For example, Dahl and Scharer (2000) suggested that future discussions about whole language and phonics must move away from an artificial, simplistic dichotomy that is not reflective of the reality of practice in most classrooms. These authors, as well as the National Reading Panel (2000), advocate for the application of phonics skills within the context of meaningful reading and writing activities, in order to maximize students' applications of phonics concepts as they read and write. Dahl and Scharer (2000) also stress the importance of attending to the individual learning needs of the students.

Reading Intervention

In planning reading interventions for students at-risk for or already experiencing delays, Torgesen (2005) submitted that one of the most important lessons to keep in mind for students with Reading Disabilities is that these students' individual needs are heterogeneous, and as such, even strongly evidence-based interventions may not be entirely applicable to or helpful in remediating the reading difficulties of any one individual student. Once this factor has been acknowledged, the evidence suggests that it is possible to teach all students to accurately apply alphabetic principles in decoding

novel words, even following a significant delay (first 3 to 4 years of schooling) in acquiring this ability (Torgesen, 2005). It is also known that the decoding and word identification accuracy and reading comprehension of students with relatively severe reading disabilities can be dramatically accelerated via carefully administered and intensive interventions (Torgesen, 2005). Hence, despite serious delays in reading achievement and reading skills acquisition, students with reading disabilities can make significant gains as long as reading interventions are sensitive in taking into account their individual learning needs.

The focus of the present work is to investigate students' reading achievement within the context of one particular reading intervention: the Peer Assisted Learning Strategies (PALS) program developed by researchers at Vanderbilt University (Fuchs, Fuchs, Mathes, & Simmons, 1997). According to Mathes, Howard, Allen, and Fuchs (1998), the PALS intervention strives to combine carefully designed systematic phonics instruction with more holistic practices that emphasize contextualized reading experiences, focusing on helping students to make improvements in multiple domains, including word recognition, fluency, and comprehension, while also emphasizing phonological skills and alphabetic knowledge.

One of the core concepts of PALS is the decentering of instruction, which involves students taking greater responsibility for their own learning while teachers serve as facilitators by arranging the environment and curriculum in an effort to enhance learning. PALS works through peer-mediated instruction, which involves the pairing of students with other children within their own classroom. One student is assigned to the role of the coach, and the other to the role of the reader. The role of the coach is designed

to be filled by a stronger reader, while the role of the reader is designed to be filled by a struggling reader. The classroom or special education teacher who is implementing the PALS program assigns the pairs based on knowledge of the students' baseline reading achievement and based on his/her knowledge of which students would work well or not well together. Each pair completes a tutoring program that has been carefully taught by the teacher. Additionally, each pair is assigned to one of two teams, for which they can earn points for academic activities and cooperative behaviour during tutoring. The PALS program is taught through a series of lessons, first by the teacher and then the student pairs take over the responsibility of learning the material together by completing a series of activities associated with each lesson. The lessons and activities are typically administered in 35-40 minute blocks, three times per week. The coach acts as a model to the reader by completing the activities associated with each lesson first and then guiding the reader through completion of these same activities. The content of the lessons varies by grade level. In the Grade 1 PALS program, the focus is on more basic skills like decoding and fluency, while students in Grades 2 and 3 practice higher-order reading skills such as reading comprehension.

The PALS program also has a built-in progress monitoring system that is based on curriculum-based measurement (CBM). CBMs have been shown to provide reliable and valid information about students' progress (Deno, Fuchs, Marston, & Shin, 2001). Within the PALS program, there are four separate CBMs. The first is Letter Sound Fluency (LSF), which measures how many letter sounds each student can decode within one minute; the second is Word Identification Fluency (WIF), which measures the number of words that each student can read within one minute; the third is Passage

Reading Fluency, which measures how many words from a connected passage of text each student can read within one minute; and the fourth is MAZE Fluency, which measures each student's ability to choose the correct word from a list of options to fill in the blanks in a connected passage of text. These CBMs are administered in the form of quick "probes" in an individual testing format by the teacher who is implementing the PALS program. As the students move through grade levels, the complexity of the probes that are administered to them increases. For example, a student in Grade 1 would not be expected to be able to complete the MAZE Fluency probe, whereas a student in Grade 3 would be expected to have moved well beyond simple letter sound decoding skills, so administration of this measure at that age would likely not provide clinically- and educationally-relevant information about that student's reading progress. CBMs have been deemed useful for measuring baseline and outcome reading levels before and after engaging in the PALS program, as well as for monitoring the development of students' reading skills in the program over time to chart whether or not individual students are progressing through literacy acquisition as expected.

In addition to describing the PALS program, Mathes and colleagues (1998) investigated the effectiveness of the program. This investigation looked at whether the implementation of the First Grade PALS program in school led to improvements in students' reading scores, and also at whether students and teachers were satisfied with and enjoyed working within the PALS framework. These authors found that PALS implementation did correspond with improvements scores on independent measures of reading achievement, and that this was especially true for the low-achieving readers who were at greater risk for delays in reading achievement. Additionally, both students and

teachers implemented PALS with relative ease, demonstrated high fidelity to the method, and reported high levels of satisfaction. Based on effectiveness research such as this, the PALS program was implemented locally within the Windsor-Essex Catholic District School Board (WECDSB), beginning in the 2008/2009 academic year.

Present Objectives

The present investigation will look at WECDSB students' reading achievement within the context of the PALS method and will consist of three separate but related studies that all relate to reading achievement and intervention during the early years of school. Study One will identify whether students who receive a psychological assessment report make greater gains in reading over the school year, presumably as a result of teachers having access to more information about these students' skills and challenges. The purpose of this study will be to determine whether there is value added by the findings from a psychological assessment (as is recommended by proponents of the cognitive assessment approach; e.g., Hale et al., 2004) on top of the ongoing data that are collected by teachers on a weekly basis within the classroom setting as part of the PALS intervention (as in the RTI method; Fletcher, 2006) in identifying students with reading disabilities.

Study Two will look at summer learning loss in students with reading disabilities. To be specific, the focus of this investigation will be on whether the reading skills of students who are low achievers in reading tend to regress to a greater extent during the traditional North American summer vacation as compared to their typically-achieving peers. Whether the low achieving students take longer to recover from summer loss will also be investigated, as will differences in the school-year learning trajectories of Low-,

Average-, and High-achieving readers. These questions are important, as additive deficits which accumulate over multiple summers could contribute to a significant Matthew Effect in reading for students who are already struggling to acquire adequate reading ability (Alexander, Entwisle, & Olson, 2007). The Matthew Effect in reading posits that the "rich get richer," or that good readers will continue to build upon and improve their reading skills, whereas the "poor get poorer," in that struggling readers will not get adequate exposure to print materials and therefore will continue to fall behind in reading. Although recent research on the Matthew Effect has partially refuted this phenomenon (Protopapas, Sideris, Mouzaki, & Simos, 2011), it cannot be fully ruled out within all populations and at all levels of reading achievement, and hence, it is subject to further investigation herein.

Finally, Study Three will look at the role of prosody in reading achievement. The purpose of this study will be to investigate the predictive power of language prosody in early reading scores. It is thought that well-developed understanding and use of language prosody might be related to the development of reading skills. Study Three will look at whether expressive and receptive prosodic processing act as predictors of reading outcome following the delivery of the PALS intervention. As mentioned, all three of these studies are related in that they all look at reading achievement within the context of the PALS intervention program.

Please note that information regarding the correlations between and the reliability of the PALS CBM data is presented in Tables 1 and 2. This information is referred to here because the reliability of and correlations between the CBMs are relevant pieces of information for all three studies, thus making them cross-cutting issues. With respect to

the observed correlations between measures, overall, the CBMs are highly correlated with each other. For the reliability data, both test-retest reliability and parallel forms reliability were examined, based on Stevens' (2009) discussion that it is important to examine more than one type of reliability coefficient when making judgments about the reliability of one's data. Most of the reliability coefficients and correlations were calculated by comparing students' scores on the two baseline measurements taken at the beginning of each school year. For example, two baseline measurements are administered in Grade 1 for the LSF measure. These two baseline measurements were the scores that were considered when calculating both test-retest reliability and parallel forms reliability. All of the data for which two baseline measurements were available demonstrated excellent reliability. For the MAZE Fluency measure, however, no baseline measurements were administered. Rather, the MAZE Fluency measure is only administered in January and June of the Grades 2 and 3 school years. Therefore, these two scores were subjected to the reliability analyses for the MAZE Fluency measure. Because they are administered months apart (and therefore are not both baseline measurements), it makes sense that both the correlations between these scores as well as the reliability coefficients for the MAZE Fluency scores would be lower than for those measures which were administered closer together in time. Even with this separation of measurements, however, both forms of reliability were deemed adequate within the present sample for the MAZE Fluency measurements.

CHAPTER II

The Roles of Psychological Assessments and Curriculum-Based Measurements in Identifying Reading Disabilities

Learning disabilities represent a significant problem in Canadian schools. According to the Learning Disabilities Association of Canada (2007a) approximately 1 in 10 children has some form of Learning Disability (LD), whereas other evidence suggests that approximately 3% of the Canadian school-aged population (ages 5-14) is affected by learning disabilities (Statistics Canada, 2006). Without research-based and effective reading intervention, Canadian children with LDs are at risk for a host of poor outcomes. According to a study which took a comprehensive look at ten different Statistics Canada datasets that range across the lifespan of Canadians (Learning Disabilities Association of Canada, 2007b), these poor outcomes can manifest in both the short- and the long-term. Short-term outcomes include early school dropout, low educational attainment, and lower levels of literacy achievement as compared to students without LD. In the long-term, poor outcomes for adults with LD include a lack of success at finding and keeping employment and related financial problems, and a three-fold increase in reported problems with physical, general, and mental health, including high levels of distress, depression, anxiety disorders, and suicidal thoughts. Given the problems encountered by individuals with LDs, both in and out of school, it is important that school staff and other professionals know how to define, identify, and remediate LD.

Within both research and practice, however, there is ongoing debate as to how LD should be defined, identified, and remediated. There appear to be two camps: the Response-to-Intervention (RTI) proponents and those who advocate for more thorough

cognitive assessments for all children suspected to have LD. Both groups have made valid contributions to the literature on LD. There also exists a faction of researchers and clinicians that propose a combination of these methods in defining, identifying, and planning interventions for students with LD. Given that trends from the United States tend to spur changes in the Canadian system, an American model of identification will be used within the context of the present study.

RTI and Cognitive Assessment Approaches in LD Identification and Treatment

Recently, a reauthorization of the Individuals with Disabilities Education Act (IDEA; U.S. Department of Education, 2004) occurred. In its reauthorized state, IDEA 2004 puts an emphasis on early intervention services and provides specific provisions that allow school districts in the U.S. to adopt models of LD that focus service allocation toward RTI (Fletcher & Vaughn, 2009). It should be noted that IDEA 2004 does not require that an RTI model be adopted, but that it does allow for this type of approach to be implemented in school districts. This change has led to a hard push for RTI practices to be adopted, and in part for good reason, as the RTI approach does pose a significant advantage in many ways over the more traditional psychometric LD assessment methods.

Specifically, one advantage that the RTI model has over traditional psychometric approaches to the assessment of LD is that it is not a "wait-to-fail" model (Fletcher, Coulter, Reschley, & Vaughn, 2004), which means that students are not required to demonstrate years of underachievement before they become eligible for a LD assessment, diagnosis, and placement for services. Related to this point is the fact that within the RTI framework, the intervention is actually a part of the assessment process. Specifically, assessment in the RTI model is ongoing and progress is monitored often. Those students

who do not respond within a given time frame are identified as needing more formalized special education services (Fletcher & Vaughn, 2009). This means that students who are simply low-achieving readers who need extra help and those who have a LD in reading can benefit from an RTI approach.

In fact, Bradley, Danielson, and Hallahan (2002) suggested that low achievement and an inadequate response to effective research-based interventions can be used as the sole determining factors for LD identification, as long as the typical exclusionary factors (i.e., intellectual disability, sensory deficits, serious emotional disturbance, limited English proficiency, and lack of opportunity to learn) are ruled out. According to Fletcher (2008), this reflects the current and historical underpinning of LD; only if the student demonstrates a lack of response to quality instruction can it be known that their low achievement is unexpected, otherwise low achievement may be explained by other factors such as a lack of opportunity to learn. Fletcher (2008) went on to claim that an intelligence-achievement discrepancy (which has been heavily criticized and largely dismissed within the LD field; Fletcher et al., 1994; Fuchs, Mock, Morgan, & Young, 2003; Hale, Kaufman, Naglieri, and Kavales, 2006) and discrepancies across cognitive domains do not provide the same assurance of unexpected underachievement.

On the other hand, the camp which suggests that the testing of cognitive processes is necessary in the determination of LD would purport that the RTI approach is too simple and does not get at the basic psychological processes which underlie LD. According to Hale and colleagues (2004), using a RTI model without also using standardized instruments relies on inferences regarding the basic psychological processes, rather than objective measurement of these constructs. Hale and colleagues (2004)

pointed out that the removal of objective individual measurement of cognitive factors may increase the likelihood of classification errors, as poor academic achievement and failure to benefit from current instructional practices has often been linked to factors other than LD, including experiencing low socio-economic status, being a person of racial or ethnic minority status, and having limited English proficiency.

While both Fletcher (2006) and Hale and colleagues (2004) highlight the need for a comprehensive evaluation of each individual student who may be eligible for a LD diagnosis, the nature of these evaluations appears to be quite different. From Fletcher's (2006) perspective, the comprehensive evaluation should address three issues: (1) that the student's response to general education instruction was below expectation, indicating the possibility of a disability; (2) the presence of low achievement scores across multiple academic domains; and (3) contextual factors and the presence of associated conditions that would better account for low academic achievement should be ruled out (e.g., intelligence tests and the assessment of adaptive functioning to rule out Intellectual Disability, assessments of language status, and assessment of behaviour problems that could interfere with the student's RTI experience). From the perspective of Hale and colleagues (2006), the conceptual definition of LD implies that a discrepancy between intact processes and those that are disordered must be established as the factors underlying the observed unexpected academic underachievement. Additionally, Hale and colleagues (2004) believe that well-validated, reliable, stable, and well-normed cognitive tests need to be a part of the assessment process in order to properly measure these areas of integrity and deficit in each individual student. Given this unresolved difference of opinion, the question remains as to which method is best suited for the task of defining,

identifying, and remediating LD: a strictly RTI method or the method which requires comprehensive evaluation of basic psychological processes prior to LD classification (Hale et al., 2006), or perhaps one that combines these approaches.

There appears to be an emerging belief that a combination of the RTI and the cognitive assessment approaches will best serve students at risk for LD (Hale et al., 2006). In Hale and colleagues' (2006) combined approach, a three-tiered RTI model is implemented, wherein Tier 1 involves standard instruction for all students. If a student is identified as a non-responder in Tier 1, an individualized problem-solving approach would be undertaken at Tier 2, allowing the teacher and support staff to define and analyze the problem that the individual student is having, come up with individualized intervention strategies, and then develop a relevant monitoring process to measure the student's progress within Tier 2. If the student still fails to respond to Tier 2 intervention, then he or she moves on to Tier 3, which would consist of a standard evaluation of basic psychological processes. Definition of LD is assured only in Tier 3, and only if the evaluation reveals that he or she has cognitive processing and achievement deficits in the context of processing integrities. For comprehensive coverage of what should be included in a thorough assessment of psychological processes, see Semrud-Clikeman (2005), Raberger and Wimmer (2003), Scarborough (1998), and Wolf and Bowers (1999). The combined RTI/cognitive processing assessment method of Hale and colleagues (2006) appears to combine the best features of both LD assessment models.

LD Diagnosis and Treatment in Canada

Within the Canadian context, the issue of how LDs are defined and identified appears to have received less attention and debate. According to the Canadian definition

of LD, "Learning Disabilities refer to a number of disorders which may affect the acquisition, organization, retention, understanding or use of verbal or nonverbal information. These disorders affect learning in individuals who otherwise demonstrate at least average abilities essential for thinking and/or reasoning." (Learning Disabilities Association of Canada, 2002). This definition also notes the following: "Learning disabilities are suggested by unexpected academic under-achievement or achievement which is maintained only by unusually high levels of effort and support." Use of the word "unexpected" leaves open the opportunity for using an ability-achievement discrepancy model to identify Canadian students with LD. Given the research suggesting that the ability-achievement discrepancy definition does not adequately discriminate between students with LD and students who are simply low-achieving readers (Fletcher et al., 1994), and the fact that Canadian school policies are usually spurred by educational practices within the United States, it appears that a shift towards a more modern method of assessment is due in Canada. This raises a number of questions: Should Canadian schools adopt a strictly RTI approach, or a strictly cognitive definition of LD, or turn toward a combined approach like that of Hale and colleagues (2006)? If schools do require that students participate in cognitive testing in order to confirm a profile of mixed strengths and weaknesses, what should these profiles look like and what should be included in the assessment in order to merit a diagnosis of LD? And once the diagnosis is given, how should clinicians and teachers go about remediating students with reading problems, given the great degree of heterogeneity in students who require this remediation?

The goal of the current study is to examine whether students whose teachers have access to a psychological assessment report demonstrate any extra gains in reading scores over their peers who did not undergo psychological evaluation and whose teachers do not have access to a report. The research question associated with the present study will be: Is there added value to conducting a full psychological assessment, or are CBM data enough to identify LDs and plan interventions? It may be that a failure to respond to early intervention efforts paired with frequent, brief functional assessments provides adequate information for educational planning for at-risk students.

Method

Original Study Design

It should be noted that there was a qualitative analysis planned that would have investigated teachers' perspectives on the utility of psychological assessment reports, but this strategy was discontinued due to a lack of interest in research participation by Windsor-Essex Catholic District School Board (WECDSB) teachers, who failed to respond to advertisements posted within school staff rooms. Additionally, the initial focus of the present study was on whether students who have a more comprehensive psychological assessment report make greater gains over students whose reports are less comprehensive. The reports were subjected to a literature-based (e.g., Semrud-Clikeman, 2005) "best-practice" checklist that measured the comprehensiveness of each report separately. The checklist indicated whether the report included information in the following domains: vocabulary (both from an intelligence-based measure, and whether the students were subjected to additional receptive and expressive vocabulary testing), phonological awareness/processing, word reading, nonword reading, letter reading

fluency, word reading fluency, nonword reading fluency, passage reading fluency, working memory, spelling, executive functioning, attention, rapid naming, reading comprehension, math, written expression, any other academic domains, and whether an intelligence test was administered. It was also noted if the report specifically applied an LD diagnosis. The checklist strategy, however, was revised because all of the reports were about equally comprehensive and there was inadequate variance in the checklist data. Therefore the question became: Do students who have a report learn more over the subsequent year than students who do not? This prompted the addition of the matched controls to the analysis. It should be noted that efforts were made to control for generally low reading scores.

Participants

In order to have received a psychological assessment through the Windsor-Essex Catholic District School Board (WECDSB), the students' parents were required to sign a consent form for services which also indicated their consent to their children's data being used for research purposes in the future. This study did, however, require Research Ethics Board clearance, given that the data from CBM probes and psychological assessment reports needed to be reviewed with names attached in order for the two sources of data to be matched at the individual student level.

Approximately 120 assessments are completed every year within the WECDSB, and the present study made use of reports from 2008 onward. In theory, this meant that approximately 600 psychological reports were available for analysis. However, the sample size was significantly reduced from this large number because there were a number of exclusionary criteria. Specifically, in order to be matched with PALS data, the

students had to be in Grade 3 or younger, as that is the time when PALS administration ends within the school board. This exclusion criterion alone ruled out approximately 450 reports. Also, reports which identified students as having an Intellectual or Developmental Disability were not considered, as these conditions must be ruled out before an LD diagnosis is considered. This ruled out another approximately 50 reports. In the end, the number of eligible reports that were completed during the proper timeline and assessed for reading problems was 103. However, these students' reports had to be matched with the PALS data of control participants. This means that only the reports from the early 2012-2013, late 2011-2012, and the summer of 2012 could be considered, as the matched controls needed to come from the same class as the report participants, and it was not possible to identify both the class and the PALS data of students in years other than the 2012-2013 school year. Demographic information such as sex and whether or not the participants were born in Canada was also noted where available.

The sample size was further reduced from the 103 total reports that were reviewed and deemed appropriate for analysis because CBM data could not be located from within the WECDSB records for all 103 students. The total sample size of students who both had a report and for whom matching CBM data could be located was 18. Each of these participants was matched by the primary investigator with another student from the same class in the same grade. The mean age of the report group was 87 months (SD = 9.2 months) at the time that the reports were written. Given that they were matched with students who were in the same class and grade during the assessment year and who therefore should have been same-age peers, it is assumed that the students in the comparison group were approximately the same age, but this assumption cannot be

confirmed as the primary investigator did not have access to the birthdates of the control participants. At the time of the assessment, nine of the students were in Grade 3 (i.e., nine assessments reports were completed at the beginning of the Grade 3 year), and the remaining nine were tested near the end of Grade 2 or over the summer between Grades 2 and 3. There were 12 male participants in the report sample (66.7%) and 6 female participants (33.3%). Efforts were made to match report participants to control participants of the same gender, but only after they were matched with students in the same grade and class whose CBM data most closely resembled their own (so only if there were two students of opposite genders whose data matched equally closely with that of the report group student was gender used as a secondary matching criterion). It is also important to note that the report group participants had undergone psychological assessment as a result of an identified deficit in reading achievement. Although attempts were made to identify students whose reading achievement at the beginning of the Grade 3 year was similarly impaired when compared to each report group participant, it is possible that the participants were presenting as qualitatively different from each other; a reading disability diagnosis is assured in the report group, while the possibility that the comparison group participants demonstrated low scores in reading could be more likely attributable to factors other than a true disability in reading.

Materials

Curriculum-based measurements (CBMs). The measures used within the current study were the CBMs associated with the Peer Assisted Learning Strategies (PALS) program, created by Fuchs, Fuchs, Mathes, and Simmons (1997). This reading intervention focuses on decentering reading instruction in classrooms by pairing peers

with each other for intensive but efficient reading practice. High levels of effectiveness, as well as a high degree of both teacher and student satisfaction with the program, have been demonstrated (Mathes, Howard, Allen, & Fuchs, 1998), and thus, the PALS intervention is widely used in school settings. Within the PALS framework, the measurement of reading skills is accomplished using the curriculum-based measurements (CBMs) associated with the PALS program. These CBMs are meant to track student progress regularly throughout the duration of the intervention. The PALS program measures four reading-related domains: Letter-Sound Fluency (LSF), Word Identification Fluency (WIF), Passage Reading Fluency (PRF), and MAZE Fluency. However, only the PRF and MAZE Fluency measures were used in the present analysis, as the sample was made up of students in Grade 3 only, and the LSF and WIF measures are not administered in Grade 3. Both measures were administered to both the control and the report group participants.

Passage Reading Fluency. The PRF test assesses students' fluency in reading grade-level text passages. In completing this task, students read aloud a grade-appropriate passage for one minute. Performance on this task is measured by counting the number of words attempted within the time limit and subtracting the number of words that were omitted and the number of words that were read incorrectly to arrive at a total score which represents the total number of words read successfully within the time limit.

MAZE Fluency. The MAZE Fluency test assesses reading comprehension by having students read a passage and circle the correct word from a group of choices for each blank. Students are given two and a half minutes to complete this task, and they receive one point for each correct response. Scoring is discontinued if three consecutive

errors are made. It should be noted that this measure was only administered twice (in January and June), whereas the PRF measure was administered at regular intervals across the entire school year.

