Assessing the Behavioral and Physiological Effects of a Reading Intervention on School-age Children with a Learning Disability in Reading

Badriah Basma, M.Ed.

Department of Education and Counselling Psychology

McGill University, Montréal, Canada

February 2024

A dissertation submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Human Development

Table of Contents

Abstract

The N400 is a brain response indicating semantic processing during language comprehension. It's sensitive to semantic incongruity in sentences. In individuals with reading disabilities (RD), N400 responses may be aberrant. Only one study, Hasko et al. (2014), has explored N400 changes following reading interventions in children with RD, but it lacked comprehensive intervention details and a randomized control design to evaluate the effect of the intervention. This limits knowledge on interventions inducing neural changes for RD. This thesis aims to implement a reading intervention, Set for Variability (SfV), and track the N400 to enhance reading abilities in children with RD. Set for Variability (SfV) is a reading strategy that focuses on building a repertoire of grapheme-phoneme variations together with tapping on lexical-semantic processing to build word reading ability. The thesis seeks to provide theoretical and empirical analysis to guide SfV development and implementation. Study 1 (Manuscript 1, Chapter 4) is a systematic review, and a meta-analysis which was conducted to explore whether there is a reliable difference in the N400 between readers with and without RD. The second objective of the study is to explore if there are moderating effects that have an impact on the N400 difference between these two populations. The meta-analysis indeed revealed, as hypothesized, that the N400 is aberrant in readers with RD among published studies. However, the aberrancy is impacted by several moderators. In the meta-analysis, the difference in N400 was most robust in the sentence incongruent task. The second study (Manuscript 2, Chapter 6) used an ERP paradigm to assess the effects of sentence incongruity on N400 where younger children below the age of 10 with and without RD read silently, to explore links more fully between N400 and reading acquisition and to explore group differences in SfV. The second objective of Manuscript 2 was to explore the correlations between the N400 and reading

measures, with a specific interest in the SfV measure. The results showed that readers without RD exhibited a typical negative peak of the N400 with incongruent sentences, while readers with RD did not have this N400 profile. The N400 was correlated with reading comprehension in typically developing readers. SfV was correlated with all reading measures. In readers with RD, SfV was correlated with N400 latency. Having established evidence of a difference in SfV and the N400 between the two populations and a correlation of N400 indices with SfV, the next and final step (Manuscript 3, Chapter 8) was twofold: to explore the effects of systematically teaching SfV to readers with RD within an experimental randomized control trial intervention study and to then explore changes in neural activity post-reading intervention. Thirty children were randomly assigned into two groups: an SfV variability group and an active control group that received Current Best Practices (CBP) (Savage et al., 2018). ERP and reading measures were taken before and after the reading intervention. The results showed, as hypothesized, that children in the SfV intervention group made significant improvement in SfV measure and irregular reading at post-test. In addition, the SfV intervention group also exhibited a significantly more negative N400 peak at post-test. This study is the first to demonstrate neural changes after a SfV reading intervention using EEG with readers with RD. In sum, the work included in this thesis supports the reliability of the N400 effect as an index for lexical-semantic processing within a cascading reading system. In addition, findings (Study 2) provide evidence that readers with and without a RD differ in terms of (i) their N400 profile when reading incongruent sentences, and (ii) the relationship between their N400 indices and reading ability (i.e., SfV). Finally, findings demonstrated that SfV intervention results in a significant increase in post-test reading ability on both SfV and other reading measures, and an associated change in N400 profile (i.e., more negative N400 peak). The research presented here used both behavioural and neurophysiological approaches to advance our knowledge of the neurophysiological correlates of reading difficulties and assess the neurological changes following well-designed reading interventions.

Resumé

La N400 est une réponse cérébrale signalant le traitement sémantique lors de la compréhension du langage. Cette réponse cérébrale est sensible à l'incongruité sémantique dans les phrases. Chez les individus présentant des troubles de la lecture (TL), les réponses N400 peuvent être aberrantes. Une seule étude, Hasko et al. (2014), a exploré les changements de la N400 à la suite d’interventions en lecture chez les enfants présentant des TL, mais cette étude manquait de détails compréhensifs sur l'intervention utilisée et de la mise en place d’un contrôle randomisée pour évaluer l'effet de l'intervention. Cela limite les connaissances sur les interventions induisant des changements neuronaux pour les TL. Cette thèse vise à mettre en œuvre une intervention en lecture, Set for Variability (SfV), et à suivre la N400 pour améliorer les capacités de lecture chez les enfants présentant des TL. Set for Variability (SfV) est une stratégie de lecture qui met l'accent sur la construction d'un répertoire de variations graphème-phonème ainsi que sur le traitement lexical-sémantique pour développer l’aptitude à la lecture de mots. La thèse cherche à fournir une analyse théorique et empirique pour guider le développement et la mise en œuvre du SfV. L'étude 1 (Manuscrit 1, Chapitre 4) est une revue systématique de littérature et une méta-analyse qui a été menée pour explorer s'il existe une différence fiable de la N400 entre les lecteurs avec et sans un TL. Le deuxième objectif de cette étude est d'explorer s'il existe des effets modérateurs ayant un impact sur la différence de la N400 entre ces deux populations. La méta-analyse a en effet révélé, comme prévu, que la N400 est aberrante chez les lecteurs avec un TL parmi les études publiées. Cependant, l'aberrance est impactée par plusieurs modérateurs. Dans la méta-analyse, la différence de la N400 était la plus robuste dans la tâche de phrases incongruentes. La deuxième étude (Manuscrit 2, Chapitre 6) a utilisé un paradigme ERP pour évaluer les effets de l'incongruité des phrases sur la N400 où les jeunes enfants de moins de 10