Psychological assessment reports. Within the WECDSB, psychological assessments that are aimed at investigating the possibility of a LD in reading tend to cover standard constructs, the first of which is general cognitive abilities. General cognitive abilities are measured using the indexes from the fourth edition of the Wechsler Intelligence Scale for Children (namely, Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed). The Verbal Comprehension index measures language-related abilities such as vocabulary, practical and general knowledge, and verbal abstract reasoning. The Perceptual Reasoning index measures the ability to make sense of visual patterns and to attend to details presented visually. The Working Memory index measures the ability to hold in mind and mentally manipulate information such as a string of letters and numbers or math problems that are presented verbally. The Processing Speed index measures the ability to perform pencil-and-paper tasks quickly. Also routinely assessed are visual motor and visuospatial abilities, attention (specifically, the ability to maintain and focus attention), learning (as measured by the ability to benefit from repetition and cues, and to recognize previously presented information), the use of memory strategies such as chunking information into semantic categories, sentence memory, coping and adjustment, executive functioning, and adaptive behaviour.

Academic achievement is also investigated within WECDSB assessments which considered a diagnosis of a LD in reading. The assessment of academic achievement involves the measurement of basic reading skills such as pseudoword decoding and

simple word reading, basic spelling skills, as well as higher-order reading and writing skills such as reading comprehension, accuracy, rate, and fluency (a combination of rate and accuracy) and writing abilities, including letter formation, proper spacing between letters, measures of expressive writing, and written fluency. A diagnosis of LD in reading is given if the student presents with severe academic problems in the absence of an Intellectual and/or Developmental Disability, as mentioned above.

If a diagnosis of LD in reading is given, the following recommendations are routinely made within the context of a WECDSB report, depending upon the problems manifested by each individual student: (1) explicit and systematic phonics instruction to improve decoding skills, (2) improving reading fluency by emphasizing repeated oral reading with adequate feedback, (3) activities that promote automaticity in reading through repetition, drill, and practice (practice must be consistent and extend over a period of weeks or months – e.g., 10 to 20 minutes every day for next several months), (4) the building in of incentives for student progress, (5) the provision of copies of notes that include visual representations of words for students who struggle with spelling (depending on the student's grade level), (6) additional time to complete assignments or tests, (7) the opportunity to review the material that was covered in order for the student to be able to seek clarification of concepts or ideas and to integrate material the he or she did not immediately understand, (8) strategies to improve encoding and recall for students who experience memory problems, (9) modified test formats, (10) scribing support, (11) vocabulary development strategies if the student is struggling with reading comprehension, (12) explicit instruction in comprehension strategies, and (13) reading support in the classroom wherein assignments are read to those students who struggle

with reading to make sure that they are getting the necessary information. The recommendations that relate to continued specific instruction in phonics and/or comprehension are often achieved by having the student continue participation in the PALS program.

Procedure

The matched controls were chosen by looking at all of the CBM scores within the report student's class, and choosing the student whose baseline data most closely matched that of the report participant. Gender was used as a secondary matching criterion when possible. Progress over the school year was measured by subtracting the average of each student's two baseline measurements from the last recorded score of the school for that student.

Theoretical Model of Change

Collins (2006) outlined a number of elements that must be considered when articulating a theoretical model of change. With respect to the general or characteristic shape of change, it is expected that reading skills on both measures will increase linearly over the school year. Although reading skills can develop in a more discrete, stage-like manner that is characterized by jumps in development overall, it is thought that each individual skill (e.g., decoding ability) develops continuously over time, increasing in a linear fashion. It is not expected that there will be periodicity or a cyclical nature associated with this change. The change in reading achievement over the school year is thought to be due to calendar time as well as the reading instruction that is provided over that calendar time to foster learning. The most relevant covariates that may predict changes aside from calendar time and the PALS intervention itself are socio-economic

status and perhaps reading encouragement within the home environment, outside of the PALS program. Attention factors and motivation to achieve may play a role as well. It is possible that some of these covariates may change along with the calendar time associated with changes in achievement, but they are expected to remain relatively stable over time. Socio-economic status is expected to remain especially consistent over time, as is encouragement to read within the home unless an intervention is put in place that will increase this support for students. Attention and motivation factors can wax and wane over time, but again, are expected to remain relatively consistent for each student The process of learning to read, within each skill set (e.g., decoding fluency development), is expected to be fairly continuous, although it could certainly be impacted by major life events (e.g., an acquired brain injury, a death in the family, etc.). Finally, it is expected that there will be meaningful inter-individual variability in change, as each student presents with their own individual reading trajectory.

Results

Paired-samples t-tests were used to compare whether the report participants gained more over the school year than did the comparison participants. The data met all of the assumptions associated with this analysis (continuous dependent variables, categorical independent variable, free from outliers. The assumption of normal distribution for the dependent variables will be reported below.

PRF Progress Analysis

The data for both groups' PRF progress appear to be normally distributed. For the report group's PRF Progress, values for skewness (.167) and kurtosis (-.911) fell within acceptable limits (Field, 2009). Additionally, the Shapiro-Wilk statistic was not

significant (p = .391), suggesting that the distribution of the data does not differ significantly from the normal curve. Likewise, for the matched group PRF Progress variable, skewness (-.493) and kurtosis (-.874) values were also found to be within acceptable limits, and the Shapiro-Wilk statistic was not significant (p = .202). It should be noted that there was some missing CBM data for this analysis, resulting in a sample size of 16 in each group (report group and comparison group). The paired samples were significantly correlated (correlation coefficient = .845, $p < .001$). The result of the paired-samples t-test for the PRF measurement showed that there was a significant difference in the amount of progress shown by the two groups, but that in fact, the comparison group demonstrated greater progress over the Grade 3 school year, $t(15) = -4.061$, p = .001, d = .57. Means and standard deviations are presented in Table 3.

MAZE Progress Analysis

The data for both groups' MAZE Fluency progress appear to be normally distributed. Values for skewness (-.147) and kurtosis (-1.322) fell within acceptable limits (Field, 2009). Additionally, the Shapiro-Wilk statistic was not significant (p = .323), suggesting that the distribution of the data does not differ significantly from the normal curve. Likewise, for the matched group MAZE Fluency Progress variable, skewness (-.529) and kurtosis (-.531) values were also found to be within acceptable limits, and the Shapiro-Wilk statistic was not significant (p = .466). It should be noted that there was also some missing data for this analysis, resulting in a sample size of 11 in each group. In this case, the paired samples were not correlated (correlation coefficient = -.070, p = .839). The result of the paired-samples t-test for the MAZE Fluency measurement

showed no significant difference in the amount of progress shown by the two groups, $t(11) = -.250$, $p = .807$, $d = .11$. Means and standard deviations are presented in Table 3.

Discussion

Overall, the hypothesis that the report group would make greater school-year gains on the CBMs was not supported. For the PRF measure, the comparison group made greater gains, and for the MAZE Fluency measure, the gains made by the two groups were not significantly different from each other. There are a number of possible reasons for the lack of expected findings, ranging from methodological limitations to interesting implications of a contrary finding in and of itself.

It may be that psychological assessment reports are not always a useful tool to teachers and other school professionals in planning and implementing interventions for children who are at risk for reading problems, especially within a school board where all students are already accessing well-validated reading intervention. The purpose of the present study was to determine whether there is value added by the findings from a psychological assessment (as is recommended by proponents of the cognitive assessment approach; e.g., Hale et al., 2004) on top of the ongoing data that are collected by teachers on a weekly basis within the classroom setting as part of the PALS intervention (as in the RTI method; Fletcher, 2006) in identifying students with reading disabilities. In the year immediately following identification with an assessment report, the students who were identified with reading disabilities did not demonstrate improvements in reading scores when compared to students who were similarly struggling at the beginning of the school year. This could mean that in the present sample having access to a formal report did not better inform the teachers about the students' individual learning needs to the point where

knowledge of the students' individual strengths and weaknesses translated into better reading scores. It is possible that even though the reports are explained to the classroom teachers in a feedback session by the psychologist who completed the assessment, and the opportunity to ask questions is provided, the teachers may not understand how to translate the results of the report and the recommendations included in it into actual teaching practice, or possible as well that they may not have the resources to do so. Teachers may even reject the validity of the report's suggestions, therefore not implementing them because they feel them to be less-than-useful. Further education for teachers and educators on being a consumer of a psychological assessment report or clearer recommendations included therein might help to bridge this potential gap between reporting the problems and working to correct them.

Curriculum-based measurement may be a good alternative way to monitor students' progress and to document their needs for specialized services. The WECDSB is in a unique position, in that the teachers already have access to a great deal of information on their students' learning needs from the CBMs that are regularly collected. Perhaps educators and administrators within the school board can trust these CBMs to provide ongoing information about progress made in reading, consistent with an RTI-type model. It is important to note that all students within the school board system are getting access to a well-validated intervention program, and that a formal diagnosis is not required for classroom accommodations within this system, so the students are not disadvantaged by not having a formal diagnosis. Rather, delaying this diagnosis to the later years of school when administration of PALS has stopped could be beneficial within this particular school board, as this is the time when the at-risk students will need

additional instruction beyond that which they are already receiving along with their typically-achieving peers. One downside to doing this, however, is that it may be unethical to delay formal testing if it is known that a student is really struggling to learn to read before the end of the PALS program. For those students, formal evaluation might still be warranted. As well, if a diagnosis of a reading disability was applied to a particular student, it would necessary to stress to his/her teaching team that this diagnosis should not mean that we "give up" on trying to remediate reading problems that persist beyond the early years of intervention. Rather, ongoing efforts should continue to be made. In other words, the reading disability label is simply a diagnostic tool for describing the student's difficulties.

Limitations

Alternatively, however, it is possible that the null and contrary findings were at least partially due to methodological limitations. A steeper reading trajectory may have been found in the non-report group because these were indeed not disabled readers. It is unfortunate that the originally planned analysis could not be completed, as a complete sample of students who all present with learning disabilities, but who received more or less comprehensive reports, would have made for a better control. While efforts were made to control for initial reading ability, this is far from a perfect proxy for a match in reading achievement levels over time. It is indeed possible that the comparison group students were qualitatively different from the report group participants, as the latter group was known to be diagnosed as having learning disabilities, whereas the LD-status of the comparison group was unclear. The report group was known to present with LDs (as this information was included within their reports if their data were used in the analysis), so

this may not represent a wholly valid comparison. Again efforts were made to control for the effects of higher initial reading scores, but this method was not ideal.

For the MAZE Fluency measure, neither of the two groups made a great deal of gain at all on this measure over the school year. It could be that the MAZE Fluency CBM is not sensitive enough to pick up on changes over only one school year, as there appears to be a floor effect associated with this measure. Additionally, because this measure is only administered in January and June, full-year trajectories could not be calculated for the present analysis. The MAZE CBM might lend itself better to trajectory analyses that are cross-year in nature, and if there was a September administration point, this would allow for full-year trajectories to be calculated. Based on the lack of progress demonstrated by both groups over the school year, it is recommended that if MAZE Fluency is to be used as a measure of reading progress, that a September (or at least earlier-than-January) administration point should be added, and there should be a goal to take into account more than one year's worth of data. The MAZE Fluency measure is administered in both Grade 2 and Grade 3 within the school board, so cross-year comparisons should be possible for students who attended schools within this board for both the Grades 2 and 3 school years. This should contribute to a more nuanced understanding of the meaning of the MAZE Fluency CBM data.

Due to small sample size, it is also likely that the study lacked sufficient statistical power to find the expected effects on both measures. The statistical power of the two (two-tailed) t-test analyses can be calculated based on the observed sample sizes, the set α level, and the observed effect sizes. For the PRF analysis, the statistical power was equal to 0.73, and for the MAZE Fluency analysis, statistical power was equal to 0.062, both of

which are quite low, with the MAZE Fluency power value being especially problematic. In this way, it is unfortunate that the analyses were reduced to such small sample sizes due to both missing data and most reports occurring either during or just prior to the Grade 3 school year, thereby limiting the numbers further because it was not possible to find matched controls by class for all of the reports that were reviewed, nor was it even possible to match all of the students' reports with their own CBM data due to issues with missing data.

Another limitation of the present study was the fact that it was not known whether or not the students in the comparison group also had undergone psychological testing. It is known that they did not receive an assessment within the school board, but this is not the only avenue for receiving such an assessment report. It is possible that these students may have undergone private assessments, or assessments at community agencies such as the Regional Children's Center or Children First. This possibility cannot be ruled out, and this potential confound might also serve to at least partially explain these somewhat unclear results. In other words, it is not known whether the comparison group participants may have been diagnosed outside of the school board with either a learning disability or another condition that could impact upon their reading trajectories (for example, Intellectual Disability).

Another limitation inherent in the present study is the possibility of regression toward the mean. In describing their methodology, Protopapas and colleagues (2011) discussed how they planned to avoid the confound of using identical criterion and outcome variables, as they felt that doing this could obscure any possible divergence in reading trajectories. Unfortunately, within the present work, using the same criterion and

outcome variables was an unavoidable design flaw. Therefore, regression toward the mean needs to be acknowledged within the context of tracking achievement over the school year.

Interpretations

Given these limitations, it is not possible to say with certainty what these results mean when considered within the context of the ongoing responsiveness-to-intervention vs. psychological assessment report debate, but the potential importance of a null/contrary result can be discussed. Fletcher (2006) purported that a responsiveness-to-intervention approach to identifying LDs is best practice, as it is not a wait-to-fail model. The CBMs lend themselves well to this type of model, as they represent ongoing progress monitoring and a way to identify students whose skills are lagging behind and who therefore are likely to be in need of more intensive remediation. On the other hand, Hale and colleagues (2004) stressed that well-validated measures of cognitive and academic strengths and weaknesses are necessary to conclusively demonstrate the presence or absence of a learning disability. In a follow-up study, Hale and colleagues (2006) suggested that a multi-tiered method for identifying LDs may be employed. In this model, standard instruction is first provided to all students, followed by more intensive remediation for those students who fail to respond to standard instruction. Finally, those students who continue to fail to respond even following more intensive remediation are subjected to psychological assessment and a report that supports the presence of the LD is generated, if this diagnosis proves appropriate.

The multi-tiered model is employed within the school board, in that only those students who fail to respond to intervention that is provided over multiple school years

are subjected to psychological assessment. This is the reason why there were not many reports completed with students in the younger grades; ongoing efforts for the prevention of learning problems were made through use of the PALS program. It was most often the case that students who were tested following persistent failure to respond were indeed diagnosed with reading disabilities, demonstrating good judgment on the part of those who decided to implement assessment procedures. The present results, however, suggest that psychological assessment may not translate to improvements in early reading scores, at least not in the year immediately following the implementation of report-based recommendations. This finding should be considered when deciding whether or not the cost of this type of assessment is justified. The school board might consider delaying completing a psychological assessment until after the students have aged out of PALS, or as they are just about to do so. Having access to a psychological assessment evaluation might prove useful once teachers are no longer able to rely on the PALS program, starting in Grade 4 when administration generally stops for most students. The data associated with the present study simply suggest that having a report may not translate to improved scores on the CBMs in the year following the implementation of the report. These results cannot speak to the utility of the report for understanding students' learning needs and remediating reading problems beyond PALS administration. It is possible that when teachers are not delivering a well-validated intervention program to all students within their classroom, they may benefit from having a report that describes the students' strengths and weaknesses. Therefore, a recommendation to delay formal LD testing within the WECDSB to the time when students are about to age out of the PALS program could be considered. This is what was done in the majority of cases already, and the

present results lend some support to the idea of waiting until after PALS administration stops to implement report-based recommendations for reading remediation.

Additionally, according to the Learning Disabilities Association of Canada (2002) and the Learning Disabilities Association of Ontario (2011) a learning disability is a lifelong condition for most people, although it can be accommodated to the point where a diagnosis would not be made once strategies for accommodating problems are learned, and given an appropriate match between demands of the environment and the individual's characteristics. Therefore, it is possible that teachers see these students as being less able to make progress in learning than their peers, which even if true, could set up a self- or teacher-fulfilling prophecy of underachievement. If all students are treated as early learners who have the potential to achieve up until the end of PALS administration, then teachers would need to delay the shift from accommodating reading problems to providing curriculum modifications. Students would be seen as simply struggling learners up until the end of PALS administration, which would mean that the teacher remains responsible for trying to educate those students.

The practice of waiting until after the end of the PALS program, although a potentially good practice for the discussed reasons, proved to be an unfortunate circumstance for the present study, as it meant that the sample size was small due to there being so few reports that were completed during the years of PALS administration. Continued use of the PALS program is often included as a recommendation in the reports of students who are diagnosed with reading disabilities upon psychological assessment. These data, however, could not be located within the dataset as a whole. It is likely that these students are indeed continuing to be exposed to the concepts associated with the

PALS program, but their CBM data are not recorded in a manner that was amenable to analysis here (i.e., these data were not included within the CBM data sheets provided by classroom teachers). It is expected that these data are not included among the regularly-charted and readily-available classroom spreadsheets if these students are enrolled in a class that is not administering the PALS program to the entire class. It would be very interesting to have had access to the ongoing CBM data of those students who continued to benefit from the PALS program, as this could have answered the question as to whether improvements in reading scores are observed over a longer period of time following the implementation of report-based recommendations.

Future Directions

Future studies could aim for a more controlled analysis, wherein for example, the comparison group participants are on a waitlist for a learning disability assessment, but they have not yet received it. This would contribute to a better-matched control group than one that is based solely on initial reading scores, as there are many explanations for low reading scores other than a learning disability diagnosis. Additionally, an analysis that takes into account longitudinal data across school years could be useful as well. It would be interesting to follow all students who are struggling to achieve in reading, starting from the very early years of schooling and continuing on to the high school years, to determine whether gaining access to a report that describes their strengths and weaknesses does benefit their long-term reading achievement, as it is possible that the expected results were not observed because the students did not have adequate time to demonstrate the eventual improvements in reading that may have followed from the implementation of report-based recommendations. Finally, a cost-benefit analysis could

be useful to look at the expense of a psychological assessment versus the cost of intervention services being provided to students who may not require the extra help in reading beyond the administration of PALS programming.

CHAPTER III

Summer Learning Loss as Measured by Curriculum-Based Measurement Data

The Matthew Effect theory of reading suggests that those children who read well and have large vocabularies will read more often, learn more new words, and continually enhance their reading skills, while those children who read slowly and do not have adequate vocabularies experience a flatter learning curve which impedes upon further growth in reading (Stanovich, 1986). Assuming that appropriate motivation to learn is present, on the "rich get richer" side, good readers are more likely to be exposed to print materials and to practice reading skills, while students with reading disabilities tend to fall into the latter "poor get poorer" group in terms of reading achievement, as they are less likely to be exposed to print materials and to practice reading skills (Grant, Wilson, & Gottardo, 2007).

There has been some recent research that appears somewhat counter to the assumptions of the Matthew Effect theory. Protopapas, Sideris, Mouzaki, and Simos (2011) studied this effect in a cross-sequential design where they followed a large sample (N = 587) of Greek students starting in Grades 2 through 4 for a two-year period, with a specific interest in tracking reading comprehension. These authors employed a hierarchical linear modelling analysis, finding that the low and high ability groups demonstrated significantly different starting points, but that their achievement trajectories on a measure of passage comprehension did not suggest that the achievement gap would widen over time, contrary to the Matthew Effect. They did find, however, that there also was no evidence that the achievement gap would fully close. Protopapas and colleagues (2011) concluded in their abstract that "although the poor students may not be getting

poorer, they do not get sufficiently richer either" (p. 402) and in their discussion of educational implications, they stated that "*partial* manifestation of the [Matthew Effect] model is severe enough to warrant remedial action" (p. 418). They went on to mention that students who score low on reading achievement may never catch up to their high- and typically-achieving peers, and that this has very important educational implications that point to the need for early intervention, as the sample that they studied was somewhat older. Their study certainly adds to the literature base on the Matthew Effect phenomenon, and Protopapas and colleagues (2011) appropriately did not stretch their findings beyond the age range and reading components that they directly studied.

Similarly, a large-scale meta-analysis conducted by Pfost, Hattie, Dörfler, and Artelt (2014) found inconsistent evidence for the Matthew Effect, finding that it is more likely to be observed in the presence of moderating factors, such as when looking at measures of decoding efficiency, vocabulary, and composite reading scores, but only when the achievement tests were not affected by deficits in measurement precision. The PALS CBMs are expected to be fairly precise measurements of reading achievement, although it is possible that inconsistency in teachers' administrations of the probes may have affected the results. In theory, the CBM probes are thought to provide an accurate and precise measurement of reading achievement (Deno, Fuchs, Marston, & Shin, 2001). However, there are scaling problems with the CBMs as well, which will be discussed in the Discussion section of the present work.

Any manifestation of the Matthew Effect, whether in full or in part, is expected to be especially likely during the summer months when students do not have easy access to printed material and encouragement to read and help with reading are not as readily

available. In other words, students who have trouble with the mastery of early reading skills are more likely to fall further behind during the summer months, given the possibility for the previously described Matthew Effect in reading. Summer learning loss, or summer regression, refers to a decline in test scores and/or a loss of memory for acquired skills that occurs during the two-month summer break from school that is part of the traditional school calendar. According to Cooper (2003), one concern associated with students taking a summer vacation is the assumption that children learn best when instruction is continuous. Cooper (2003) posited that an extended break in the school year disrupts the rhythm of instruction, leads to forgetting, and requires a significant amount of review when students return to school in the fall.

In a meta-analysis of 39 articles dealing with summer learning loss, it was reported that, on average, one month's worth of information tends to be lost by students in general over the summer (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996). It should be noted that Cooper and colleagues (1996) were working with data collected within American schools, which traditionally have a shorter school year (by approximately 20 days, most of these during the month of June), as compared to Canadian schools. These extra 20 days are likely to be beneficial for Canadian students' learning by having a protective effect against summer regression. As will be addressed in this paper, however, even Canadian students tend to demonstrate summer regression in specific domains of reading achievement (Menard & Wilson, 2013).

Cooper and colleagues (1996) also describe group similarities and differences in summer learning loss. For example, these authors report that the severity of observed summer regression tends to increase with students' grade level. Cooper and colleagues

(1996) do specify, however, that this phenomenon could be explained by a floor effect wherein students at the lower grades are already scoring at the bottom of the range of possible scores and therefore it is more difficult to chart regression in the earlier grades. Cooper and colleagues (1996) also found that students of lower socio-economic status tend to demonstrate greater regression over the summer months, a finding which was further supported by Burkam, Ready, Lee, and LoGerfo (2004). This effect is likely due to their relative lack of learning resources outside of school. Finally, Cooper and colleagues (1996) found that summer learning loss does not differ by gender or race as long as socio-economic status is held constant.

The incidence of additive summer regression across many years appears to have long-lasting negative effects on students who lose or forget a substantial amount of information while out of school. Alexander, Entwisle, and Olson (2007) conducted an examination of the consequences of seasonal learning differences during the elementary school years on academic achievement during the high school years. Knowing that socio-economic status is an important predictor of summer regression (Cooper et al., 1996), these authors separated their sample into three groups: low, mid, and high socio-economic status. They then charted the students' initial test scores on a measure of reading comprehension at the beginning of Grade 1, the gains (or losses) that the students made over the school year and over the summer, and their test scores at the end of Grade 9. Alexander and colleagues (2007) found that the high-socio-economic status group demonstrated significantly better reading comprehension in Grade 1 and in Grade 9, and made significantly greater gains during the summer months. However, the gains made during the school year by the high-socio-economic status group were not significantly

different from those made by the low-socio-economic status group. This suggests that the summer is indeed an important time for students of low-socio-economic status, as they appear to make comparable gains during the school year, but their out-of-school learning tends to lag behind that of their high-socio-economic status peers. This is important, given the finding that achievement in high school is further related to long-term educational attainment and performance (Baron & Norman, 1992), which research shows is related to long-term labour market outcomes (Allmendinger, 1989). It is currently unclear whether such cumulative effects of summer regression would affect other at-risk groups, but it is expected that a similar pattern would emerge for students who are at risk for reading problems as for students of low-socio-economic status.

In terms of group-level differences between students with reading disability diagnoses and those without this diagnosis, the literature is sparse. Although it is generally recognized that students with reading disabilities or other special education placements may require extra help to catch up following the summer break (Sargent & Fidler, 1987; Katsiyannis, 1991), very few systematic studies have been conducted to determine the extent to which students with reading disabilities fall further behind their peers over the summer. Shaw (1982) compared the relative rates of retention of arithmetic and reading skills of students with a learning disability diagnosis and those without, finding that the diagnosed students tended to regress to a greater extent in both subject areas than did their non-learning disabled peers. This served to provide early evidence that students with a learning disability diagnosis do in fact tend to lose more information over the summer than do their non-learning disabled peers.

As a follow-up to Shaw (1982), Menard and Wilson (2013) conducted a similar study comparing students with reading disability diagnoses to students without this diagnosis, this time using standardized and well-validated measures of reading and language achievement and phonological processing. The scores of the control group students either remained stable or increased over the summer, while the scores of the reading disabled group regressed only on certain measures. Specifically, significant regression was noted on those measures which were timed and therefore required students to read with fluency and automaticity. In other words, students with reading disabilities tended to lose the ability to apply their phonological processing skills in an automatic manner, but when given additional time, they remained able to read as effectively as they had before the summer break.

One of the limitations of Menard and Wilson (2013) was the lack of sensitivity to change of the standardized measures used. Specifically, some tests had steep item gradients due to raw scores being converted into standard scores. This means that students must demonstrate a relatively large change in their raw scores in order for these changes to translate to a large difference in standard scores. In other words, a large degree of change at the raw score level translates into a much smaller degree of change at the standard score level, meaning that the test is less sensitive to very small changes in the number of items answered correctly. Therefore, changes in standard scores may not be appropriate for charting change over a relatively short period of time. Additionally, because parallel forms were not available for all measures, it is possible that practice effects could have masked significant regression that would otherwise have been observed. The present study will seek to extend and refine the findings of Menard and

Wilson (2013) by using Curriculum-based measurement (CBM) as the method for measuring regression in students who are and are not at risk for reading problems.

CBM has been shown to provide reliable and valid information about students' progress (Deno, Fuchs, Marston, & Shin, 2001) and it is expected that this method will be more sensitive to change than standardized measures, as the item gradient is not as steep due to the fact that there are generally more items per test and the raw scores are not converted into standard scores. Furthermore, multiple data points for each child are available as CBM is measured on an ongoing and regular basis. Thus, CBM allows for the analysis of learning trajectories, and so the present study can include a longitudinal component in which change is charted not only over the summer, but also over the school year. Consistent with Shaw (1982) and Menard and Wilson (2013) it is expected that the group of students with the lowest initial performance in reading will demonstrate the greatest magnitude of regression over the summer on the CBM measures of reading skills, while the CBM scores of average- and high-achieving students will demonstrate less severe regression. It is also expected that the students who struggle in reading will take longer to recover from summer regression and will have a significantly flatter school-year learning trajectory in reading.

Method

Participants and Procedure

The data for the current study were archival CBM data that are regularly collected by classroom teachers within the WECDSB. Data were taken from CBM probes starting in Junior Kindergarten and up until Grade 3. Demographic information was not available for this study, as most of the data are de-identified.

Materials

The CBM probes associated with the Peer Assisted Learning Strategies (PALS) program (Fuchs, Fuchs, Mathes, & Simmons, 1997) were used to chart summer regression within the current study. The PALS program measures performance in four reading-related domains: Letter-Sound Fluency (LSF), Word Identification Fluency (WIF), Passage Reading Fluency (PRF), and Maze Fluency (MF). The PRF and MAZE data, however, were not appropriate for the present analysis. Although the PRF measure was administered across school years, it cannot be subjected to further analysis within the context of the present study, as each year when the students return to school, they are administered a more difficult passage. Reading a more difficult passage upon returning to school should lead to a lower score that is, in fact, not attributable to summer learning loss, but rather a function of the nature of the test. Because the inflation of the appearance of summer learning loss cannot be assumed to be uniform across achievement levels, the PRF test cannot be used for cross-year comparisons and therefore will not be considered further here. With respect to the MAZE measure, it is not useful in the assessment of summer learning loss because it was only administered in January and June in Grades 1 and 2; there was never a September administration point, so charting summer loss upon returning to school is not possible. Therefore, because the LSF and WIF scores are the only ones which can be subjected to summer loss analysis, and they are only administered in JK, SK, and Grade 1, achievement in these grade levels only will be considered for the main analyses herein.

Letter Sound Fluency. The LSF test assesses students' speed and accuracy in identifying letter sounds. This measure was used in the earlier grades (Junior

Kindergarten, Senior Kindergarten, and Grade 1) within the WECDSB sample. It was administered eight times at regular intervals throughout the school year. In completing this task, students are presented with a page of 26 letters and then given one minute to pronounce as many letter sounds as they are able to pronounce. This measure is scored by counting the number of correct responses to arrive at a total raw score. If the student reads all or part of the list in under one minute, a correction formula is applied as follows: (number of letters read correctly/number of seconds it took to read list) x 60 = estimated number of letters read correctly in one minute. In the end, the test measures the number of letters that students are able to read in one minute, regardless of whether they can decode all of the letters on the page.

Word Identification Fluency. The WIF test assesses the students' speed and accuracy in identifying whole words, and it is also intended for early readers. Within the present sample, it was administered to students in SK and Grade 1. Within the SK sample, WIF was administered twice (January and June measurements), and within the Grade 1 sample, it was administered seven times at regular intervals throughout the school year. In completing the WIF measure, the student is presented with a list of 100 words and he or she is given one minute in which to read as many words as possible from the list. This task is scored by assigning a value of '1' to every word correctly read and '0' to those not attempted or read incorrectly. If the students finishes reading the list in under one minute, the same correction formula as described above is applied.

Data Analysis

CBM data were available across multiple school years and multiple summers for each student. Specifically, the data from across five school years and four summers (JK

to Grade 3) were available, but only the data from JK, SK, and Grade 1 were considered due to the problems with using the PRF and MAZE Fluency data mentioned previously. The data were accessed in multiple forms. For the 2009-2010 and 2010-2011 school years, all of the data had been aggregated into one spreadsheet. These two spreadsheets, therefore, only needed to have duplicate and nonsensical data removed (see data cleaning section below). The data for the 2011-2012 and 2012-2013 school years, however, came in the form of one spreadsheet per participating classroom teacher. Each teacher had submitted end-of-year data that showed his/her students' CBM scores over the school year. The columns on these spreadsheets were inconsistently titled, so it was necessary to run each of the approximately 400 spreadsheets per year through a specially-designed transformation sheet. This transformation sheet made it so that all inconsistent column headers were recognized. The inconsistent column headers were manually converted when necessary, and then manually copied/pasted to the aggregated spreadsheets (one per school year). The data were then matched across school years using the VLOOKUP procedure on Microsoft Excel 2010, which matched the data based on Ontario Education Number whenever possible, with a secondary matching rule criterion using WECDSB student number where OEN was not available and student number was available. In all, data from 10,821 total participants were included in the analysis, although not all of these students had longitudinal data. The data were then grouped by grades instead of school years. Specifically, all of the JK data were grouped together, then all of the SK data were grouped together and entered next to the JK data in the same spreadsheet, and so on, regardless of what academic year each student was enrolled in each grade. In other

words, the data were then grouped by the grade during which it was collected, rather than by the school year in which it was collected.

Data cleaning. Next, it was necessary to remove text-based data (e.g., where a teacher had written 'A' for absent, the entire word 'absent,' 'V' or 'vacation,' or a number of other text-based data that were not appropriate for this analysis). This was done using a find+replace strategy. Then, duplicate data were removed. As mentioned, the students were identified by Ontario Education Number and WECDSB student number within the dataset. Although these numbers should be unique, duplicate data were found. Where there was an identifiable explanation for the duplication, the data were copied into one line and left in the master sheet. The most common example of an identifiable reason for duplicate data was when the student had moved to a new school within the WECDSB, as evidenced by his/her data ceasing to be collected at one school at the same time as they appeared in another school's data sheets. When the reason for the duplication could not be identified, both lines of data were deleted from the analysis. It was much more common that a reason for the duplication could be identified than not. Duplicate data were identified using a "highlight cell rule" under the Conditional Formatting option on Excel, and then the data were visually scanned through for highlighted cells. Next, nonsensical data were removed. This most often involved data where number-based typos were likely made. For example, on the LSF test, a score of over 200 is highly unlikely, to the point of being unreasonable, as this would represent letter-sound decoding at a rate of over 200 sounds produced per minute by beginning readers. It should be acknowledged that an arbitrary cutoff score (120 letter sounds/minute) was used. This same cutoff score was used on the WIF test, as it was thought to be highly unlikely that a reading rate of

over 120 words per minute was an achievable score by beginning readers, and would therefore be considered outlying data.

Calculation of summer learning loss. Baseline (beginning-of-year) scores were calculated by averaging the two baseline points that were collected at the beginning of the school year in September. Outcome scores were obtained from the final point of administration in June of the preceding year. The extent of summer learning loss was calculated as the difference between the June outcome score prior to summer, and the September baseline score following the end of summer break. Larger negative scores on the learning loss variable indicate more summer loss. Summer gains, where observed, are indicated by positive values.

Calculation of time to recover from summer loss. For each student, an outcome score was obtained at the final point of administration in June. Over the subsequent year, following the learning loss event, that student's progress was tracked at regular intervals. The time taken to recover from summer learning loss (i.e., to again attain achievement at the June outcome level) was recorded as (t_x). In other words, t_x is the time in days from the beginning of the school year that it takes each student to again reach a score greater than or equal to their outcome score of the previous year. If the student's test score at t_x was perfectly matched with the outcome score of the previous year, then t_x was the student's time to recover from summer learning loss. If a student's score at t_x was slightly higher than his/her outcome score of the previous year, then an interpolation was done between the test score at t_x and the test score at the time point preceding it (t_{x-1}) to calculate the student's time to recover from summer learning loss. This interpolation calculation was necessary because there were only 9 sampling points throughout the year.

If the students did not regain the learning that they had lost over the subsequent school year at all, then they were assigned a score of 300 for the time to recover variable, which was meant to express that they were still in the same place at the end of a 300-day school year. Similarly, any students who maintained the same scores when they returned in the subsequent fall, or had made summer gains, were assigned a score of zero for time to recover, indicating that they needed no time to recover from summer learning loss.

Calculation of school-year trajectory. Students' individual trajectories over the school year were calculated by subtracting their baseline scores from their outcome scores later in the same school year. It should be noted that the school year trajectory is reported as the gain in test score over the school year (not dividing by time and getting a slope), as this simplifies the discussion.

Grouping method. It was expected that the group of low-achieving readers would demonstrate significant regression over the summer on the CBM measures of reading skills, while the CBM scores of average- to high-achieving students would generally remain stable over the summer months. More specifically, it was expected that the data for those students who are poorer readers would show significantly deeper valleys following the summer break, and that this group's reading scores would recover at a slower rate than will the post-summer reading scores of their non-reading-impaired peers. It was also expected that the learning trajectories of the students at risk for reading disabilities would not be as steep during the school year as those of their non-impaired peers. Hence, for students with reading problems, the summer vacation will likely be a time when, every year, their reading skills would dip lower and they would less readily recover from this summer regression across domains of reading achievement.

In order to compare Low, Average, and High achievers in reading, it was necessary to group the students' data thusly. Reading group status was defined on a standard deviation basis, wherein those students who fell more than one standard deviation below the mean on outcome scores at the end of each year were considered to fall within the "Low" group, those who fell within one standard deviation above or below the mean were considered to fall within the "Average" group, and those whose outcome scores fell above one standard deviation from the mean outcome score were considered to fall within the "High" group. While this grouping method did lead to a larger cell size for the Average group (by definition, as the majority of students should fall within one standard deviation of the mean), this grouping method was considered to be superior over other methods. In an article by McMaster, Fuchs, Fuchs, and Compton (2002), these authors outline a number of different ways to define non-responsiveness to intervention. One is by assuming that students who are reading fewer than 40 words per minute by the end of Grade 1 are seriously deficient in reading. The one standard deviation-based grouping method fell in line with this assumption, as reading at a rate of 40 words per minute at the end of Grade 1 coincided with falling approximately one standard deviation below the mean. Another way to group non-responders according to McMaster and colleagues (2002) is to consider those who are greater than .5 standard deviations below the mean as being reading-impaired. Their study, however, included fewer participants than the present one, and by having a greater number of participants, it was thought that a more extreme cutoff would contribute to a more thorough investigation of the reading trajectories of the most impaired readers. Finally, McMaster and colleagues (2002) outlined a proposed method that considers both slope and intercept when considering the

difference between responders and non-responders (the dual-discrepancy approach). However, this approach was deemed not to be appropriate for the present analysis because it was too rigid. McMaster and colleagues (2002) set out to identify non-responders to intervention, whereas the purpose of the present study was to identify low achievers. Using the dual-discrepancy method would have only identified the non-responders, as the students who are identified need to be low on both performance level and learning trajectory. In the end, a performance-level cutoff of one standard deviation above or below the mean was used to group the students.

In both JK and SK, the students were grouped based on their end-of-year outcome scores. Re-grouping was necessary between school years as some students move to a new group after a year passes. Descriptive data regarding the average absolute levels of achievement by the three groups at key time points (i.e., transitions between school years) are presented in Table 4. There were 1029 students with complete longitudinal data spanning the JK-to-SK summer, referred to in Table 4 as Longitudinal Group A. There were 939 students with complete longitudinal data spanning the SK-to-Grade 1 summer, referred to in Table 4 as Longitudinal group B. The distinction between A and B was necessary because some students changed achievement groups across the years, so they have to be regrouped each year so that the groupings are as accurate as possible. So, Longitudinal Group A reflects the groupings of students based on JK LSF outcome data, while Longitudinal Group B reflects the groupings of students based on SK LSF data. Longitudinal Groups A and B have only 199 students in common, representing a 20% overlap in the number of students whose scores are included in both longitudinal groups. Based on JK outcome (Longitudinal Group A), Average students had LSF scores

between 11.0 and 48.6 (mean = 29.8, SD = 18.8, n = 1940). Based on SK outcome (Longitudinal Group B), Average students had LSF scores from 30.1 to 78.9 (mean = 54.5, SD=24.4, n=2029). In both cases, High and Low achieving students were those whose scores fell above or below these ranges.

Main statistical analysis. The data were analysed using ANOVA, where the independent variable in each case was the Low/Average/High reading ability status. There were three dependent variables of interest: (1) summer learning loss, (2) time required to recover losses in learning, and (3) steepness of learning trajectory over the school year. For the summer learning loss analysis, students were grouped based on their performance at the end of each school year and then the absolute amount of loss was examined between that year and the subsequent year. Comparisons of this type were performed three times; between JK and SK on the LSF measure, between SK and Grade 1 on the LSF measure, and between SK and Grade 1 on the WIF measures. Three time to recover analyses were also performed, using the interpolated time to recover measurement described previously for the same three measures. Finally, three school year trajectory analyses were also performed; one for each of the measures listed above.

Although it may have provided a more sophisticated design had a structural equation modeling strategy been used to analyze the present data, this type of strategy was not employed because an ANOVA-based strategy was thought to adequately and most parsimoniously address the research questions.

Checking for floor and ceiling effects. There was cause for concern about possible floor effects for the Low group over the JK summer, as their scores are quite low. However, all students had scores larger than zero prior to the summer break, and

only 6% of them decreased to a score of zero following the summer break. Because the floor effect only affects 6% of students, and because their summer loss scores are simply underestimated (as opposed to being not captured at all in the data), the floor effect should only have a small effect on the results, although possibly an important one. There was also cause for concern about a potential ceiling effect in the LSF scores of the High group in Grade 1. Upon further inspection, the learning trajectory increases at the same rate until the end of the year. If there had been a ceiling effect, the trajectory should have begun to form an asymptotic shape instead of the linear one that was observed. Thus, the data appear to be free from a ceiling effect, although it is important to acknowledge that a score in the higher range versus one in the average range may not be clinically meaningful on a measure of letter sound decoding (Kirby et al., 2010).

Estimation of the opportunity-costs of summer vacation. For this analysis, the summer learning trajectory was recalculated as the change in test score per month. These were usually negative, as the test scores decreased over the summer when there was a loss. The school year trajectory was then also recalculated on a monthly basis. This represented the rate at which the students learn while in school. The difference between the school year trajectory and the summer trajectory therefore provided information about the scores that the students might have achieved per month, had they been in school instead of on summer break. This number was then multiplied by two to get an estimate of the test scores that could have been achieved across the entire two-month summer break from instruction. The opportunity-cost was then standardized to the students' reading scores to give a more interpretable opportunity-cost as months' worth of reading skills that could have been achieved. The resultant score represents how many months'

worth of reading achievement that the students could have theoretically been further advanced, had there not been a summer break.

Missing data. In general, missing data were left as missing because this is a large-scale dataset, and cases were excluded if there were inadequate data to perform the analysis on that case (for example, if a student was missing both baseline measurements in the subsequent year, then his/her data were excluded from the analysis altogether). The exception to this is in calculating the interpolated data for the time to regain summer learning loss analysis. These data, if not precisely available, were calculated using the interpolation method described previously in order to arrive at a more precise date at which summer learning loss was regained.

Theoretical Model of Change

Collins (2006) outlined a number of elements that must be considered when articulating a theoretical model of change. With respect to the general or characteristic shape of change, it is expected that reading skills on both measures (LSF and WIF) will increase linearly over the school year. Although reading skills can develop in a more discrete, stage-like manner that is characterized by jumps in development overall, it is thought that each individual skill (e.g., decoding ability) develops continuously over time, increasing in a linear fashion. It is not expected that there will be periodicity or a cyclical nature associated with this change. The change in reading achievement over the school year is thought to be due to calendar time as well as the reading instruction that is provided over that calendar time to foster learning. The change in reading achievement over the summer is thoughts to be due to the combination of calendar time and the stoppage of continuous, intensive instruction in reading. The most relevant covariates that

may predict changes in school-year trajectory, summer loss, and time to recover from summer loss, aside from calendar time and the PALS intervention itself, are socio-economic status and reading encouragement within the home environment, outside of the PALS program. As well, attention factors and motivation to achieve may play a role as well. It is possible that some of these covariates may change along with the calendar time associated with changes in achievement, but they are expected to remain relatively stable over time. Socio-economic status is expected to remain especially consistent over time, and so is encouragement to read within the home unless an intervention is put in place that will increase this support for students. Attention and motivation factors can wax and wane over time, but again, are expected to remain relatively consistent for each student. The process of learning to read, within each skill set (e.g., decoding fluency development), is expected to be fairly continuous, although it could certainly be impacted by major life events (e.g., an acquired brain injury, a death in the family, etc.). Finally, it is expected that there will be meaningful inter-individual variability in change, as each student presents with their own individual school-year and summer reading trajectories.

Results

Beginning-of-year and end-of-year LSF and WIF scores are given in Table 4, providing information on the absolute reading abilities of each achievement group at key time points. Descriptive data for all main change score analyses are presented in Table 5. Table 4 is included to provide a reference to absolute reading scores so that the summer learning loss and school year trajectories in Table 5 can be compared to the absolute levels in Table 4 to better gauge their effects. Estimates for the opportunity-cost of summer break are presented in Table 6.

JK to SK LSF analysis

Testing of assumptions. The assumption of a normal distribution was examined by visual inspection of a Q-Q plot, which plots the expected normal distribution against the observed values for the actual distribution. The values are then examined visually for non-linear deviation from the line of best fit. The values for skewness and kurtosis were also considered, as significance tests are less useful when working with a large sample size (Field, 2009). The data did not appear to be normally distributed, showing instead a positively skewed, leptokurtic distribution. The assumption of homogeneity of variance was also violated, according to Levene's test, $F(2,936) = 69.93$, $p<.001$. Because both of these basic assumptions were seriously violated, a bootstrapping approach was employed using SPSS v.22 software (Field, 2009).

JK/SK LSF summer learning loss. Overall, a statistically significant difference was found in the degree of summer learning loss demonstrated by the groups, $F(2,936) = 208.19$, $p<.001$, $\omega^2 = .31$. Statistically significant differences (at less than the .001 level) were also found between all group pairs (Low-Average, Average-High, and Low-High) upon bootstrapped pairwise comparisons. It is clear that the greatest degree of summer learning loss was experienced by the High achieving group, contrary to the hypothesis that the Low achieving group would lose more of what they had learned over the summer.

JK/SK LSF time to recover from summer learning loss. An overall significant difference was also found in terms of the time needed to regain learning that was lost over the summer, $F(2,925) = 52.41$, $p < .001$, $\omega^2 = .10$. Again, all bootstrapped pairwise

comparisons were also significant (at less than the .001 level), with the High group needing the longest amount of time to recover from summer loss.

SK LSF school-year trajectory. An overall significant difference was also found in terms of the amount of gain made by the three groups of students during the year subsequent to the summer event (in this case, the SK year), $F(2,822) = 24.20$, $p<.001$, $\omega^2 = .053$. Again, all of the bootstrapped pairwise comparisons were significant (Low-Average $p<.001$, Average-High $p=.016$, and Low-High, $p<.001$). Students in the High group made the greatest amount of gain over the SK school year.

SK to Grade 1 LSF analysis

Testing of assumptions. Whether or not the data had met the assumption of normality was assessed again by visual inspection of a Q-Q plot and by considering the values for skewness and kurtosis. The data were again not normally distributed; this time showing a negatively skewed and leptokurtic distribution. The assumption of homogeneity of variance was also violated within the SK to Grade 1 summer learning loss data according to Levene's test, $F(2,1027) = 25.96$, $p < .001$. Again, a bootstrapping approach was employed using SPSS v.22 software in an attempt to correct for the violation of these assumptions in this data sample (Field, 2009).

SK/Grade 1 LSF summer learning loss. Overall, a statistically significant difference was found in the degree of summer learning loss demonstrated by the groups, $F(2,1027) = 178.86$, $p<.001$, $\omega^2 = .26$. Statistically significant differences (at less than the .001 level) were also found between all group pairs (Low-Average, Average-High, and Low-High) upon bootstrapped pairwise comparisons. This analysis also shows that the greatest degree of summer learning loss is demonstrated by the High achieving group,

contrary to the hypothesis that the Low achieving group would lose more of what they had learned over the summer.

SK/Grade 1 LSF time to recover from summer learning loss. An overall significant difference was also found in terms of the time needed to regain learning that was lost over the SK-to-Grade 1 summer on the LSF measure, $F(2,1035) = 117.16$, $p<.001$, $\omega^2 = .18$. Again, all bootstrapped pairwise comparisons were also significant (at less than the .001 level), with the High group needing the longest amount of time to recover from summer loss.

Grade 1 LSF school-year trajectory. Overall, no significant difference was found in terms of the amount of gain made by the three groups of students during the year subsequent to the summer event (in this case, the Grade 1 year), $F(2,978) = 0.28$, $p=.76$, $\omega^2 = .001$. In other words, there was no significant differences in the amount of gain made over the Grade 1 school year by the three groups. Therefore, the bootstrapped pairwise comparisons were not considered.

SK to Grade 1 WIF analysis

It should be noted that because of a serious floor effect on the WIF measure at the end of SK, the Low, Average, and High groupings were determined based on LSF scores at the end of the SK year, rather than based on SK WIF outcome scores.

Testing of assumptions. Again, based on examination of the Q-Q plot and skewness/kurtosis values, the WIF data again did not conform to a normal distribution; instead showing a negatively skewed, leptokurtic distribution. The assumption of homogeneity of variance was also violated based on the results of Levene's test,

$F(2,1031) = 66.51$, $p < .001$. Therefore, a bootstrapping method was again employed using SPSS v.22 software (Field, 2009).

SK/Grade 1 WIF summer learning loss. Overall, a statistically significant difference was found in the degree of summer learning loss demonstrated by the groups, $F(2,1031) = 8.07$, $p<.001$, $\omega^2 = .013$. Statistically significant differences (at less than the .05 level) were also found between all group pairs (Low-Average $p=.016$, Average-High $p=.007$, and Low-High, $p<.001$) upon bootstrapped pairwise comparisons. These results indicate that the Low group does not demonstrate gains that are consistent with the Average and High groups. This is consistent with the hypothesis that the Low group would fare worse over the summer than the Average and High groups.

SK/Grade 1 WIF time to regain summer learning loss. An overall significant difference was also found in terms of the time needed to regain learning that was lost over the SK-to-Grade 1 summer on the WIF measure, $F(2,915) = 3.25$, $p = .039$, $\omega^2 = .0049$. This time, not all of the bootstrapped pairwise comparisons were significant. Specifically, in terms of the Grade 1 WIF time needed to regain summer loss, the difference between the Low and Average group was significant ($p=.018$), as was the difference in trajectory between the Low and High group ($p=.036$), but the difference in time to regain summer loss between the Average and High group was not significant ($p=.63$). The means show that the Low group took longest to recover from summer learning loss, which is consistent with the initial hypothesis.

Grade 1 WIF school-year trajectory. Overall, a significant difference was found in the amount of gain made by the three groups of students on the WIF measure during the year subsequent to the summer event (in this case, the Grade 1 year), F(2,914) =

3.355, p = .035, ω^2 = .0051. As with the Grade 1 WIF time to regain analysis, not all of the bootstrapped pairwise comparisons were significant. Specifically, with respect to the Grade 1 WIF school-year trajectory, the difference between the Low and Average group was significant (p=.017), as was the difference in trajectory between the Low and High group (p=.033), but the difference in school year trajectory between the Average and High group was not significant (p=.63). Students in the Low group made the greatest gain in WIF scores over the Grade 1 school year, contrary to the initial hypothesis.

Post-hoc analysis: Transitioning between achievement groups

It is likely that the Low group includes both low achievers and non-responders, or those who would meet the dual-discrepancy definition of a non-responder described by McMaster and colleagues (2002). An attempt was made to remove the non-responders from the analysis. However, in doing so, the "non-responder" class was found to be highly unstable. Specifically, of those students who were identified by McMaster and colleagues' (2002) dual-discrepancy definition in one year, only approximately 59% (67/112 students) remain in the Low achieving group the next year. By the next year, approximately 39% of the sample (44/112 students) had moved up into the Average group, while approximately 1% (1/112 students) of the Low sample had actually improved enough to fall within the High group by the subsequent year. Given that the dual-discrepancy definition does not really seem to capture actual non-responders within this sample, the removal of students who were identified as non-responders was discontinued. This work did, however, spawn an additional question: Do students who are in the Low achieving group at the beginning of their academic careers (whether non-responders or not) tend to stay within the Low group over time?

This question was investigated by looking at the rate of low achievers who remain in the Low group over time. Both graphical and data-based methods were used. Figure 1 is a graphical representation of this. Data from SK were used, as by this time, students are expected to have learned and retained enough of the reading instruction that they have received. The separations in the bars represent the groupings across the years. As can be seen in Figure 1, students in the Low and in the High groups tend to regress toward the mean. Approximately half of the Low group students tend to enter the Average group by Grade 1 and this trend continues into Grade 3. It is thought that those who stay in the Low group are likely to be the true non-responders to intervention, thus suggesting a potential new method of identifying these students based on multi-year longitudinal data.

This finding is also demonstrated in number form in Table 7. Interpretation of the table is done as follows: The SK column indicates into which group the students fell during their SK year. The G1, G2 and G3 columns indicate the group to which these students have moved, in the subsequent years. The JK column indicates the group into which these students fell in the previous year. The values indicate the number of students in each group. For example, during the SK year, there were 332 students in the Low group (the values for Average and High within the "Low" category are zeroes because this is the baseline year, so no one had moved to a different group yet). By Grade 1, 84 of the 332 Low students (25%) remained in the Low group, 72 had moved to the Average group (22%), and 5 had moved to the High group (1.5%). The remaining 51.5% is accounted for by attrition, as the same students are followed every year, so some attrition is expected. Going forward another year, 36 of the 84 students that were in the Low group in Grade 1 remained in the Low group in Grade 2, 48 of the 72 remained in the

Average group, and 4 of the 5 remained in the High group. Every year, the number of students in the sample gets smaller, as these are the same students being followed forward in time; no new students are added to this particular analysis. Although simple statistical regression toward the mean is evident, it may yet be interesting to note that those students who consistently remain in the Low group based on their CBM scores may be the students who are best qualified for special education services.

Discussion

The present results describe the consequences of summer learning loss during the crucial early years of reading instruction. The summer break negatively impacts all three achievement groups, however, it has the largest effects on the LSF scores of the High achievement group over the JK-to-SK summer. The summer break most negatively affects the Low group over the SK-to-Grade 1 summer, in that they do not show gains that are consistent with those of their Average- and High-achieving peers..

Summer Learning Loss

With respect to summer loss of reading skills, overall, based on the early reading skill of gaining fluency at decoding letter sounds, low-achieving students did not lose more information over the summer months than their average- and high-achieving peers. In fact, the opposite trend was observed. The low group showed little loss in LSF scores over the summer, while the LSF scores of the high achieving group dropped by over 30%. When comparing the absolute levels of loss, it would be tempting to draw the conclusion that the students who demonstrate the highest levels of achievement are extremely disadvantaged by the summer break in the school year. While they do regress to a greater extent than their low-achieving peers, and are therefore most negatively

impacted by the summer break from a statistical standpoint, it is important to take into consideration the mean levels of achievement as well when evaluating the impact of these losses. For example, in looking at the mean levels of achievement on the LSF measure spanning the JK-to-SK summer vacation, it should be noted that although the High group experienced a large 23-point (-37%) loss over the summer – while the Low group lost only 0.6 points (-11%), the scores of the high group remain approximately 33 points higher (800%) than those students in the Low group following this summer loss event. Over the SK-to-Grade 1 summer, the same trend is seen in LSF scores. The Low group continues to demonstrate less regression on the LSF measure as compared to the Average and High groups, but the absolute scores of the two upper groups remain much higher than those of the students in the Low group.

It is possible that regression toward the mean effects accounts for at least some of this effect, in that the High group especially has much more to lose on the LSF measure because, by definition, their pre-summer scores are higher than those of students in the Average and Low-achieving groups. Relatedly, although there was only a small floor effect directly identified, it is certainly possible that the High group having that much higher scores (and therefore that much more to lose) means that regression toward the mean might create something like a floor effect phenomenon wherein the regression in the High group is attributable to this group's "floor scores" being significantly higher than that of the Low group.

It is also possible that although the high group does lose a significant amount of fluency in their letter sound decoding scores, they might still return to school in September at a level that would be considered "adequate" for reading fluently. According

to Kirby and colleagues (2010), there may not be a clinically significant difference between "Average" and "High" in terms of letter sound decoding skills. In other words, the utility of quick and accurate decoding of letter sounds might "top out" in the Average range. Therefore, it is possible that the losses in letter sound decoding that were found in the Low group's scores might actually be more clinically significant, although these losses were statistically smaller.

For the higher-order skills of word identification fluency, the Low group did not make gains consistent with those of the Average and High groups. The Low group demonstrates almost no change, on average, on word identification fluency over the summer months. Although the Low group students did not demonstrate the expected losses, the high-achieving students' scores appear to improve by approximately five words read per minute, on average, over the summer between SK and Grade 1. Recall that, by definition, the Average and High groups have a more solid foundation in reading than does the Low group, as they have higher absolute LSF and WIF scores. These results could point to the beginnings of a potential Matthew Effect in reading (Stanovich, 1986). The Low group likely does not seek out the same opportunities for print exposure over the summer months because reading is, an onerous task for struggling readers, so they tend not to practice their reading skills or take opportunities to read when reading is not actively encouraged (Grant, Wilson, & Gottardo, 2007), like during the summer months when they are out of school. Additionally, these results might represent a shift in the focus of reading instruction from teachers to parents over the summer months. It would be more intuitive for engaged parents to help their children to learn new words, rather than drilling them in LSF-type tasks. This difference in focus over the SK summer

might at least partially explain the large drop in LSF scores despite a gain in WIF scores. As previously mentioned, however, it is possible that the High and Average achievers are still returning to school at a level of letter sound decoding that we would deem "adequate" for fluent whole word and passage reading.

The present results complement the results of Menard and Wilson (2013), who tested students in Grades 4 to 6, thus giving the students enough time to have built a solid foundation in reading, or alternatively, to have failed to build a solid foundation in reading. Menard and Wilson (2013) found that by Grades 4 to 6, low-achieving readers demonstrate losses as compared to their typically-achieving peers. It is possible that the present results show the very beginning of when summer differences between low-, average-, and high-achieving students can begin to be measured; the transition from SK to Grade 1 may be a very important marker for later reading achievement, so special attention should be paid to students who continue to struggle at the onset of Grade 1. WECDSB staff might consider meeting at the end of the school year to note which SK students are continuing to struggle at the end of that year and then make sure that they are afforded extra help and attention upon returning to school for the Grade 1 year.

The improvement in WIF over the summer by high achieving students also aligns with a developmental theory of reading. Chall (1983) and Kaplan and Walpole (2005) posit that reading progresses in qualitatively distinct stages. From the beginning of reading acquisition in early elementary school, up through more advanced reading refinement into and beyond high school, students transition from learning the mechanics of reading to being able to develop more advanced and synthesized knowledge based on what they read. The Grade 1 year maps onto these stages by marking the transition from

learning basic decoding skills to learning to read with increased fluency, which is expected to occur at about age 6 or 7. Those students who have mastered early reading skills are more likely to practice more advanced skills such as whole-word reading over the summer vacation, and to therefore come back in the fall at both absolute and relative levels of achievement that exceed that of their struggling peers.

Time to Recover from Summer Loss

The results from the time to recover from summer loss partially supported the initial hypothesis that students in the Low achieving group would take longer than their Average- and High-achieving peers to recover from observed summer learning loss. Students in the Low group who did demonstrate losses recover at a significantly slower rate from summer losses in WIF despite rapid recovery in LSF scores. One hypothesis that could explain the slower recovery in WIF achievement in the Low group is that this group includes the students who present with reading disabilities. This would fit with the Responsiveness-to-Intervention model of identifying learning disabilities, which posits that reading disabilities should be identified only when students fail to respond to well-validated intervention efforts aimed at remediating problems in reading (Fuchs, Mock, Morgan, & Young, 2003). Given that low achievement in reading is another important diagnostic criterion for reading disability identification (American Psychiatric Association, 2013), it is likely that the Low achievement group is made up of a subset of students who present with reading disabilities. However, it is important to note that the difference in recovery time, although significant, is only approximately 10 days more for the Low group than the Average group. This suggests that students in general respond well to the PALS intervention, as they rapidly regain lost skills when they return to

school in September, at least as measured by the CBMs. In future studies, it would be interesting to investigate the response rates of students who are specifically diagnosed with reading disabilities, as the Low group in the present analysis is made up of a mix of "garden-variety" low achievers in reading (Gough & Tunmer, 1986), as well as students with identified reading disabilities. Unfortunately, it was not possible to access information on reading disability status for the present investigation.

School Year Trajectory

The present results showed that, on average, the students in the Low group sample learned less over the SK school year on the LSF measure, as compared to the Average and High groups. Thus, students in the Low group tend to demonstrate a shallower trajectory during the school year on this measure. In terms of the absolute scores, students in the Low group tend to trail their peers by learning to fluently decode about 10 fewer letter sounds over the SK year. In contrast, however, when looking at the Grade 1 school year WIF trajectory, the Low group tends to make the greatest amount of gain over the Grade 1 school year. It is very important to consider, however, than students in this group continue to present with much lower absolute scores as compared to those in the Average and High groups at Grade 1 outcome on this measure. Therefore, although they do tend to make greater gains over the Grade 1 school year, by no means do they catch up to the levels of achievement observed in their typically- and high-achieving peers. This is likely to be an important difference that would affect the rate of reading passages and could also affect later reading comprehension if more cognitive resources are being allocated to decoding if fluency is not yet achieved. As mentioned previously, it is suspected that the low absolute scores observed in this group are at least partially driven by students with

reading disabilities whose non-responsiveness to intervention negatively impacts their word reading fluency. Future studies should seek to separate students with reading disabilities from "garden-variety" low-achieving readers because these two groups represent qualitatively distinct classifications of readers (Gough & Tunmer, 1986). Comparisons between the absolute scores and the achievement trajectories of these two groups of low-achieving readers may yield interesting conclusions regarding their differences in level of responsiveness to intervention.

Interesting implications also arise when considering the LSF learning trajectory comparison between the Average and High groups, without considering the Low group's school year trajectories. Over the school year in SK, students in the High and Average groups demonstrate statistically similar school-year trajectories. When considered in accordance with the summer learning loss data, it appears as though the High group tends to fare worse than the Average group, as they do not make up for greater losses during the JK-to-SK summer by demonstrating a steeper learning trajectory as a compensatory mechanism. In other words, the gap between the Average and High groups appears to narrow during the summer between SK and Grade 1, and by the end of the Grade 1 year, the High students have not caught up to maintaining the same degree of competitive "edge" that they had over the Average-achieving students, although the High group's mean absolute level of achievement does remain higher than that of the Average group. Furthermore, only 55% of students from the high group at the end of JK remain in the high group again at the end of SK. The present results suggest that this regression of high achievers towards average could be a consequence of a summer break being detrimental to the high group over the JK-to-SK summer. Therefore, given the detrimental effect of

the summer on the High group that is not made up for by school year trajectory (as compared to the Average group), summer instruction may also be essential for the High group to maintain their lead in reading achievement, if this is deemed important by educators, students, and/or parents. Alternatively, however, it is certainly possible that this "edge" in letter sound decoding fluency does not represent a clinically meaningful difference, as the High group is likely continuing to achieve at an "adequate" level for fluent word decoding. Therefore, the greater decline in their LSF scores may not represent an important loss for these readers. Rather a similar loss in those students who struggle with the mastery of basic reading skills might be very much more meaningful than a drop of this magnitude in the High group.

Additionally, it is important here to acknowledge the thoughts of Protopapas and colleagues (2011) on the use of the same measure for charting changes over time as it relates to the Matthew Effect. Again, we return to the statistical phenomenon of regression toward the mean. Protopapas and colleagues (2011) were able to avoid the confound of using identical criterion and outcome variables, as they felt that doing this could obscure any possible divergence in the groups' trajectories. Unfortunately, within the present work, using the same criterion and outcome variables was an unavoidable design flaw. Therefore, regression toward the mean needs to be acknowledged within the context of school year trajectories as well as summer learning loss. Furthermore, according to Protopapas, Parrila, and Simos (2014), there are problems with scaling when using literacy measures to compare performances in the present manner. These authors explained that the comparison of performance differences requires interval-level data, but only ordinal-level data are available with current reading

measures. The CBM probes, although linear in their scaling, cannot be considered interval-level data because we cannot say about a student who scores, for example, 30 points on the Word Identification Fluency probe that he/she is twice as strong a reader as a student who scores 15 points. Therefore, although these types of comparisons are routinely made throughout the reading literature, this limitation certainly needs to be acknowledged as a limitation within the present work.

Opportunity-cost of summer vacation

Taking a prolonged break from continuous instruction slows down the learning trajectories of students, especially in these early years. Because CBMs were taken at frequent intervals throughout the school year, estimates regarding the consequences of the summer break can be proposed. The present results suggest that low-achieving students could have been further ahead by about 2.5 months on both LSF and WIF. This additional time to continue improving their early reading skills could be important for these students who struggle with learning to read. Surprisingly, however, the opportunity-costs are the highest for high-achieving students over the JK-to-SK summer. They might have been seven months further advanced had there not been a JK-to-SK summer break. However, the summer break is less detrimental to the high group over the SK-to-Grade 1 summer, as their LSF scores are already very high, which may contribute to the improvement in this group's WIF scores. Providing parents and students with resources to teach more effectively and track progress, or enrolling students in summer learning programs (ranging from participation in a summer book club to summer engagement in a well-validated reading intervention) could lessen the apparent impact of the summer break for all students.

Transitions between Achievement Groups

Important results were also identified within the post-hoc analysis. It is clear, whether due to a statistical phenomenon like regression to the mean or due to simply needing extra time to catch up to their peers, that many students in the Low group do manage to enter into the Average group by Grade 1. In general, the dual-discrepancy definition of non-responsiveness described by McMaster and colleagues (2002) did not accurately predict low achievement over time within this sample. Thus, longitudinal data (based on either a dual-discrepancy model or the single-discrepancy model used herein) may be necessary to identify true non-responders. This is in line with the responsiveness-to-intervention approach to defining reading disabilities in which progress is continually monitored and those children who still do not respond to a given number of waves of increasingly intensive instruction qualify for special education services and/or a reading disability diagnosis (Fuchs et al., 2003; Fletcher et al., 2007). The present data highlight the importance of ongoing monitoring of reading scores, at least in the early years of reading instruction, as opposed to a single time-point measurement that will or will not qualify a particular student for special services.

Strengths of CBM tracking

It is believed by this researcher that within the present sample, the CBMs did well at tracking and describing students' changing levels of achievement. For example, the CBMs showed that approximately 50% of students who were in the low achievement group at the end of JK were able to transition to the average group by the end of SK. Also, 75% of students who are in the Low group in JK transition out by the end of Grade 1. High turnover of this Low group may mean that the CBM probes are able to identify

them as low achievers, and the PALS program may then foster growth in reading. These data also suggest that only 25% of the low group is comprised of students with reading disabilities who do not respond to instruction; 4% of the total experimental sample. This finding complements the results of Fuchs, Fuchs, and Compton (2004), who found that approximately this percentage of students are identified as being non-responsive to intervention by the end of Grade 2, and also approximately matches the Learning Disabilities Association of Canada's (2007) report of incidence rates in Canada. By using the CBMs associated with PALS (as Fuchs et al., 2004 did), the present results were able to produce complementary results a full school year earlier, using a larger sample.

Lack of Control Group

In addition to those limitations already discussed, it is important to acknowledge one additional limitation inherent in the present study. All of the students included within the present sample were receiving the PALS intervention; there was no control group of students who were administered the CBM probes outside of the context of PALS. Although this appears to be a considerable strength in terms of the school board's approach to instruction, as PALS is a well-validated reading intervention from which the students are expected to have benefitted, it is nonetheless a limitation of the present investigation that the analysis could not be separated from the context of the PALS intervention by comparing to the achievement of students who are not receiving PALS intervention. Widespread participation in the PALS program may have optimized all students' school year trajectories, and diminished their time to recover from summer learning loss, compared to students who were not engaged in a well-validated reading intervention, making it impossible to generalize the present results to school boards

where participation in a well-validated reading program is not a requirement for all students.

Conclusions

During the early years of reading instruction, within a school board where well-validated intervention is the norm, the summer break appears to have a negative impact on students from all levels of reading achievement. Over the JK-to-SK summer, when improving letter sound fluency is critical, taking a summer break appears to have the largest consequence on the High achievement group. These-high achieving students might be up to seven months further ahead in reading scores, had they gone to school instead of taking a break. The summer break most negatively affects the Low achievement group over the SK-to-Grade 1 summer when improvement in word identification fluency is observed in their average- and high-achieving peers. This lack of growth consistent with their peers may compound reading problems in low-achieving readers, as their reading abilities are already low relative to their peers and their peers are demonstrating progress in this area over the summer. At the beginning of Grade 1 in particular, students whose reading achievement was low prior to the summer may require extra individual attention and remediation. Taken together, the present results point to the utility of summer practice and/or instruction during the summer break, so that students continue to learn to their full potential. The present results also suggest that year-round programming or alternate school calendars (e.g., those with more frequent, shorter breaks spread out throughout the school year) might benefit all students, not just those with identifiable reading problems.

Furthermore, regular CBMs as part of the PALS program were useful in tracking the progress of all three groups of students. All achievement groups recovered quickly from summer learning loss and enhanced their reading ability at a linear rate over the school year. Even low-achieving students were able to rapidly recover from summer learning loss, perhaps because they were easily identified as low achievers by teachers at the beginning of the school year. Approximately half of the low achieving students were even able to transition into the Average group within one school year. Given the appropriate responsiveness of most students to the PALS intervention, the present results also lend support to the choice to implement PALS as a school board-wide intervention strategy to target reading achievement in students, regardless of achievement group status, although this hypothesis would require comparison with a control group and agreement from a measure of concurrent validity to be considered officially "supported."

Future Directions

Future studies should more thoroughly investigate the possibility of Grade 1 reading scores as being a tipping point for potential ongoing reading problems. It would be interesting to develop a study that would bridge the gap between the present results and the results of Menard and Wilson (2013) by looking closely at the early primary school years, either using a CBM strategy or using standardized measures, , or ideally a combination of both strategies so that concurrent validity can be assessed. This study could either be cross-sectional in nature by testing students in Grades 1 to 4, or longitudinal in nature by testing students in Grade 1 and following these same students until they reach Grade 4. Given the emphasis herein on the importance of longitudinal data, a longitudinal approach would be ideal. To go a step further, this hypothetical study

could even include students in the older elementary grades and into high school, including passing or failing the Ontario Secondary School Literacy Test (traditionally written in Grade 10) as an outcome variable. It would be very interesting to chart reading development over the course of students' entire early academic careers in a Canadian context. Although this might be a very ambitious undertaking for any single study, this undertaking could be accomplished via a series of related studies.

CHAPTER IV

Prosody as a Predictor of Responsiveness-to-Intervention in the Peer Assisted Learning Strategies Program

The successful development of literacy skills is one of the most important challenges for school-age students to meet. Reading problems in the early years and onward are shown to have significant short- and long-term outcomes for those affected. According to a study which took a comprehensive look at ten different Statistics Canada datasets that range across the lifespan of Canadians (Learning Disabilities Association of Canada, 2007), poor short-term outcomes for individuals with Learning Disabilities (LD) in reading and in other areas include early school dropout, low educational attainment, and lower overall levels of literacy achievement as compared to students without LD. In the long-term, poor outcomes for adults with LD include a lack of success at finding and keeping employment and related financial problems, and a three-fold increase in reported problems with physical, general, and mental health, including high levels of distress, depression, anxiety disorders, and suicidal thoughts. Given the greater likelihood of experiencing poor life outcomes such as these, the importance of developing a solid academic foundation in reading in is apparent.

Knowing how to read well is an undeniably important skill set to possess, yet reading is also a very complicated process to master. Reading achievement is dependent on a student's level of proficiency in areas such as phonological processing, phonological awareness, reading fluency, vocabulary, short-term memory, and reading comprehension. Chall's (1983) five-step model of the development of reading ability proposes that reading occurs in discrete stages that are qualitatively distinct (Kaplan & Walpole, 2005).

Chall (1983) posits that reading begins with the acquisition of language skills in the pre-reading stage (stage zero). Then, in the beginning stages of true reading development (ages 6 to 7), the child must learn the letter-sound relationships in order to decode printed words. Learning the correspondences between letters and sounds is a process which requires increasing proficiency in phonological processing and phonological awareness. Then in the second stage (ages 7 to 8), decoding fluency is gained through practice. The third stage of reading development (ages 9 to 13) marks the transition from "learning to read" to "reading to learn." The child begins to acquire a store of background information and a growing vocabulary through further practice in reading and by reading a wide variety of materials, and in doing so, the child is acquiring new information, thoughts, and ideas through reading. In stage four (ages 14 to 18), this knowledge is compared and evaluated. Different viewpoints and multiple interpretations of the text can be considered in stage four reading. Finally, stage five reading (ages 18 and up) involves the synthesis of information and the formation of advanced-level hypothetical thinking; reading in this stage is constructive, in that the reader can construct knowledge from the text. Snider and Tarver (1987) specify that each stage is dependent on the mastery of the previous one. Based on Chall's (1983) model of reading development, is it clear that skilled reading depends on adequate decoding, sight word knowledge, and automaticity.

In addition to other component skills, prosody is also at play in the development of students' reading skills. Prosody is a linguistic term which describes the rhythmic and tonal aspects of speech; the "music" of oral language (Hudson, Lane, & Pullen, 2005, p. 704). Prosodic features of speech include variations in pitch (intonation), stress patterns (loudness and syllable prominence), and duration (the length of time spent on each

speech sound; Dowhower, 1991). Prosodic reading also includes "appropriately chunking groups of words into phrases or meaningful units in accordance with the syntactic structure of the text" (Kuhn & Stahl, 2000, p. 5). Taken together, prosodic features are considered suprasegmental because they usually contribute to, or cover, more than one speech sound or segment (e.g., syllables, words, or larger units of speech; Dowhower, 1991). Overall, prosodic features contribute to the reader's level of expressiveness when reading a passage of text, and reading with expressive rhythmic and melodic patterns is called prosodic reading (Dowhower, 1987). Thus, a prosodic reader segments the text into meaningful units marked by appropriate prosodic cues such as pauses, and demonstrates variation in the duration of those pauses, the raising and lowering of pitch, the lengthening of certain vowel sounds, and greater emphasis on certain words (Dowhower, 1991). In order for children to read in a way that is meaningful both to themselves and to their listeners, they must read prosodically, as opposed to reading in a way that is monotonous and lacks flow.

While prosody and fluency are purportedly connected, they remain separate aspects of reading ability. The development of reading fluency appears to be closely related to the ability to read a passage or a list of words quickly, without laborious decoding of individual phonemes. Prosody is the higher-order of the two skills, in that expressiveness rarely occurs without fluency, but fluency occurs without expressiveness (Cowie, Douglas-Cowie, & Wichmann, 2002). The following discussion is not so much concerned with fluency per se, despite its close relation to prosody, but with the development and use of prosody and its relation to the acquisition of skilled reading in school-age children.

According to Dowhower (1991), prosodic features, or "the melodies and rhythms of our language," play a significant role in how children process both written and spoken language (p. 168). From the processing of "motherese" in infancy – a variation on normal speech which is characterized by exaggerated prosody – to learning to read with expression as their reading skills progress, prosody remains an important part of language learning in early to middle childhood. To this end, there have been many studies conducted which link prosodic awareness and use to language processing skills. As one example, Marshall, Harcourt-Brown, Ramus, and van der Lely (2009) found that children with specific language impairments and dyslexia, both of which are disorders characterized by difficulties in the broad area of phonology, have an impaired ability to extract meaning from certain linguistic structures when compared to age-matched controls. This means that while children with specific language impairments and dyslexia perform well on auditory discrimination and imitation tasks, they are less able to use prosodic elements to comprehend, or impose meaning on, spoken language. This finding suggests that children with language or phonological processing impairments experience higher-order impairment in the application of prosody when attempting to extract meaning from spoken language.

Greater sensitivity to prosody in spoken language has also been linked to greater attainment of early reading skills. Holliman, Wood, and Sheehy (2010) looked at sensitivity to speech rhythm and non-speech rhythm and how these two types of sensitivity relate to reading development. Receptive sensitivity to both speech rhythm and non-speech (i.e., musical or metrical) rhythm predicted a significant amount of unique variance in reading attainment that was independent of the contributions to reading

attainment made by variables such as age, vocabulary, phonological awareness, and working memory. The implication of this finding is that reading attainment appears to depend not only on the child's ability to apply prosodic features when reading a passage of text, but also on the more general sensitivity to the use of prosody in processing speech-based and non-speech-based sounds.

Furthermore, studies have been conducted which link oral receptive prosody with the attainment of higher-order reading abilities such as text comprehension, as opposed to simple word reading. For example, Cohen, Douaire, and Elsabbagh (2001) investigated the influence of prosody and its written equivalent, punctuation, in text comprehension. In their first experiment, participants listened to passages in which the prosodic features were classified as normal, monotonous, or altered. In the normal condition, the prosodic features matched participants' expectations, the monotonous condition lacked prosodic cues, and in the altered condition, there was a mismatch between the participants' expectations for pronunciation and the way that the passages were actually articulated. As such, the altered condition presented conflict between prosodic structure and syntactic structure where these structures would normally overlap (as in the normal condition). Similarly, in their second experiment, participants were asked to read passages which contained manipulations in punctuation; appropriate punctuation, no punctuation, or altered punctuation. The results suggested that a similar pattern exists across aural and graphical media, in that altered prosody and punctuation affect performance in a similarly deleterious fashion which seriously impairs listening and text comprehension and subsequent word recognition. This finding suggests that even normally-developed adults

rely heavily on the prosodic features of both spoken language and text in order to bolster comprehension.

Furthermore, Miller and Schwanenflugel (2006) also studied oral reading prosody and its apparent link to reading comprehension. These authors used sophisticated spectrographic analysis; an imaging technique which renders a three-dimensional representation of speech that can be compared across samples. In analyzing Grade 3 students' and adults' use of prosody when reading aloud a syntactically complex passage of text, in addition to standardized measures of oral reading skill, these authors found that children who had quick and accurate oral reading skills had shorter and more adult-like pause structures, larger declines in pitch at the ends of declarative sentences, and larger rises in pitch at the ends of yes/no questions. Also, children who demonstrated more advanced use of prosody also tended to demonstrate better reading comprehension skills, further suggesting that an understanding of and the ability to use the prosodic features of language is an important factor in the ability to extract meaning from text.

In addition to measuring prosodic processing of speech sounds, research has also been conducted to look at whether readers insert prosodic features when reading silently. Because silent reading by definition involves no outward verbalization of the text, researchers who study prosody in silent reading must take a more indirect measurement approach, often measuring variables such as eye movements that occur while reading. These eye movements are used as a proxy for prosody in that they suggest that the reader is inserting pauses and other prosodic features while they read a passage silently. In one such study, Ashby and Clifton (2005) tested Fodor's (1998) implicit prosody hypothesis, which claims that readers impose similar prosodic features as would be observed in

speech when they are reading silently. Ashby and Clifton (2005) asked whether it takes longer to read words which, if spoken aloud, would have two stressed syllables (e.g., maladjusted) than words of similar length that have only one stressed syllable (e.g., significant). These authors found that readers' eyes do indeed pause longer on words that have two syllables on which stress is placed, suggesting that they are inserting the stress while reading silently. The implication of this finding is that prosody is not limited solely to speech or when reading aloud, but is also used in strictly silent reading.

Furthermore, there is evidence to suggest that readers not only insert prosodic features onto what they read, but that they also use these features as an aid to comprehension. Breen and Clifton (2011) conducted a study which investigated whether readers would become confused and therefore slowed down in their reading when their expectations for how the prosodic features of the text should occur did not align with the actual text that they were assigned to read. Specifically, readers were assigned to read limericks, which have a prescribed metrical structure. These limericks included alternating-stress homophones such as "CONvict" or "conVICT," the pronunciation of which is ambiguous in the absence of context. The stress of these words, presented in the context of the limericks, either matched or mismatched the metrical pattern that is associated with limerick poems. The results of Breen and Clifton's (2011) investigation demonstrated that readers experience a reading cost, in that they need to pause longer to comprehend what they have read when their prosody-based expectations do not match reality. This finding further supports the hypothesis that there is a close link between prosody and comprehension.

One of the major problems in studying prosody and reading ability is in determining the direction of causality between prosody and reading comprehension (Dowhower, 1991; Kuhn & Stahl, 2000; Hudson et al., 2005). Dowhower (1991) describes the relationship between prosody and comprehension as a "chicken-and-the-egg dilemma" (p.170) and specifies that it is unclear which one comes first or if one is necessarily an indicator of the other. There appears to be a reciprocal relationship between prosody and reading comprehension (Hudson et al., 2005) in which readers who score higher on measures of comprehension also tend to read more expressively, and students who read with expression tend to better comprehend what they have read. As yet, there are no known studies which have attempted to directly investigate the direction of causality between comprehension and prosodic reading. Although the proposed study will not seek to directly address this issue, it will seek to more clearly elucidate the relationship between prosody and early reading achievement.

Prosodic reading ability was previously used as a mediator of reading comprehension skill by Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl (2004). These authors treated prosody as a potential partial mediator of the relationship between decoding speed and comprehension skill, with all variables measured at the same time point. Decoding speed was measured using the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), reading comprehension was measured using same-named subtest from the Wechsler Individual Achievement Test (1992), and expressive oral reading prosody was measured using computerized spectrographic analyses of prosodic elements such as intra- and inter-sentential pauses and variations in pitch across and within words. According to Schwanenflugel and colleagues (2004), the

rationale behind using prosody as a mediator of reading comprehension is that as a child acquires better decoding skills, attentional resources are freed up since they are no longer being used for the onerous task of decoding individual words. These resources become available to be allocated to higher-order functions of reading such as comprehension. Prosody is hypothesized as a feedback process by which the child can bolster his or her comprehension of the text. In other words, prosody is proposed as a mechanism by which decoding ability can develop into reading comprehension. Schwanenflugel et al.'s (2004) results suggest that there is only minimal support for the role of prosody as a mediator of reading comprehension, in that while more fluent reading does appear to free up resources that lead to more advanced use of prosodic reading, prosody does not appear to be acting as a scaffolding or feedback process for comprehension.

As with any study, however, there were limitations to Schwanenflugel and colleagues' (2004) results. For example, the authors address the fact that their choice of a measure of reading comprehension may not have been satisfactory, in that the reading comprehension subtest of the Wechsler Individual Achievement Test (a) may measure comprehension too globally, (b) may not be adequately challenging to allow for the sensitivity to find a relationship between prosody and comprehension, and (c) did not use the same passages that were used for the measure of prosodic reading, so direct comparisons may not have been ideal. Additionally, the fact that Schwanenflugel and colleagues (2004) collected measures of both decoding skill and reading comprehension at the same time point does not allow for an analysis of whether prosodic processing ability might affect students' reading trajectories over the school year.

The goal of the current study is to investigate prosodic processing within the context of the Peer Assisted Learning Strategies (PALS) program developed by Fuchs, Fuchs, Mathes, and Simmons (1997). This reading intervention focuses on decentering reading instruction in classrooms by pairing peers with each other for intensive but efficient reading practice. High levels of effectiveness, as well as a high degree of both teacher and student satisfaction with the program, have been demonstrated (Mathes, Howard, Allen, & Fuchs, 1998), and thus, the PALS intervention is widely used in school settings. Within the PALS framework, the measurement of reading skills will be accomplished using the curriculum-based measurements (CBMs) which are associated with the PALS program. These CBMs are meant to track student progress regularly throughout the duration of the intervention. CBMs have been shown to provide reliable and valid information about students' progress (Deno, Fuchs, Marston, & Shin, 2001) and it is expected that they will be more sensitive to change than standardized measures, as the item gradient is not as steep due to the fact that there are generally more items per test and the raw scores are not converted into standard scores. Another advantage is that the proposed study made use of archival data that measure early reading skills prior to the point when the students entered the PALS intervention and also measures their reading outcomes following participation in PALS. This is a notable improvement on Schwanenflugel and colleagues (2004) since it adds a longitudinal component to the study of prosody and reading achievement. In sum, it was hypothesized that a better understanding and use of prosody (receptive and expressive prosody, respectively), would predict a greater slope in reading scores over the school year.

Method

Participants

The data for the current study were a mix of archival and newly collected information. The archival portion of the data consisted of CBM reports from the Windsor-Essex Catholic District School Board (WECDSB). These reports provided the baseline and outcome assessments of early reading ability, and the difference between start- and end-of-year scores were used as a proxy for reading achievement over the school year. All of the children who participated in the present study were fluent English speakers.

The participants consisted of 104 students in Grades 1 to 3. The average age of the sample was 91.86 months, which corresponds to an age of approximately 7 years, 8 months (range: 72 to 111 months). In terms of the grade distribution, 37 of the students were in Grade 1, 28 in Grade 2, and 39 in Grade 3 during the year of testing. These participants attended four different schools within the WECDSB. Data collection at the first school occurred during the months of November to March ($n_{school1} = 39$), data from students attending the second and third schools were collected concurrently during the months of March to May ($n_{school2} = 22$; $n_{school3} = 30$), and data were collected from students at the fourth school during the month of June ($n_{school4} = 13$). The gender breakdown of the sample was even; 52 female and 52 male participants. Information regarding the race/ethnicity of the participants came from parental report. According to this report, 65.4% of the sample was identified as being non-Hispanic White/European descent (n = 68), 6.7% as other/mixed (n = 7), 1.9% as Hispanic/Latino (n = 2) and non-Hispanic Black/African descent (n = 2), 1.0% as Aboriginal (n = 1) and Asian or Asian

descent (non-Arab) (n = 1), while 22.1% of the parent sample did not answer the question regarding their child's race (n = 23 left blank). In terms of parental education, 100% of the sample had at least one parent who had completed at least a high school education (3 participants, or 2.9% of the sample, had only one parent who had completed high school). With respect to higher education, 43.3% of the sample had at least one parent who had completed at least some college/trade school, while 58.6% of the sample had at least one parent who had completed at least some university as their highest level of education. Finally, it should be noted that 28 (26.9%) of the participants were siblings of other children in the sample. This will be acknowledged within the data analysis.

Materials and Procedure

Curriculum-based measurements (CBMs). The measures which were used to measure the students' learning trajectories within the proposed study are the CBMs associated with the Peer-Assisted Learning Strategies (PALS) program created by Fuchs et al. (1997). The PALS program measures achievement in four reading-related domains: Letter-Sound Fluency (LSF), Word Identification Fluency (WIF), Passage Reading Fluency (PRF), and Maze Fluency (MF).

Letter Sound Fluency. The LSF test assesses students' speed and accuracy in identifying letter sounds. This measure was used in the earlier grades (Junior Kindergarten, Senior Kindergarten, and Grade 1) within the WECDSB sample. It was administered eight times at regular intervals throughout the school year. In completing this task, students are presented with a page of 26 letters and then given one minute to pronounce as many letter sounds as they are able to pronounce. This measure is scored by counting the number of correct responses to arrive at a total raw score. If the student

reads all or part of the list in under one minute, a correction formula is applied as follows: (number of letters read correctly/number of seconds it took to read list) x 60 = estimated number of letters read correctly in one minute. In the end, the test measures the number of letters that students are able to read in one minute, regardless of whether they can decode all of the letters on the page.

Word Identification Fluency. The WIF test assesses the students' speed and accuracy in identifying whole words, and it is also intended for early readers. Within the present sample, it was administered to students in SK and Grade 1. Within the SK sample, WIF was administered twice (January and June measurements), and within the Grade 1 sample, it was administered seven times at regular intervals throughout the school year. In completing the WIF measure, the student is presented with a list of 100 words and he or she is given one minute in which to read as many words as possible from the list. This task is scored by assigning a value of '1' to every word correctly read and '0' to those not attempted or read incorrectly. If the students finish reading the list in under one minute, the same correction formula as described above is applied.

Passage Reading Fluency. The PRF test assesses students' fluency in reading grade-level text passages. In completing this task, students read a grade-appropriate passage for 1 minute. Performance on this task is measured by counting the number of words attempted within the time limit and subtracting the number of words that were omitted and the number of words that were read incorrectly to arrive at a total score which represents the total number of words read successfully within the time limit. Within the present sample, this measure was administered to students in Grade 1, Grade 2, and Grade 3.

MAZE Fluency. Finally, the MAZE Fluency test assesses reading comprehension by having students read a passage and circle the correct word from a group of choices for each blank. Students are given 2.5 minutes to complete this task, and they receive one point for each correct response. Scoring is discontinued if three consecutive errors are made.

Profiling Elements of Prosodic Systems – Child Version (PEPS-C). According to its creators, Peppé and McCann (2003), the PEPS-C is designed to assess children's ability to understand and express prosody. All descriptions of the test come from Peppé and McCann's (2003) outline of the test. The PEPS-C is suitable for all ages above 4 years. The test is administered using a computer and elicits responses through auditory stimuli and pictures presented on the screen. A pre-test vocabulary check is completed to make sure that the tester and the participant agree on the terms to describe the stimuli that are seen on the computer. For the present study, the version of the PEPS-C that is recorded in a "North American" accent was used.

The PEPS-C is comprised of twelve tasks which address both receptive and expressive skills. The tasks fall across two levels, examining both prosodic function and prosodic form. The first-level (prosodic function) PEPS-C tasks cover four main linguistic functions that are conveyed by prosody, and include a receptive and an expressive task for each (eight subtests in total). These four prosodic functions include: (1) "turn end type" – indicating whether an utterance requires an answer or not (question versus statement), (2) "affect" – indicating mood, emotions, and/or opinions (specifically, signaling liking or reservation with respect to food items), (3) "chunking" – indicating the chunking of prosodic phrase boundaries into meaningful units (e.g., the difference

between ‘fruit, salad, and milk’ and ‘fruit-salad and milk’), and (4) “contrastive stress or focus” – indicating emphasis on one word in an utterance to focus attention on it (e.g., ‘white COW’ as opposed to ‘WHITE cow’). As mentioned, the PEPS-C also measures prosodic form. There are two auditory discrimination tasks which measure receptive prosodic form and two imitation tasks which measure expressive prosodic form (these tasks make up the remaining four out of twelve subtests). To successfully complete the auditory discrimination tasks, the participant has to indicate whether the stimuli presented have the same or different meanings. The stimuli for these two tasks exemplify the prosodic variations that convey the different meanings within the receptive prosodic function tasks, described above. To successfully complete the two imitation tasks, the participant must be able to produce, as a whole, the types of prosodic variations needed for completing the expressive function tasks, also described above.

In terms of the actual administration of the tasks, all six of the receptive tasks present as an auditory stimulus with two pictures as response options on the computer screen; the participant points to or clicks on the half of the screen which represents his or her response. For the expressive function tasks (four, in total) pictures appear on the screen and the participant is required to say aloud what he or she sees. Their utterances are scored by the tester on a separate keypad. For the expressive form, or imitation, tasks (two, in total) the tester similarly evaluates the participants’ responses as Good, Fair, or Poor.

In all, the twelve subtests of the PEPS-C take approximately 40 minutes to administer. They were administered in the order in which they were described above. As mentioned, prior to the administration of the subtests, a vocabulary check was completed

to ensure that the participants were familiar with the objects depicted in the tasks. The administration also included a number of practice items prior to the test items (which do not contribute to the participants' final scores), in order to familiarize participants with the tasks.

Theoretical Model of Change

Collins (2006) outlined a number of elements that must be considered when articulating a theoretical model of change. With respect to the general or characteristic shape of change, it is expected that reading skills on all measures (LSF, WIF, PRF, and MAZE Fluency) will increase linearly over the school year. Although reading skills can develop in a more discrete, stage-like manner that is characterized by jumps in development overall, it is thought that each individual skill (e.g., decoding ability) develops continuously over time, increasing in a linear fashion. It is not expected that there will be periodicity or a cyclical nature associated with this change. The change in reading achievement over the school year is thought to be due to calendar time as well as the reading instruction that is provided over that calendar time to foster learning. The most relevant covariates that may predict changes aside from calendar time and the PALS intervention itself are socio-economic status and perhaps reading encouragement within the home environment, outside of the PALS program. As well, attention factors and motivation to achieve may play a role. It is possible that some of these covariates may change along with the calendar time associated with changes in achievement, but they are expected to remain relatively stable over time. Socio-economic status is expected to remain especially consistent over time, and so is encouragement to read within the home unless an intervention is put in place that will increase this support for students. Attention

and motivation factors can wax and wane over time, but again, are expected to remain relatively consistent for each student. The process of learning to read, within each skill set (e.g., decoding fluency development), is expected to be fairly continuous, although it could certainly be impacted by major life events (e.g., an acquired brain injury, a death in the family, etc.). Finally, it is expected that there will be meaningful inter-individual variability in change, as each student presents with his or her own individual reading trajectory.

Data Analysis

According to Little's MCAR test, missing data were found to be missing completely at random for both the Grade 1 sample, χ^2 (36) = 31.559, p = .680, and the Grade 2/3 sample, χ^2 (19) = 19.842, p = .404. Two methods of dealing with missing data were compared. First, for any case where a subtest score was missing, the mean of the relevant expressive or receptive subtest scores was substituted for the missing score. For example, Participant #17 was missing a score on Chunking Output. Therefore, the mean of the other five output variable scores was substituted for the missing Chunking Output score. Such substitutions were required for seven participants within the sample. This was thought to be the most intuitive solution to the missing data problem, as this method used existing data from each participant's own related scores to generate a total score on either the receptive or expressive prosody scale. In other words, existing data were used to make a prediction at the missing score, and then this predicted mean value was substituted as if it were an actual obtained value. Stevens (2009) suggests that mean substitution is usually a viable option for dealing with missing data. Alternatively, an estimation-maximization strategy was also employed using SPSS software. The

differences in the individual values of all prosody variables were so negligibly small that it was thought that both methods of dealing with missing data were equally effective and valid. Therefore, all regression values reported below come from the original mean substitution method of dealing with missing data. No differences in trends were observed when the regression analyses were run using the EM method. The reported results reflect the data as imputed using the mean substitution method.

Following imputation of missing data values, the prosody variables were aggregated into composite variables; the total scores on all six of the input variables and all six of the output variables were added together to arrive at composite receptive and expressive prosody variables, respectively. The reliability of the PEPS-C data was tested using Cronbach's alpha coefficients. For the Grade 1 Receptive PEPS-C data, Cronbach's $\alpha = .70$, and for the Grade 1 Expressive PEPS-C data, Cronbach's $\alpha = .75$. For the Grades 2/3 Receptive PEPS-C data, Cronbach's $\alpha = .46$ and for the Grades 2/3 Expressive PEPS-C data, Cronbach's $\alpha = .67$. Correlation matrices for the subscales of the PEPS-C composites can be found in Tables 8 through 11.

Beginning-of-year scores on the PALS program variables were established by taking the average of two reported baseline scores. End-of-year scores were established similarly, by taking the average of two post-intervention scores (where available; if there was only one post-intervention score available, then this score was used). Progress on the PALS program variables was established by subtracting the beginning-of-year score from the end-of-year score. This difference score was taken as a measure of how much improvement each participant made over the school year (positive scores indicate

improvement, negative scores indicate regression, and a progress score of 0 would indicate no change).

The predictive power of prosody on reading progress was tested using a regression analysis. Given that different CBMs are used to assess reading skill in Grade 1 students, as compared to those used for Grades 2 and 3 students, the data were split thusly, as mentioned above. Three regression analyses were run for the Grade 1 data (as three separate CBMs are administered in Grade 1), and two regression analyses were performed for the Grade 2/3 data. It should be noted that running multiple analyses may have contributed to an increase in the probability of making a Type I error by increasing family-wise error. The assumptions of multiple regression analysis were examined before the outcome-specific output was considered. There are three assumptions that are not analysis-specific. First, all of the dependent variables were measured on a continuous scale. Second, all regression analyses included at least two independent variables. Furthermore, in an attempt to control for time of administration effects, two dummy codes were created to differentiate between that data which were collected from School 1, Schools 2 and 3 (as these data were collected concurrently), and School 4. Aside from concern about time of administration differing between schools, school effects are assumed to be of no great concern, as all of the schools in which testing was conducted were chosen due to their high adherence to the PALS method. In other words, it is assumed that all of the four schools included in the analysis were approximately equal in their adherence to the PALS method, thereby preserving the integrity, or the "sameness" of the PALS data collected from each school. For all regression analyses, the two dummy codes were entered into Block 1, and the receptive and expressive prosody composites

were entered into Block 2. The regressions were run using the Enter method, as there was no theoretical reason to believe that receptive prosody would be a stronger predictor than expressive prosody, or vice versa.

Results

Means and standard deviations for PEPS-C and progress on the CBM variables are presented in Tables 12 to 16.

Grade 1 Analyses

The two IVs were first assessed for multicollinearity within the Grade 1 sample by entering them both into a regression and predicting a random variable. The variance inflation factor value was 1.909, so the assumption of lack of multicollinearity was not violated, as Tabachnick and Fidell (2012) recommend that variance inflation factor values greater than 10 be treated as problematic.

Grade 1 LSF analysis. The independence of observations assumption was tested using the Durbin-Watson statistic. A value of 1.94 indicates no serial correlation between the residuals. However, given that the data are nested within schools and classrooms, and the sample did include some siblings, it is thought that this assumption has been violated regardless of a Durbin-Watson statistic within normal limits, thereby making the standard errors slightly smaller than they should be, and contributing to an increase in the likelihood of Type I error. To assess for linearity of the relationships between the predictors and the criterion variable, scatterplots were examined. It appeared that both receptive and expressive prosody showed no relationship with the LSF data. Furthermore, in assessing for homoscedasticity (variances along the line of best fit remaining similar as you move across the line), a scatterplot of the residual by the predicted values was

examined. The data were evenly spread across the fit line, so this assumption was not violated. The data were also inspected for outliers, high leverage points, and influential points. There were no z-values greater than 2.5 standard deviations above or below the mean, and therefore no significant outliers. Likewise, the cutoff for high leverage points [(2k+2)/n, or in this case ((2*4)+2)/37 = .270] revealed no problematic points of data. Also, all Cook's Distances were well below 1 (highest Cook's value = .2), so these data also contain no influential points. Finally, the Shapiro-Wilk test revealed that the residuals were normally distributed (p = .586).

As mentioned above, the regression was run by first entering the two dummy codes for administration time into Block 1. The overall sample size for the Grade 1 LSF analysis was 37 participants. This model was not significant, which suggests that time of administration of the PEPS-C did not affect the outcome, F (2, 34) = 1.705, p = .197. The two prosody variables were entered into Block 2. Adding these two variables did not contribute to a significant increase in the predictive power of the model (ΔR^2 = .014, p = .779). Therefore, the model which included the prosody variables also was not significant, F (4, 32) = .941, p = .453, R^2 = .105, Cohen's f^2 = .117, nor were receptive or expressive prosody significant predictors of change in LSF scores (receptive: β = -.020, b = -.047, t = -.077, p = .939, expressive: β = -.115, b = -.220, t = -.450, p = .656). Overall, prosody scores were not good predictors of scores on the LSF measurement in the Grade 1 sample.

Grade 1 WIF analysis. The independence of observations assumption was again tested using the Durbin-Watson statistic (Durbin-Watson = 1.821). Again, given that the data are nested within schools and classrooms, and the sample did include some siblings,

it is thought that this assumption has been violated regardless of a Durbin-Watson statistic within normal limits. With respect to linearity, scatterplots again suggested that there was no relationship between both receptive and expressive prosody and the WIF data. In terms of homoscedasticity, a scatterplot of the residual by the predicted values revealed that the data were relatively evenly spread across the fit line, although there were two points that were more toward the higher side of the x-axis, while the rest of the points were clustered closer to the lower side of the x-axis (the predicted axis). As above, there were no outliers, high leverage points, or influential points identified within the data. Finally, the Shapiro-Wilk test revealed that the residuals were normally distributed (p = .357).

This regression was also run by first entering the two dummy codes for administration time into Block 1. The overall sample size for the Grade 1 WIF analysis was also 37 participants. This model was not significant, which suggests that time of administration of the PEPS-C did not affect the outcome, F (2, 34) = .347, p = .710. The two prosody variables were entered into Block 2. Adding these two variables did not contribute to a significant increase in the predictive power of the model (ΔR^2 = .088, p = .222). Therefore, the model which included the prosody variables also was not significant, F (4, 32) = .968, p = .439, R^2 = .108, Cohen's f^2 = .121, nor were receptive or expressive prosody significant predictors of change in WIF scores (receptive: β = .432, b = .810, t = 1.686, p = .102, expressive: β = -.199, b = -.302, t = -.781, p = .440). Overall, neither were prosody scores were good predictors of scores on the WIF measurement in the Grade 1 sample.

Grade 1 PRF analysis. The independence of observations assumption was again tested using the Durbin-Watson statistic (Durbin-Watson = 2.101). Again, given that the data are nested within schools and classrooms, and the sample did include some siblings, it is thought that this assumption has been violated regardless of a Durbin-Watson statistic within normal limits. With respect to linearity, scatterplots again suggested that there was no relationship between both receptive and expressive prosody and the PRF data. In terms of homoscedasticity, a scatterplot of the residual by the predicted values revealed that the data were relatively evenly spread across the fit line, although there was one point that appeared quite high on the residual y-axis. There was one outlier within the PRF data, which corresponded to a student who gained 97 points on this measurement (z = 3.202). However, this datum was not removed, as this was thought to correspond to a real improvement of 97 points over the school year. It should be noted, however, that within the Grade 1 sample, administration of the PRF measure did not begin until January, indicating progress only from January to June, rather than the duration of the school year. There were no high leverage points or influential points within the data. The Shapiro-Wilk test revealed that the residuals were not normally distributed (p = .035), but a secondary test suggested that the residuals were normally distributed (Kolmogorov-Smirnoff: p = .200). It is likely that the normality of the residuals in this case is in the borderline range.

This regression was also run by first entering the two dummy codes for administration time into Block 1. The overall sample size for the Grade 1 PRF analysis was 34 participants as a result of missing data on the PRF measure for three of the participants that made it impossible to calculate a school-year progress score. This model

which included only the dummy codes was not significant, which suggests that time of administration of the PEPS-C did not affect the outcome, F (2, 31) = 2.436, p = .104. The two prosody variables were entered into Block 2. Adding these two variables did not contribute to a significant increase in the predictive power of the model (ΔR^2 = .096, p = .182). Therefore, the model which included the prosody variables also was not significant, F (4, 29) = 2.187, p = .095, R^2 = .232, Cohen's f^2 = .302, nor were receptive or expressive prosody significant predictors of change in PRF scores (receptive: β = .183, b = .328, t = .728, p = .472, expressive: β = .181, b = .252, t = 712, p = .482). Overall, neither were prosody scores were good predictors of scores on the PRF measurement in the Grade 1 sample.

Grade 2/3 Analyses

The two IVs were first assessed for multicollinearity within the Grade 2/3 sample by entering them both into a regression and predicting a random variable. The variance inflation factor value was 1.284, so the assumption of lack of multicollinearity was not violated, as Tabachnick and Fidell (2012) recommend that variance inflation factor values greater than 10 be treated as problematic.

Grade 2/3 PRF analysis. The independence of observations assumption was tested using the Durbin-Watson statistic. A value of 2.128 indicates no serial correlation between the residuals. However, given that the data are nested within schools and classrooms, and the sample did include some siblings, it is thought that this assumption has been violated regardless of a Durbin-Watson statistic within normal limits. In assessing for linearity of the relationships between the predictors and the criterion variable, scatterplots again suggested that both receptive and expressive prosody showed

no relationship with the PRF data. Furthermore, in assessing for homoscedasticity, a scatterplot of the residual by the predicted values revealed that the data were relatively evenly spread across the fit line, with most of the scatter falling near the left side of the x-axis, and six points falling closer to the right (more extreme) side. The data were also inspected for outliers, high leverage points, and influential points. There was one z-score that fell outside of normal limits (z = 2.66), but this was thought to represent a true gain in PRF score over the school year, so it was not removed, although it should be acknowledged that this score may have altered the influence of the analysis. There were no high leverage values or influential points. Finally, the Shapiro-Wilk test revealed that the residuals were normally distributed (p = .188).

As mentioned above, the regression was run by first entering the two dummy codes for administration time, as well as the dummy code which differentiated Grade 2 from Grade 3 students, into Block 1. The overall sample size for the Grade 2/3 PRF analysis was 63 participants. This model was not significant, which suggests that time of administration of the PEPS-C and being in Grade 2 vs. Grade 3 did not affect the outcome, F (3, 59) = .605, p = .614. The two prosody variables were entered into Block 2. Adding these two variables did not contribute to a significant increase in the predictive power of the model (ΔR^2 = .005, p = .870). Therefore, the model which included the prosody variables also was not significant, F (5, 57) = .409, p = .481, R^2 = .035, Cohen's f^2 = .0362 nor were receptive or expressive prosody significant predictors of change in LSF scores (receptive: β = -.069, b = -.163, t = -.439, p = .662, expressive: β = -.014, b = -.026, t = -.093, p = .926). Overall, prosody scores were also not good predictors of scores on the PRF measurement in the Grade 2/3 sample.

Grade 2/3 MAZE analysis. The independence of observations assumption was tested using the Durbin-Watson statistic. A value of 2.069 indicates no serial correlation between the residuals. However, given that the data are nested within schools and classrooms, and the sample did include some siblings, it is thought that this assumption has been violated regardless of a Durbin-Watson statistic within normal limits. In assessing for linearity of the relationships between the predictors and the criterion variable, scatterplots again suggested that both receptive and expressive prosody showed no relationship with the MAZE data. Furthermore, in assessing for homoscedasticity, a scatterplot of the residual by the predicted values revealed that the data were relatively evenly spread across the fit line, meaning that this assumption was not violated. The data were also inspected for outliers, high leverage points, and influential points. There was one z-score that fell outside of normal limits (z = 3.35), but this was thought to represent a true gain in MAZE score over the school year, so it was not removed, although it should be acknowledged that this score may have affected the outcome of the analysis. There were no high leverage values or influential points. Finally, the Shapiro-Wilk test revealed that the residuals were normally distributed (p = .233). As with the Grade 1 LSF measure, the MAZE measure was only administered within this sample in January and June, so progress in MAZE fluency scores between those months does not represent progress over the entire school year.

The model which included the two dummy codes for administration time, as well as the dummy code which differentiated Grade 2 from Grade 3 students, was significant. This suggests that time of administration of the PEPS-C and being in Grade 2 vs. Grade 3 did affect the outcome, F (3, 58) = 10.608, p < .001. The overall sample size for the

Grade 2/3 MAZE analysis was 62 participants as a result of missing data on the MAZE measure for one of the participants that made it impossible to calculate a school-year progress score. The two prosody variables were entered into Block 2. Adding these two variables did not contribute to a significant increase in the predictive power of the model ($\Delta R^2 = .020$, $p = .415$), but the model which included the prosody variables was also significant, $F (5, 56) = 6.669$, $p < .001$, $R^2 = .374$, Cohen's $f^2 = .597$. However, it is clear that this significant model was due mostly to variance explained by the dummy coded control variables, as receptive and expressive prosody were again not significant predictors of change in MAZE scores (receptive: $\beta = -.036$, $b = -.020$, $t = -.281$, $p = .780$, expressive: $\beta = .165$, $b = .070$, $t = 1.313$, $p = .195$). Prosody scores were also not good predictors of scores on the MAZE measurement in the Grade 2/3 sample.

Discussion

It was expected that more advanced understanding and use of prosody in early readers would be related to a greater improvements on the PALS reading measure across one school year. Although it would not have been possible to determine the direction of causality based on the present data (given that students are not randomly assigned to receive a prosody-based intervention or not), the finding that prosody would have predictive power with respect to reading trajectory was expected. However, this hypothesis was not supported; neither higher expressive nor receptive prosody were found to predict greater improvements in reading scores over a single school year.

It could be that within the present sample, prosodic processing is truly not related to reading achievement. This would mean that prosody and reading ability are two independent constructs that are unrelated to each other in terms of predictive power. It is

possible that prosodic processing and reading achievement both develop alongside each other while maintaining functional independence. Perhaps language variables are indeed not good predictors of early reading achievement. The possibility that prosodic processing and reading fluency and comprehension are independent constructs that may simply develop in parallel with each other cannot be ruled out. In the context of the present student, however, it is important to address the other factors that might have impacted upon the ability to identify the potential predictive power of prosody with respect to reading scores.

One possibility for the lack of significant predictive power of the prosody measure in reading achievement is that the broad domain of *language* prosody was measured, as opposed to specific *reading* prosody. When Schwanenflugel and colleagues (2004) designed their investigation of the relation between prosody and reading skill, they measured prosody using spectrographic analysis of students' verbal output while they were actually reading, getting much more directly at their ability to read prosodically. It is possible that in early readers, language prosody is indeed not related to reading achievement, which is an interesting finding in itself. If language prosody is not related to reading achievement, then it would not be a good point of intervention for remediating reading problems. This finding, however, does not rule out the possibility that interventions that target specific reading prosody would not be beneficial to struggling readers. The addition of the longitudinal component (although problematic in its own right) to the measurement of reading achievement is considered to be an improvement on the design used by Schwanenflugel and colleagues (2004), but the change from reading to language prosody may have proven to be a significant downfall

of the present study in terms of finding significant effects, should they exist. As stated, however, the finding that language prosody may not be related to reading achievement is an interesting one in its own right.

On the other hand, as alluded to above, it is possible that the longitudinal nature of the present analysis might have added a significant challenge to finding the expected effects. One of the key challenges in conducting longitudinal research relates to the timing and the number of measurements, as well as the sensitivity of those measurements. With respect to the prosody measure, the PEPS-C is a validated measure of prosodic processing (Peppé & McCann, 2003), but it was only possible to administer it once to each student over the school year. Attempts were made to control for differences due to time of administration, but this did not make up for the fact that the prosody measure captured students' performances at only one time point during the school year, while the reading progress scores represent progress made over the entire school year.

Also, it is possible that the PEPS-C was not adequately sensitive to subtle differences in prosodic processing in the demographic group sampled within the present study. Peppé and McCann (2003) described the PEPS-C as having been "normed on 120 [Southern British English] children (aged 5–14) in 1995–7" (p. 346). Although raw scores were used within the present analysis (i.e., the scores were not norm-referenced), it is possible that the tasks may be less appropriate for the age group of the present sample, in that they may be either too easy or too difficult. Also, the test was originally developed in the United Kingdom, and therefore some of the terms associated with the stimuli are less familiar to North American children (e.g., "cream buns"). The present study did make use of the available North American accent option, and the task which teaches the

students about each of the stimuli was administered, but it is possible that regional differences in familiarity with the stimuli may have affected the sensitivity of the PEPS-C measure within the present sample.

Another possibility is that, although they did allow for the calculation of reading trajectories, the earlier CBMs (namely, LSF and WIF) may represent too rudimentary an assessment of reading skill, at a time in reading development when understanding and use of prosody is not yet crucial. Specifically, there may not be reason to believe that reading (or in this case, speaking and listening) with expression would be related to higher scores on a test where one is asked to read a list of words or letter sounds. Alternatively, as discussed by Dowhower (1991), expressiveness becomes a more important element when fluency and passage comprehension come into play. This hypothesis is partially supported by the fact that prosody as a predictor of reading skills within the present study was trending on significance only on the higher-order measures that require reading fluency and comprehension. Unfortunately as well, the higher-order measures were administered least often (only in January and June for MAZE Fluency and Grade 1 PRF). Had full-year trajectories been available for these higher-order measures, it is expected that a statistically significant result would have been more likely. Again, however, it is possible that the understanding and use of language prosody does not affect the very earliest of reading skills, and even the higher-order CBMs may have been too rudimentary for prosody, at that stage of reading, to be a contributor to higher scores. The hypothesized link between language prosody and reading achievement may not be observable until the later years of reading instruction, after the basic skills have been

mastered, and prosody may only contribute to longer-term progress in reading than was available for the higher-order measures within the present analysis.

Another limitation inherent in the present study is the possibility of regression toward the mean. In describing their methodology, Protopapas and colleagues (2011) discussed how they planned to avoid the confound of using identical criterion and outcome variables, as they felt that doing this could obscure any possible divergence in reading trajectories. Unfortunately, within the present work, using the same criterion and outcome variables was an unavoidable design flaw. Therefore, regression toward the mean needs to be acknowledged within the context of tracking achievement over the school year.

Another possible explanation for the lack of significant results is that the present study may have lacked sufficient statistical power to find an effect, should it exist. The statistical power of the main regression F-test analyses can be calculated based on the observed sample sizes, the set α level, and the observed effect sizes. Based on these values, the statistical power for all regression analyses herein was found to range between .0538 (for Grade 2/3 PRF) and .967 (for Grade 2/3 MAZE Fluency, although as mentioned, the large observed effect size is obviously attributable to the significance of the dummy codes for this last analysis). All of the other values for observed statistical power much more closely resembled the lower end estimate, and were well below the statistical power that was desired for finding the expected effect, if present. It is possible that multi-level modelling might have contributed to a greater likelihood of finding an effect, but unfortunately, the sample size was not adequately large for that type of analysis.

Conclusions and Future Directions

Overall, it was found that receptive and expressive language prosody were not related to reading scores within the present samples. There are, however, a number of possible explanations for this lack of significant findings, as discussed. The results of the present study do not point to prosodic processing as being an effective point of intervention for remediating early reading problems, but neither do these results conclusively rule out the possibility that more advanced readers, or students whose *reading* prosody is impaired, may benefit from this type of intervention. There are two avenues for future research that arise from these results. Future studies could investigate: (1) the predictive power of reading prosody (as opposed to language prosody) in predicting early reading trajectories, and (2) the possibility that impaired understanding and use of language prosody may not be a crucial deficit in learning the very early reading skills, but instead that this variable may indeed have an effect on higher-order skills such as fluency and comprehension.

CHAPTER V

General Discussion

The ability to read is a critically important life skill. Failure to acquire basic reading skills can have a far-ranging negative impact on many aspects of an individual's success. From the Canadian perspective (Learning Disabilities Association of Canada, 2007), poor short-term outcomes for individuals with reading disabilities and other learning disabilities include early school dropout, low educational attainment, and lower overall levels of literacy achievement as compared to typically-achieving students. In the longer-term, poor outcomes for adults with learning disabilities include a lack of success at finding and keeping employment and related financial problems, and a three-fold increase in reported problems with physical, general, and mental health, including high levels of distress, depression, anxiety disorders, and suicidal thoughts. Also as mentioned previously, given the greater likelihood of experiencing poor life outcomes such as these, the critical importance of developing a solid foundation in reading is apparent. The question remains then, how can we improve upon the chances of success for those students who struggle in reading, and who are at risk for illiteracy? This document serves to shed some light on some possible areas of intervention and of importance for students at risk for reading failure.

Definition and Assessment of Reading Disabilities within the WECDSB

This work has especially important implications for the way in which reading disabilities are defined and assessed. Although hesitation is indicated based on methodological limitations, the present results generally support the importance of the responsiveness-to-intervention model, with a caveat. In considering that it appears as

though having access to a lengthy psychological assessment report may not contribute to teachers' ability to improve students' early reading scores, these assessments may not be worth the cost of implementation until the later years of schooling (at least beyond Grade 3). This is in line with the thought process of Hale and colleagues (2006), who purported that a three-tiered RTI model is ideal. In this model, Tier One involves standard instruction for all students. If a student is not responsive to intervention in Tier One, he or she moves on to Tier Two, in which an individual problem-solving approach to intervention is taken. In Tier Two, the teacher and support staff develop individualized intervention strategies, as well as a monitoring process for tracking the student's progress in Tier Two. It is hoped that Tier Two intervention will prove successful at remediating reading problems, but if the student continues to fail to respond to intervention, then he or she moves on to Tier Three, in which access to a psychological assessment and resultant report is granted to further investigate the specific areas of psychological functioning in which the student is manifesting deficits. The WECDSB is currently operating under a variant of the three-tiered model, in that there is no "standard" instruction at the Tier One level. Rather, a well-validated, standardized reading intervention is provided to all students, regardless of their reading ability. This is a real strength of the reading instruction within this school board. At Tier Two, however, the reading instruction for those at risk for reading failure does not become more intensive. Instead, the at-risk students continue to participate in the PALS intervention, as they would have done regardless of their reading status. Although attempts are made to pair them with stronger readers, or "coaches," the intervention strategy itself remains the same, albeit possibly stronger than in many other school boards in which well-validated, standardized

interventions are not used at all. Those students who continue to struggle either near the end of the implementation of the PALS program or beyond it are then allowed to access a Tier Three psychological assessment, as described above. This assessment report is then used to plan further individual programming for those students.

Following psychological evaluation, school staff should have access to a profile of the student's areas of strength and challenge. Intervention should therefore become targeted at those specific areas in which the student is manifesting deficits, as highlighted by the psychological assessment report. For example, if the student struggles with executive skills such as planning and organizing what he or she reads, or keeping just-read information in mind for the purpose of using it to understand a passage, then intervention strategies would be concentrated in this area, rather than on learning basic phonics/decoding skills. Alternatively, if the student has not mastered reading at the most basic letter-sound level, even when entering Grade 4 and beyond, then this should be the focus of ongoing intervention. Students who continue to struggle with the basics of reading are often recommended to continue to access the PALS program within the WECDSB, although CBM data do not appear to be collected and/or recorded for these students. Therefore, one recommendation would be to continue to track the progress of those students who continue in the PALS program beyond Grade 3.

Additionally, as administrators within the WECDSB seem to already be doing (given the trouble herein with finding appropriate psychological assessment reports for analysis), it might be advisable, at least for students suspected of having a reading disability, to wait until these students are about to age out of the PALS program to implement psychological assessment. The years beyond PALS are the period of time

when an advanced understanding of students' individual learning needs will be necessary, as school board personnel can no longer rely on PALS as a wide-ranging intervention strategy that will target early reading skills in general, unless a recommendation for continued participation in the PALS program is suggested as part of the psychological assessment report. The PALS program is assumed to be doing a good job of remediating early reading problems, as well as identifying those students who need further remediation and attention beyond the discontinuation of the program, although again, a measure of concurrent validity and a control group are needed to further support this hypothesis. Based on a synthesis of two of the present studies, it could be most useful to recommend psychological assessment for those students whose scores have consistently fallen at least one standard deviation below the mean across the early years of school, as these are the students who are thought most likely to be accurately diagnosed with a reading disability upon psychological assessment. One downside to doing this, however, is that it may be unethical to delay formal testing if it is known that a student is really struggling to learn to read before the end of the PALS program. For those students, formal evaluation might still be warranted. As well, if a diagnosis of a reading disability was applied to a particular student, it would necessary to stress to his/her teaching team that this diagnosis should not mean that we "give up" on trying to remediate reading problems that persist beyond the early years of intervention. Rather, ongoing efforts should continue to be made. In other words, the reading disability label is simply a diagnostic tool for describing the student's difficulties.

Implications for the Developmental Model of Reading

The present results point to Grade 1 as being a very important time in terms of the development of reading skills. As mentioned previously, Chall (1983) posits that reading progresses in qualitatively distinct stages. From the beginning of reading acquisition in early elementary school up through more advanced reading refinement in high school, students transition from learning the mechanics of reading to being able to develop more advanced and synthesized knowledge based on what they read. The Grade 1 year maps onto these stages by marking the transition from learning basic decoding skills to learning to read with increased fluency, which is expected to occur at about age 6 or 7. Those students who have mastered early reading skills are expected to be more likely to practice these skills over the summer vacation, and to therefore come back in the fall at a level of achievement that exceeds that of their struggling peers. In Grade 1, the students who are at risk for reading failure do not yet tend to regress over the summer; rather, they do not make gains consistent with their typically-achieving peers. Taken together with the results of Menard and Wilson (2013), however, they are likely to eventually fall behind their own pre-summer achievement when they return to school in the fall, as by Grades 4 to 6, students with reading disabilities tend to regress to a greater degree than their non-disabled peers on measures of automatic decoding. It appears as though the summer discrepancy between achievement groups may begin in Grade 1, and appears to become more pronounced by Grades 4 to 6. Looking forward into the future, it is expected that summer regression will continue to have a deleterious effect on the reading ability of the lowest-scoring readers; one that could become worse every year.

The non-significance of language prosody as a predictor of reading achievement also has implications for the development of reading skills. It appears as though language prosody may not be related to reading achievement in the earliest years of schooling, although the results for the Grade 2/3 sample were trending on significance, meaning that language prosody may start to have an influence on higher-order reading skills. According to Chall's model, at about age 9, students are beginning to acquire a store of background information and a growing vocabulary. It stands to reason that this could be the time when language prosody might become important, as it maps onto a time when students' understanding of the world that they are reading about is becoming more nuanced and integrated. In the "learning to read" stages, students are working on mastering the very basics of reading (decoding, and eventually decoding fluency), without much concern for the intersection of language variables. Early readers are operating with limited cognitive resources to begin with, and especially so when reading is a cognitive challenge that takes up more resources than would be expected (Troia, 2004). It could be that only when the basics have been mastered do students have the mental resources to account for language prosody when reading. It is conceivable that language prosody might still be an interesting point of intervention in the later years of reading, as a more holistic or integrated mental set is becoming possible.

Use of CBMs to Monitor Reading Acquisition

The use of curriculum-based measurements as a tool for tracking students' progress can also be evaluated. The CBMs associated with the PALS program were designed to track students' progress in reading over the course of the school year, and the measurements do provide a lot of information about students' school-year learning

trajectories. The CBMs, however, are less ideal for tracking progress over the summer. This is especially true of the higher-order CBMs (PRF and MAZE Fluency), as the grade-based jumps in the complexity of the passages administered made cross-year achievement tracking impossible from a statistical standpoint. This is not to say that the higher-order variables cannot be considered anecdotally across school years, however. It could be highly informative to teaching staff if, for example, a student ends the year with a high PRF score, and then comes back with a much lower score in September. A change such as this could indicate that the student may not yet be ready to make the jump to a more difficult grade-appropriate passage, and may therefore need extra help in catching up to grade-based expectations.

Additionally, WECDSB personnel might consider adding a September or near-September administration point for the MAZE Fluency test, as it is currently only administered in January and June of the Grades 2 and 3 school years. Adding a September administration point would (a) allow for the calculation of full-year learning trajectories, and (b) allow for anecdotal comparisons between end- and beginning-of-year scores such as those described above. Based on the present data, it also appears as though not a lot of progress is made on the MAZE Fluency measure during the Grade 2 school year. Therefore, it is especially important to consider these data based on multiple school years when attempting to chart learning growth within WECDSB students.

One of the major strengths of using CBMs is that students' performances are compared to their own previous levels of achievement. Thus, the CBMs represent a self-contained tracking method wherein students' individual scores are not dependent on the performance of any other student, as would be true with norm-referenced scores, so it is

possible to track students' trajectories relative to their own earlier scores. Also, at least two facets of reading achievement are assessed at any given point in time within the PALS program as it is administered within the WECDSB. For example, letter-sound fluency and word identification fluency are both assessed early on, and then measurement shifts to word identification fluency and passage reading fluency, and finally to passage reading fluency and MAZE fluency. Thus, the CBMs also provide a more comprehensive assessment of reading progress across multiple domains than would be achieved by simply administering one standardized test of reading achievement. Finally, when taken altogether, the CBMs demonstrate a low floor and a high ceiling, with many possible scores in between. In doing this, CBMs come much closer than any one standardized test does to capturing the full spectrum of reading achievement.

Comment on Study Design

The present three studies make use of both archival and newly-collected data, although all three of them included an archival component. Jones (2010) provided a discussion of the advantages and disadvantages of using archival data in research in psychology, offering a critique of archival data use with respect to general research procedures, research design, measures, and samples. Within the listed advantages, Jones (2010) discussed the opportunity for collaboration between researchers on a single dataset. While this was not an opportunity for the present analysis, as this data had not been aggregated and analysed prior to these studies, it could be an opportunity in the future if interest were to be expressed by staff at the WECDSB in having access to the full dataset that was put together by this researcher. In terms of research design, Jones (2010) described the advantage of being able to address research questions that involve

past historical time periods. This is not an especially relevant advantage within the present work. However, Jones (2010) also discussed the advantage of having access to longitudinal data without having to personally wait years for data collection to occur. This is a particularly relevant advantage within the present work, as the present data were collected over multiple school years. Finally, Jones (2010) described sampling procedures as being possibly advantageous when utilizing archival data. Within the present work, the possibility for a large sample size and good sample representativeness are relevant. Particularly for the summer loss study, this work had a large sample size that was thought to be representative of the population from which it was taken (i.e., students attending schools within the WECDSB).

Jones (2010) also discussed the disadvantages of using archival data within research in psychology, which are described as being most commonly centered on general research procedures and measures. First, it is often necessary to go to great lengths to obtain permission to use an extant dataset, although fortunately, that was not a concern in this case. Furthermore, Jones (2010) described the amount of effort needed to truly understand the research design used, the sample and population from which it was drawn, the measures, and the general procedures used for collecting and coding the data. It is sincerely hoped that this effort was noted by the reader throughout the preceding pages of this document, but it should be acknowledged that human error is possible in interpreting any of the above-listed issues. An additional problem that was acknowledged by Jones (2010) and relevant here is that the data that were collected via the CBM probes were not originally intended for use in multi-year longitudinal research. The fact that these data have been appropriated for this type of research, when the original intention

was not such, is important to acknowledge. Furthermore, Jones (2010) also discussed disadvantages that relate to the measures used in archival research, which center on the appropriateness of the measures in addressing the research question of interest. In this case, it was felt that the CBMs adequately addressed the research questions, although it would have been especially useful to have a measure of concurrent validity to strengthen the conclusions that could be drawn from these data.

Outside of the issues discussed by Jones (2010), there are some additional issues with the present data that need to be acknowledged as well. Specifically, this researcher was not able to personally oversee the data collection. A great deal of trust, therefore, is placed in the competence of the educators who did collect the CBM data. As described in the data cleaning section in the summer learning loss study, there were a large number of alphabetical anomalies and nonsensical numerical data that needed to be removed. Attempts were made to clean the data as best as possible, but there is no guarantee that every datum accurately captures each student's achievement. However, there is human error inherent in newly-collected data as well. It is never guaranteed that any method of data collection or measurement is free from error, but it is important to acknowledge that for the present three studies, the use of archival data meant that there was yet another level of removal from the data when compared to data collection methods that are overseen by the researcher.

Furthermore, another unfortunate circumstance was that it was necessary to work within the parameters set by the available data. This meant that this researcher had diminished control over the study design. Examples of less-than-ideal situations regarding the available data include, but are not limited to, (1) the discontinuation of the

PALS program prior to the implementation of the psychological assessment reports in many cases, (2) the jumps in difficulty levels of the PRF measure between school years, making it impossible to obtain a reliable cross-summer metric in the later years of PALS administration, and (3) the MAZE Fluency measure only being administered twice per school year. These circumstances were unavoidable within the present studies, as they reflect the way that the data are collected, over which this researcher had no control. Again, although the CBM data are thought to provide a strong metric of reading achievement over time, it would have been far more ideal to have had control over the way that the measures are administered, collected, and recorded. As well, it would have been ideal to have control over the timing of the administration of the measures, as it is possible that the intervals at which the measurements were administered and the timing of the measures in general may not have been optimal. For example, based on knowledge of how schools operate, there is reason to believe that the September and June measurements might be less reliable when compared to mid-year administrations due to increased chaos and confusion at the beginning and end of the school years. However, there was no way to get around using these administration points within the present work, so this must be acknowledged as an additional constraint on the part of the researcher and the research design.

Another design difficulty inherent in these data is that it may be seen as tautological to try to separate the CBMs from responsiveness to PALS, as the measurements and the intervention program are intimately linked with each other. In other words, progress in the PALS program is measured by scores on tests that were specifically designed for measuring progress in the PALS program. Although this

strategy could be considered a strength in terms of measurement relevance, it also presents a design issue when used within a research context. It would have been ideal to have access to a measure of concurrent validity when offering conclusions about the meaning of the CBMs, but because the school board is not administering these measurements within the direct context of a research protocol, a measure of reading progress that can be better separated from the PALS intervention was not administered to the students. Therefore, no such measure of concurrent validity was available for the present analyses. Given this design issue, conclusions regarding responsiveness to intervention were offered with special hesitancy at times when confusion between (a) the PALS program itself, and (b) the CBMs that are administered as a part of the PALS program, was possible. Ideal assessments for research purposes would constitute the regular PALS CBMs being administered, along with standardized reading measures like those on the Wechsler Individual Achievement Test or the Gray Oral Reading Test, for example. And finally, based on the work of Protopapas and colleagues (2011), it is known that using the identical criterion and outcome variables was a design flaw that likely resulted in the statistical phenomenon of regression toward the mean when tracking reading achievement over the school year in all three of the present studies, so it would be ideal to have different standardized achievement measures administered at the beginning and the end of each school year.

Conclusions Regarding the Utility of CBMs

Student achievement tracking tends to be accomplished by ministry-wide standardized testing as well as teacher-specific methods of tracking how their individual students are progressing in reading. The CBMs associated with the PALS program might

provide information that would serve to bridge the gap between the very global information provided by standardized provincial testing and the very detailed information that is provided by teacher-specific methods, which may or may not be systematic and research-informed. Another currently-accepted method of tracking students' achievement is by standardized testing administered by a school psychologist, which is then written up in a psychological assessment report. As highlighted previously, it appears as though psychological assessment reports may not be useful tools to teachers who already have access to ongoing measurements of their students' reading trajectories, at least while these students are still in the early years of school and still engaged in a well-validated reading intervention program. Therefore, CBMs may provide a good alternative strategy for identifying those students who will require more intensive intervention and specialized reading intervention services.

REFERENCES

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review, 72,* 167-180.

Allmendinger, J. (1989). Educational systems and labor market outcomes. *European Sociological Review, 5,* 231-250.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA.

Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade children. Journal of Educational Psychology, 96, 43-55.

Ashby, J. & Clifton Jr., C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96,* B89-B100. doi:10.1016/j.cognition.2004.12.006

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Baron, J., & Norman, M. F. (1992). SATs, achievement tests, and high-school class rank as predictors of college performance. *Educational and Psychological Measurement, 52,* 1047-1047.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology, 74, 506-521.

Bjaalid, I.-K., Hoien, T., & Lundberg, I. (1997). Dual-route and connectionist models: A step towards a combined model. *Scandinavian Journal of Psychology, 38,* 73-82. DOI: 10.1111/1467-9450.00011

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., Shaywitz, B. A., & Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology, 96,* 444-461. DOI: 10.1037/0022-0663.96.3.444

Bradley, R., Danielson, L. E., & Hallahan, D. P. (2002). *Identification of learning disabilities: Research to practice.* Mahwah, NJ: Erlbaum.

Breen, M. & Clifton Jr., C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language, 64,* 153-170. doi:10.1016/j.jml.2010.11.001

Bruck, M. (1990). Word recognition skills of adults with childhood dyslexia. Developmental Psychology, 26, 439-454.

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

Burkam, D. T., Ready, D. D., Lee, V. E., & LoGerfo, L. F. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. *Sociology of Education, 77,* 1-31.

Bus, A. G., & von IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. Journal of Educational Psychology, 91, 403-414.

Chall, J. (1983). Stages of reading development. New York: McGraw-Hill.

Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse processing. *Brain and Cognition, 46,* 73-82. doi:10.1016/S0278-2626(01)80038-5

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology, 57,* 505-528. DOI: 10.1146/annurev.psych.57.102904.190146

Cooper, H. (2003). *Summer learning loss: The problem and some solutions.* Retrieved October 14th, 2014 from http://ecap.crc.illinois.edu/eecearchive/digests/2003/cooper03.pdf.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66,* 227-268.

Cowie, R., Douglas-Cowie, E., & Wichmann, A. (2002). Prosodic characteristics of skilled reading: Fluency and expressiveness in 8 – 10-year old readers. *Language and Speech, 45,* 47-82.

Dahl, K. L., & Scharer, P. L. (2000). Phonics teaching and learning in whole language classrooms: New evidence from research. The Reading Teacher, 584-594.

Dally, K. (2006). The influence phonological processing and inattentive behaviour on reading acquisition. Journal of Educational Psychology, 98, 420-437.

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30,* 507-524.

Donfrancesco, R., Iozzino, R., Caruso, B., Ferrante, L., Mugnaini, D., Talamo, A., ... & Masi, G. (2010). Is season of birth related to developmental dyslexia? *Annals of dyslexia*, *60*, 175-182.

Dowhower, S.L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, *22*, 389-406.

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice, 30,* 165-175.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. Reading Research Quarterly, 36, 250-287.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9,* 167-188.

Field, A. (2009). Discovering statistics using SPSS, Third Edition. London, UK: Sage Publications Ltd.

Flannery, K. A., Liederman, J., Daly, L., & Schultz, J. (2000). Male prevalence for reading disability is found in a large sample of black and white children free from ascertainment bias. *Journal of the International Neuropsychological Society, 6,* 433-442.

Fletcher, J. M. (2006). The need for response to instruction models of learning disabilities. Retrieved August 31st, 2011 from http://www.nhsped.net/docs/RTIfletcher.pdf.

Fletcher, J. M. (2008). Identifying learning disabilities in the context of response to intervention: A hybrid model. Retrieved August 31st, 2011 from http://blogs.utexas.edu/dmp/files/2010/01/FletcherHybridModelRTI.pdf.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stubing, K. K., Francis, D. J., et al. (1994). Cognitive profiles of reading disability: Comparisons of the discrepancy and low achievement definitions. *Journal of Educational Psychology, 86,* 6-23.

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54,* 304-331. DOI: 10.1007/s11881-004-0015-y

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*(3), 506-522.

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From Identification to Intervention.* New York: Guilford Press.

Fletcher, J. M., & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives, 3,* 30-37.

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research, 27,* 285–319.

Foorman, B. R. (1995). Research on "the Great Debate": Code-oriented versus whole language approaches to reading instruction. *School Psychology Review, 24,* 376-392.

Francis, D. J., Fletcher, J. M., Stuebing, K., K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38,* 98-108. doi: 10.1177/00222194050380020101

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer Assisted Learning Strategies in reading: A manual. Unpublished manuscript, Vanderbilt University, Nashville, TN.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18,* 157-171.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education, 7,* 6-10.

Grant, A., Wilson, A. M., & Gottardo, A. (2007). The role of print exposure in reading skills of postsecondary students with and without reading disabilities. *Exceptionality Education Canada, 17,* 175-194.

Hale, J. B., Naglieri, J. A., Kaufman, A. S., & Kavale, K. A. (2004). Specific learning disability classification in the new Individuals with Disabilities Education Act: The danger of good ideas. *The School Psychologist.* Retrieved October 14th, 2014 from http://www.wce.wwu.edu/Depts/SPED/Forms/Kens%20Readings/LD/LD%20SLD%20classification%20in%20the%20new%20IDEA%20Hale%202004.pdf.

Hale, J. B., Kaufman, A., Naglieri, J. A., Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools, 43,* 753-770. DOI: 10.1002/pits.20186

Holliman, A. J., Wood, C., & Sheehy, K. (2010). The contribution of sensitivity to speech rhythm and non-speech rhythm to early reading development. *Educational Psychology, 30,* 247-267.

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how? *The Reading Teacher, 58,* 702-714. doi:10.1598/RT.58.8.1

International Dyslexia Association (2008). Definition of dyslexia. Retrieved October 14th, 2014 from http://www.interdys.org/ewebeditpro5/upload/Definition.pdf.

Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and the reading recovery program. Journal of Educational Psychology, 85, 112-126.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage, 20,* 693-712. DOI: 10.1016/S1053-8119(03)00343-4

Jones, C. (2010). Archival data: Advantages and disadvantages for research in psychology. *Social and Personality Psychology Compass, 4,* 1008-1017. DOI: 10.1111/j.1751-9004.2010.00317.x

Kaplan, D., & Walpole, S. (2005). A stage-sequential model of reading transitions: Evidence from the Early Childhood Longitudinal Study. *Journal of Educational Psychology, 97,* 551-563. DOI: 10.1037/0022-0663.97.4.551

Katsiyannis, A. (1991). Extended school year policies: An established necessity. *Remedial and Special Education, 12,* 24-28.

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes, 39,* 125-128, DOI: 10.1080/0163853X.2005.9651676

Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. Reading Research Quarterly, 45, 341–362.

Kuhn, M. R. & Stahl, S. A. (2000). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95,* 3-21.

Learning Disabilities Association of Canada (2002). LD defined: Official definition of learning disabilities. Retrieved October 14th, 2014 from http://www.ldac-acta.ca/learn-more/ld-defined.

Learning Disabilities Association of Canada (2007a). Putting a Canadian Face on Learning Disabilities and the Prevalence of Learning Disabilities. Retrieved September 11th, 2011 from http://www.ldac-acta.ca/en/learn-more/ld-basics/prevalence-of-lds/56-putting-a-canadian-face-on-learning-disabilities-pacfold-and-the-prevalence-of-learning-disabilities.html.

Learning Disabilities Association of Canada (2007b). *Putting a Canadian Face on Learning Disabilities.* Retrieved October 14th, 2014 from http://www.pacfold.ca/download/WhatIs/en/Highlights.pdf.

Learning Disabilities Association of Ontario (2011). Official definition of LDs. Retrieved April 29th, 2015 from http://www.ldao.ca/introduction-to-ldsadhd/introduction-to-ldsadhd/what-are-lds/official-definition-of-lds/.

Lyon, G. R. (1996). Learning disabilities. *The Future of Children: Special Education for Students with Disabilities, 6, 54-76.*

Marshall, C. R., Harcourt-Brown, S., Ramus, F., & van der Lely, K. K. J. (2009). The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *International Journal of Language and Communication Disorders, 44,* 466-488. doi: 10.1080/13682820802591643

Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-Assisted Learning Strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 33,* 62-94.

McBride-Chang, C. (1995). What is phonological awareness? Journal of Educational Psychology, 87, 179-192.

McMaster, K., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2002). Monitoring the academic progress of children who are unresponsive to generally effective early reading intervention. *Assessment for Effective Intervention, 27,* 23-33. doi: 10.1177/073724770202700404.

Menard, J. & Wilson, A. M. (2013). Summer learning loss among elementary school children with reading disabilities. *Exceptionality Education International, 23*, 72-85.

Miller, J. & Schwanenflugel, P. J. (2006). Prosody or syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology, 98,* 839-853. doi: 10.1037/0022-0663.98.4.839

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. Developmental Psychology, 40, 665-681.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups. National Institute of Child Health and Human Development, National Institutes of Health.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. Journal of Educational Psychology, 98, 554-566.

Otto, J. (1982). The new debate in reading. *Reading Teacher, 36,* 14-18.

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics and Phonetics, 17,* 345-354. doi: 10.1080/0269920031000079994

Perfetti, C. A., Landi, N. & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In *The Science of Reading: A Handbook* (M. J. Snowling and C. Hulme, Eds.). Blackwell Publishing Ltd: Oxford, UK. DOI: 10.1002/9780470757642.ch13

Pfost, M., Hattie, J., Dörfler, T., & Artlet, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew Effects in reading. *Review of Educational Research, 84,* 203-244. DOI: 10.3102/0034654313509492.

Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. Journal of Educational Psychology, 80, 319-323.

Proctor, B., & Prevatt, F. (2003). Agreement among four models used for diagnosing learning disabilities. *Journal of Learning Disabilities*, *36*(5), 459-466. DOI: 10.1177/00222194030360050701

Protopapas, A., Sideridis, G. D., Mouzaki, A., & Simos, P. G. (2011). Matthew Effects in reading comprehension: Myth or reality? *Journal of Learning Disabilities, 44,* 402-420. DOI: 10.1177/0022219411417568

Protopapas, A., Parrila, R., & Simos, P. G. (2014). In search of Matthew Effects in reading. *Journal of Learning Disabilities,* 1-16. DOI: 10.1177/0022219414559974.

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., ... & Carroll, J. (2004). Sex differences in developmental reading disability: new findings from 4 epidemiological studies. *Journal of the American Medical Association, 291,* 2007-2012. DOI:10.1001/jama.291.16.2007

Sargent, L. R., & Fidler, D. A. (1987). Extended school-year programs: In support of the concept. *Education and Training in Mental Retardation, 22,* 3-9.

Sattler, J. M., & Hoge, R. D. (2006). *Specific learning disabilities: Assessment and intervention.* In Assessment of Children: Behavioral, Social, and Clinical

Foundations, Fifth Edition (pp.411-430). Jerome M. Sattler, Publisher, Ltd: La Mesa, CA.

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology, 96,* 119-129. doi: 10.1037/0022-0663.96.1.119

Semrud-Clikeman, M. (2005). Neuropsychological aspects for evaluating learning disabilities. *Journal of Learning Disabilities, 38,* 563-568.

Shaw, T. V. (1982). *Retention of selected reading and arithmetic skills by learning disabled pupils and non-disabled pupils over summer vacation.* Unpublished master's thesis, California State College, Stanislaus, CA.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, Vol. 13, pp. 290–312.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-407.

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

Stevens, J. P. (2009). Applied multivariate statistics for the social sciences, Fifth edition. New York, NY: Routledge.

Torgesen, J. K. (2005). Recent discoveries from research on remedial interventions for children with dyslexia. *The Science of Reading,* 521-537.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). Test of Word Reading Efficiency. Austin, TX: Pro-Ed.

Troia, G. A. (2004). Phonological processing and its influence on literacy learning. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel, Handbook of Language and Literacy: Development and Disorders, (pp. 271-301). New York, NY: The Guilford Press.

U.S. Department of Education (2004). Individuals with Disabilities Education Act of 2004, Pub. L. 108-466. *Federal Register, 70,* 35802-35803.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. Psychological Bulletin, 101, 192-212.

Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. Journal of Educational Psychology, 85, 83-103.

Wechsler Individual Achievement Test. (1992). San Antonio, TX: Psychological Corporation.

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

Wells, B. & Peppé, S. (2003). Intonation abilities of children with speech and language impairments. *Journal of Speech, Language, and Hearing Research, 46,* 5–20.

Yopp, H. K., & Yopp, R. H. (2000). Supporting phonemic awareness development in the classroom. The Reading Teacher, 54, 130-143.

Table 1.

Correlations between CBM variables.

Correlations

		JKLSFOutcome	JKWIFOutcome	SKLSFOutcome	SKWIFOutcome	G1LSFOutcome	G1WIFOutcome	G1PRFOutcome	G2PRFOutcome	G2MAZEOutcome	G3PRFOutcome	G3MAZEOutcome
JKLSFOutcome	Pearson Correlation	1	.600**	.641**	.609**	.392**	.533**	.505**	.488**	.522**	.199	.130
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.001	.000	.514	.687
	N	1940	1336	754	818	407	410	361	43	43	13	12
JKWIFOutcome	Pearson Correlation	.600**	1	.325**	.495**	.247	.358	.635**	.579**	.470*	-.078	-.161
	Sig. (2-tailed)	.000		.000	.000	.255	.093	.001	.002	.020	.792	.617
	N	1336	1534	473	549	23	23	22	25	24	14	12
SKLSFOutcome	Pearson Correlation	.641**	.325**	1	.609**	.524**	.672**	.489**	.545**	.548**	.372**	.484**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	754	473	2029	1949	956	954	910	365	376	96	119
SKWIFOutcome	Pearson Correlation	.609**	.495**	.609**	1	.262**	.550**	.559**	.544**	.469**	.514**	.391**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000
	N	818	549	1949	3096	1672	1679	1648	984	1004	231	267
G1LSFOutcome	Pearson Correlation	.392**	.247	.524**	.262**	1	.577**	.417**	.403**	.343**	.429**	.316**
	Sig. (2-tailed)	.000	.255	.000	.000		.000	.000	.000	.000	.000	.000
	N	407	23	956	1672	3488	3470	2661	1872	1768	1071	1070
G1WIFOutcome	Pearson Correlation	.533**	.358	.672**	.550**	.577**	1	.789**	.779**	.625**	.715**	.606**
	Sig. (2-tailed)	.000	.093	.000	.000	.000		.000	.000	.000	.000	.000
	N	410	23	954	1679	3470	3603	2688	1912	1803	1095	1094
G1PRFOutcome	Pearson Correlation	.505**	.635**	.489**	.559**	.417**	.789**	1	.805**	.616**	.753**	.657**
	Sig. (2-tailed)	.000	.001	.000	.000	.000	.000		.000	.000	.000	.000
	N	361	22	910	1648	2661	2688	2859	1519	1504	554	603
G2PRFOutcome	Pearson Correlation	.488**	.579**	.545**	.544**	.403**	.779**	.805**	1	.653**	.827**	.673**
	Sig. (2-tailed)	.001	.002	.000	.000	.000	.000	.000		.000	.000	.000
	N	43	25	365	984	1872	1912	1519	3718	2680	2060	1974
G2MAZEOutcome	Pearson Correlation	.522**	.470*	.548**	.469**	.343**	.625**	.616**	.653**	1	.572**	.635**
	Sig. (2-tailed)	.000	.020	.000	.000	.000	.000	.000	.000		.000	.000
	N	43	24	376	1004	1768	1803	1504	2680	2855	1307	1344
G3PRFOutcome	Pearson Correlation	.199	-.078	.372**	.514**	.429**	.715**	.753**	.827**	.572**	1	.700**
	Sig. (2-tailed)	.514	.792	.000	.000	.000	.000	.000	.000	.000		.000
	N	13	14	96	231	1071	1095	554	2060	1307	4025	2849
G3MAZEOutcome	Pearson Correlation	.130	-.161	.484**	.391**	.316**	.606**	.657**	.673**	.635**	.700**	1
	Sig. (2-tailed)	.687	.617	.000	.000	.000	.000	.000	.000	.000	.000	
	N	12	12	119	267	1070	1094	603	1974	1344	2849	3049

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 2.

Test-retest and parallel forms reliability estimates for baseline scores on CBM variables.

	Test-retest correlation coefficient	Parallel forms reliability coefficient
JK LSF	.892**	.935
SK LSF	.897**	.938
Grade 1 LSF	.862**	.922
Grade 1 WIF	.954**	.975
Grade 2 PRF	.925**	.955
Grade 2 MAZE Fluency	.648**	.781
Grade 3 PRF	.929**	.961
Grade 3 MAZE Fluency	.768**	.867

***Correlation is significant at the .01 level (2-tailed).*

Table 3.

Means and Standard Deviations of PRF and MAZE progress over the Grade 3 school year.

	N	Mean	SD
Report PRF Progress	16	35.656	22.509
Control PRF Progress	16	48.594	23.254
Report MAZE Progress	11	.909	3.080
Control MAZE Progress	11	1.363	4.965

Table 4.

Average test scores of each achievement group at key time points, split by first grouping (Longitudinal Group A, grouped based on LSF outcome in JK) vs. second grouping (Longitudinal Group B, grouped based on LSF outcome in SK).

Longitudinal Group A					
Achievement Group (Group thresholds from LSF test at JK outcome)	Test		End-of-Year JK Outcome	Start-of-Year SK Baseline	End-of-Year SK Outcome
Low (< 11.0)	LSF	Mean	5.46	4.82	34.82
		SD	3.48	4.69	17.52
		N	156	156	131
Average (11.0 – 48.6)	LSF	Mean	28.20	19.80	60.63
		SD	10.84	9.93	21.08
		N	674	674	615
High (> 48.6)	LSF	Mean	61.61	38.23	83.85
		SD	12.54	15.03	23.17
		N	109	109	94
Longitudinal Group B					
Achievement Group (Group thresholds from LSF test at SK outcome)	Test		End-of-Year SK Outcome	Start-of-Year Grade 1 Baseline	End-of-Year Grade 1 Outcome
Low (< 30.1)	LSF	Mean	20.46	17.32	65.62
		SD	9.22	15.92	25.21
		N	201	201	189
	WIF	Mean	2.56	2.51	33.20
		SD	3.83	4.15	22.16
		N	200	200	192
Average (30.1 - 78.9)	LSF	Mean	54.86	40.60	90.56
		SD	13.12	14.24	20.27
		N	702	702	694
	WIF	Mean	12.19	14.13	68.02
		SD	13.14	15.88	21.90
		N	707	707	692
High (> 78.9)	LSF	Mean	95.28	61.98	112.52
		SD	13.36	15.92	19.86
		N	127	127	129
	WIF	Mean	34.09	38.71	92.19
		SD	23.31	24.73	17.87
		N	127	127	127

Table 5.

Descriptive data on summer learning loss, time to regain loss, and school year trajectory.

	Grade	Test		Low	Average	High
Summer learning loss	JK-to-SK	LSF	Mean	-0.64***	-8.40	-23.37***
			SD	4.41	8.58	14.60
			n	156	674	109
	SK-to-Grade 1	LSF	Mean	-3.14***	-14.27	-33.30***
			SD	13.60	13.72	16.56
			n	201	702	127
		WIF	Mean	-0.05*	1.94	4.61**
			SD	33.02	23.58	21.35
			n	200	707	127
Time to regain (days)	JK-to-SK	LSF	Mean	39.60***	85.75	141.51***
			SD	55.88	80.49	93.12
			n	144	673	111
	SK-to-Grade 1	LSF	Mean	43.22***	92.82	187.38***
			SD	64.79	84.94	103.44
			n	198	708	132
		WIF	Mean	32.84*	23.58	21.3
			SD	58.95	42.99	44.50
			n	185	611	121
School-year trajectory	SK	LSF	Mean	30.10***	40.64	45.49*
			SD	16.35	17.77	20.98
			n	130	603	92
	Grade 1	LSF	Mean	48.69	49.99	49.75
			SD	22.96	20.56	19.95
			n	184	674	123
		WIF	Mean	33.02*	23.58	21.33
			SD	59.07	42.99	44.48
			n	185	611	121

Asterisks are used to indicate whether the post-hoc test showed that the Low or High groups' mean scores were significantly different as compared to the Average group (***$p<0.001$, **$p<0.01$, *$p<0.05$).

Table 6.

Opportunity-cost of summer break.

Summer	Test	Achievement Group	Summer Trajectory (Change in score/month)	School Year Trajectory (Change in score/month)	Opportunity Cost of Summer	
					Test scores forgone	Months worth of reading skills forgone
JK-to-SK	LSF	Low	-0.32	3.01	6.66	2.21
		Average	-4.20	4.06	16.53	4.07
		High	-11.69	4.55	**32.47**	**7.14**
SK-to-Grade 1	LSF	Low	-1.57	4.87	12.88	2.64
		Average	-7.14	5.00	26.27	5.25
		High	-16.65	4.98	**43.25**	**8.68**
	WIF	Low	-0.03	3.07	6.18	**2.01**
		Average	0.97	5.38	**8.82**	1.64
		High	2.31	5.28	5.95	1.27

Table 7.

Chart of changes in students' learning statuses across school years.

Group at end of SK	Group in subsequent and preceding years	JK	SK	G1	G2	G3
Low	Low	55	332	84	36	8
	Average	37	0	72	48	15
	High	0	0	5	4	0
	Low	56	0	56	18	6
Average	Average	426	1397	491	194	51
	High	39	0	86	34	9
	Low	2	0	0	0	1
	Average	91	0	69	15	4
High	High	48	300	47	16	2

Table 8.

Correlation matrix for receptive prosody variables in the Grade 1 sample.

	Affect Input score	Chunking Input score	Focus Input score	Intonation Input score	Prosody Input score	Turn-end Type Input score
Affect Input score	1.000	.363	.214	.445	.361	.543
Chunking Input score	.363	1.000	-.014	.296	.080	.205
Focus Input score	.214	-.014	1.000	-.021	.048	.086
Intonation Input score	.445	.296	-.021	1.000	.810	.309
Prosody Input score	.361	.080	.048	.810	1.000	.141
Turn-end Type Input score	.543	.205	.086	.309	.141	1.000

Table 9.

Correlation matrix for expressive prosody variables in the Grade 1 sample.

	Affect Output score	Chunking Output score	Focus Output score	Intonation Output score	Prosody Output score	Turn-end Type Output score
Affect Output score	1.000	-.024	.266	.378	.246	.579
Chunking Output score	-.024	1.000	-.147	.175	.259	.112
Focus Output score	.266	-.147	1.000	.526	.622	.208
Intonation Output score	.378	.175	.526	1.000	.760	.414
Prosody Output score	.246	.259	.622	.760	1.000	.440
Turn-end Type Output score	.579	.112	.208	.414	.440	1.000

Table 10.

Correlation matrix for receptive prosody variables in the Grades 2/3 sample.

	Affect Input score	Chunking Input score	Focus Input score	Intonation Input score	Prosody Input score	Turn-end Type Input score
Affect Input score	1.000	.139	-.008	.010	-.094	.026
Chunking Input score	.139	1.000	.226	-.025	.097	.148
Focus Input score	-.008	.226	1.000	.076	-.003	-.012
Intonation Input score	.010	-.025	.076	1.000	.787	.197
Prosody Input score	-.094	.097	-.003	.787	1.000	.250
Turn-end Type Input score	.026	.148	-.012	.197	.250	1.000

Table 11.

Correlation matrix for expressive prosody variables in the Grades 2/3 sample.

	Affect Output score	Chunking Output score	Focus Output score	Intonation Output score	Prosody Output score	Turn-end Type Output score
Affect Output score	1.000	.341	.224	.354	.221	.401
Chunking Output score	.341	1.000	.242	.346	.408	.294
Focus Output score	.224	.242	1.000	.264	.414	.221
Intonation Output score	.354	.346	.264	1.000	.394	.285
Prosody Output score	.221	.408	.414	.394	1.000	.310
Turn-end Type Output score	.401	.294	.221	.285	.310	1.000

Table 12.

Descriptive data for expressive/receptive PEPS-C variables for Grade 1 LSF analysis.

	Receptive Prosody Composite	Expressive Prosody Composite	Grade 1 LSF Progress
Mean	65.58	54.57	45.79
SD	11.67	14.42	27.59
N	37	37	37

Table 13.

Descriptive data for expressive/receptive PEPS-C variables for Grade 1 WIF analysis.

	Receptive Prosody Composite	Expressive Prosody Composite	Grade 1WIF Progress
Mean	65.58	54.57	42.95
SD	11.67	14.42	21.86
N	37	37	37

Table 14.

Descriptive data for expressive/receptive PEPS-C variables for Grade 1 PRF analysis.

	Receptive Prosody Composite	Expressive Prosody Composite	Grade 1PRF Progress
Mean	64.87	54.40	30.14
SD	11.63	14.98	20.88
N	34	34	34

Table 15.

Descriptive data for expressive/receptive PEPS-C variables for Grades 2/3 PRF analysis.

	Receptive Prosody Composite	Expressive Prosody Composite	Grades 2/3 PRF Progress
Mean	77.24	67.34	43.97
SD	8.90	11.68	21.03
N	63	63	63

Table 16.

Descriptive data for expressive/receptive PEPS-C variables for Grades 2/3 MAZE Fluency analysis.

	Receptive Prosody Composite	Expressive Prosody Composite	Grades 2/3 MAZE Fluency Progress
Mean	77.30	67.51	4.42
SD	8.97	11.70	4.95
N	62	62	62

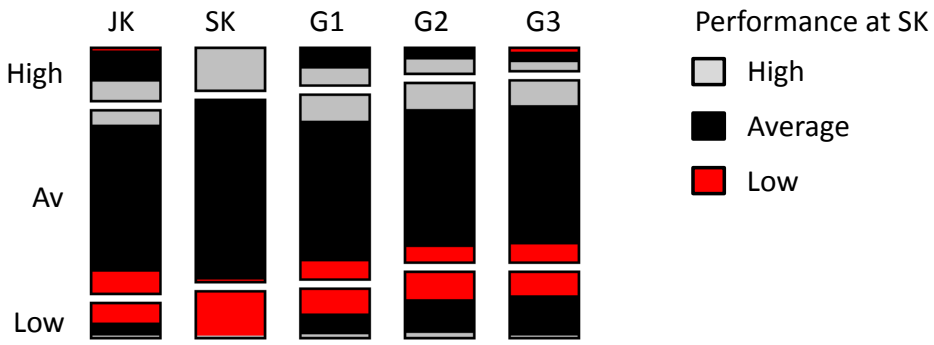

Figure 1. *Graphical representation of students' learning statuses across school years.*

APPENDICES

CONSENT TO PARTICIPATE IN RESEARCH

Title of Study: Prosody as a Mediator between Early Reading Achievement and Peer Assisted Learning Strategies Progress

Your child is asked to participate in a research study conducted by Jessica Menard, from the Psychology Department at the University of Windsor This research will contribute to the completion of Ms. Menard's Ph.D. dissertation project.

If you have any questions or concerns about the research, please feel to contact Ms. Menard's supervisor, Dr. Carlin J. Miller, at 519-253-3000 ext. 2226.

PURPOSE OF THE STUDY

The purpose of this study is to look at the connection between the processing of prosody and reading achievement. Prosody refers to the musical elements of speech, and includes melodic and rhythmic variations in pitch, stress, and duration. I expect that prosodic processing may associated with early reading achievement and success in learning to read via the Peer-Assisted Learning Strategies (PALS) method of instruction, which is currently administered every year within the Windsor-Essex Catholic District School Board (WECDSB).

PROCEDURES

It is anticipated that your child's participation will require approximately 1 hour and he or she is expected to be able to complete this assessment in one appointment. Students will be tested in a private room within the schools they attend. If he or she is uncomfortable or needs a break at any time, the child will be encouraged to let the examiner know. If your child volunteers to participate in this study, he or she will be asked to complete the following assessment tool:

- The Profiling Elements of Prosodic Systems – Child version (PEPS-C) tests both receptive and expressive prosodic ability. The test is computer-administered and involves a receptive component wherein participants will point to the answer on a laptop screen, and an expressive component wherein the students' responses will be recorded and then analysed by both the primary researcher and a research assistant.
- I will also gain access to his or her PALS assessment data (available through the school board) for the purposes of comparing the data from the prosody measure to the child's progress in reading over the school year.

POTENTIAL RISKS AND DISCOMFORTS

We do not think there is any significant risk associated with this study. Your child may feel somewhat worried or uncomfortable while completing the PEPS-C, especially if he or she has trouble with the items on this measure. If your child feels worried or upset, he or she will be encouraged to discuss any concerns with the examiner. If he or she continues to feel badly after leaving the assessment, you may contact the research supervisor, Dr. Carlin Miller, at her office (519-253-2000 ext. 2226).

POTENTIAL BENEFITS TO PARTICIPANTS AND/OR TO SOCIETY

Society may benefit if we better understand the relationships between prosody and reading achievement. If positive results are found, prosodic processing could constitute a new point of early intervention for children with reading problems.

COMPENSATION FOR PARTICIPATION

There will be no monetary of other direct compensation associated with participation in this study.

CONFIDENTIALITY

Any information you provide and that can be connected to you or your child will remain confidential. No one will be told what your child reported without your permission. We will not discuss your child's results with anyone without your written permission. Once your child's data is collected, his or her unique subject identification number will be written on every form and his or her name will be removed from all forms. These unique subject identification numbers will be used for data entry involving your child's responses. Consent forms and rating forms will be store separately in locked cabinets in a locked on-campus office. Only Dr. Miller will have access to your personally-identifying information. In the event these data are ever to be destroyed, their destruction will be carried out in a manner that will preserve you and your child's privacy.

There is one set of circumstances that would possibly necessitate a breach in confidentiality. In the event you or your child discloses that any child is in danger currently or experiencing abuse/neglect, we may tell the appropriate authorities. As individuals who work with children and families, both Dr. Miller and Ms. Menard are mandatory reporters for child abuse/neglect and are required by law to protect the rights of their research participants.

PARTICIPATION AND WITHDRAWAL

You can choose whether to let your child participate in this study or not. If you volunteer your child to be in this study, you may decide to remove your consent at any time without consequences of any kind. Specifically, we will not report to your decision to anyone, including WECDSB staff. You may also refuse to let your child answer any questions you don't want to answer and still have him or her remain in the study. The investigator may withdraw your child's from this research if circumstances arise which warrant doing so, such as it is discovered that your child does not meet eligibility criteria. In that event, Ms. Menard or Dr. Miller will discuss with you the reasons why your child is not eligible.

FEEDBACK OF THE RESULTS OF THIS STUDY TO THE PARTICIPANTS

The findings of this study will be posted on Dr. Miller's website.
Web address: http://www.uwindsor.ca/cjmiller
Date when results are available: Summer 2013

SUBSEQUENT USE OF DATA

These data may be used in subsequent studies, and in publications and presentations, however, the data will not be identifiable by subject name. All data will be reported on a group-level basis.

RIGHTS OF RESEARCH PARTICIPANTS

You may withdraw your consent at any time and discontinue your child's participation without penalty. If you have questions regarding your child's rights as a research participant, contact: Research Ethics Coordinator, University of Windsor, Windsor, Ontario, N9B 3P4; Telephone: 519-253-3000, ext. 3948; e-mail: ethics@uwindsor.ca

SIGNATURE OF RESEARCH PARTICIPANT/LEGAL REPRESENTATIVE

I understand the information provided for the study **"**Prosody as a Mediator between Early Reading Achievement and Peer Assisted Learning Strategies Progress" as described herein. My questions have been answered to my satisfaction, and I agree to let my child participate in this study. I have been given a copy of this form.

_____
Name of Child

_____

Name of Parent/Legal Guardian

_____ _____
Signature of Parent/Legal Guardian Date

SIGNATURE OF INVESTIGATOR

These are the terms under which I will conduct research.

_____ _____
Signature of Investigator Date

DEMOGRAPHIC INFORMATION

Child's Name: _____

Child's Current School: _____ Current Grade: ____

Parents'/Guardians Name/s: _____

Home Address: _____

Home Phone: _____

Cell Phone: _____

Email: _____

Name/phone number for another person who will know how to find you if we cannot reach you:

_____

May we contact you again in the future for other studies? YES _____ NO _____

Instructions: For questions that include numbered choice options, please circle the number(s) that best describes your answer. Other items will provide you with space(s) to provide a written response. Be sure to read each item carefully, and if you do not understand a question, please ask the person working with you. Please try to answer each item, however, if you feel uncomfortable with any question, you do not need to answer it. Your answers will be kept completely confidential. ***Please do not write your name or your child's name on any page but this front page. (This cover page will be detached and***

stored with your consent forms to protect your confidentiality.)

(FOR PROJECT USE ONLY – ID # _____)

NO TEXT ON THIS PAGE

FAMILY DEMOGRAPHIC INFORMATION

	Parent/Guardian 1 - YOU	**Parent/Guardian 2**
Gender	[1] FEMALE [2] MALE [3] TRANSGENDER [4] PREFER NOT TO ANSWER	[1] FEMALE [2] MALE [3] TRANSGENDER [4] PREFER NOT TO ANSWER
Lives in the home?	[1] YES [2] NO	[1] YES [2] NO
Employed outside the home?	[1] YES [2] NO	[1] YES [2] NO
Job title?		
Highest grade completed?	[1] Less than high school [2] High school [3] Some college/trade-school [4] College/trade-school [5] Some university [6] University [7] Graduate degree	[1] Less than high school [2] High school [3] Some college/trade-school [4] College/trade-school [5] Some university [6] University [7] Graduate degree

Are there any other adults living in the home? [1] YES [2] NO

If yes, please describe: _____

CHILD'S DEMOGRAPHIC INFORMATION

Date of Birth (MM/YY): ___/___ Today's Date (DD/MM/YY): ___/___/___

Sex: [1] FEMALE [2] MALE

Race/ethnic background: (please circle)

[1] ABORIGINAL

[2] ASIAN OR ASIAN DESCENT (NON-ARAB)

[3] HISPANIC/LATINO

[4] NON-HISPANIC BLACK OR AFRICAN DESCENT

[5] NON-HISPANIC WHITE OR EUROPEAN DESCENT

[6] ARAB OR MIDDLE-EASTERN DESCENT

[7] OTHER/MIXED (please describe) _____

VITA AUCTORIS

NAME: Jessica Menard

PLACE OF BIRTH: Kingston, Ontario

YEAR OF BIRTH: 1986

EDUCATION: Frontenac Secondary School; Kingston, ON

2000 – 2004: High School Diploma

Mount Allison University; Sackville, NB

2004 – 2008: Bachelor of Science (Honours) degree

University of Windsor; Windsor, ON

2008 – 2010: Master's of Arts degree

University of Windsor; Windsor, Ontario

2010 – 2015: Doctorate of Philosophy degree