ans avec et sans un TL ont lu en silence, pour explorer plus pleinement les liens entre le N400 et l'acquisition de la lecture et pour explorer les différences de groupe du SfV. Le deuxième objectif du Manuscrit 2 était d'explorer les corrélations entre la N400 et les mesures de lecture, avec un intérêt spécifique pour la mesure du SfV. Les résultats ont montré que les lecteurs sans un TL présentaient un pic négatif typique du N400 avec des phrases incongruentes, tandis que les lecteurs avec un TL n'avaient pas ce profil de la N400. La N400 était corrélée à la compréhension en lecture chez les lecteurs ayant un développement neurotypique. Le SfV était corrélé à toutes les mesures de lecture. Chez les lecteurs avec un TL, le SfV était corrélé à la latence de la N400. Après avoir établi la preuve d'une différence du SfV et de la N400 entre les deux populations et d'une corrélation des indices de la N400 avec le SfV, la prochaine et dernière étape (Manuscrit 3, Chapitre 8) était double : explorer les effets de l'enseignement systématique du SfV aux lecteurs avec un TL dans une étude d'intervention expérimentale contrôlée et randomisées ensuite explorer les changements dans l'activité neurale après l'intervention en lecture. Trente participants ont été assignés de manière aléatoire à deux groupes : un groupe de variabilité SfV et un groupe de contrôle actif qui a reçu les meilleures pratiques actuelles (MPA) (Savage et al., 2018). Les mesures ERP et de lecture ont été prises avant et après l'intervention en lecture. Les résultats ont montré, comme prévu, que les enfants du groupe d'intervention SfV ont fait des progrès significatifs dans la mesure du SfV et de la lecture irrégulière au post-test. De plus, le groupe d'intervention SfV a également présenté un pic N400 significativement plus négatif au post-test. Cette étude est la première à démontrer des changements neuronaux après une intervention en lecture SfV en utilisant l'EEG avec des lecteurs ayant des TL. En résumé, le travail inclus dans cette thèse soutient la fiabilité de l'effet N400 comme indice du traitement lexical-sémantique dans un système de lecture en cascade. De plus, les résultats (Étude 2)

fournissent des preuves que les lecteurs avec et sans un TL diffèrent en termes de (i) leur profil N400 lors de la lecture de phrases incongruentes, et (ii) la relation entre leurs indices N400 et leur capacité de lecture (c'est-à-dire SfV). Enfin, les résultats ont démontré que l'intervention SfV entraîne une augmentation significative de l'aptitude à la lecture au post-test sur SfV et d'autres mesures de lecture, ainsi qu'un changement associé dans le profil N400 (c'est-à-dire un pic N400 plus négatif). La recherche présentée ici a utilisé à la fois des approches comportementales et neurophysiologiques pour faire progresser nos connaissances sur les corrélats neurophysiologiques des difficultés en lecture et évaluer les changements neurologiques à la suite d’interventions en lecture bien conçues.

Acknowledgments

To those familiar with the challenges encountered on my journey to bring this thesis to completion, my profound gratitude knows no bounds as I pen these words. The completion of this thesis stands as a testament to the invaluable support extended by each of you, without which this achievement would have remained out of reach.

Thank you to Dr. Armando Bertone for welcoming me to your lab. Thank you for your infinite warmth and support. I am grateful that not only did I have a supervisor but a dear friend too. Thank you for all the advice, your words of support, your trust in me, and your valuable feedback.

I want to express my deepest appreciation to Dr. Robert Savage, whose unwavering guidance and support have been indispensable throughout the journey of completing this thesis. Rob, my gratitude knows no bounds for your belief in me. When I first stepped into your office in 2014 with just an idea, little did I know that every bit of knowledge I possess today was taught by you. Thank you for having me in your academic realm. Your endless support, guidance, and mentoring helped me through countless long nights and weekends. Thank you for all the meetings—whether in the night or the early hours of the morning—providing feedback and meticulously reading my work to ensure I could progress. Your guidance has made a permanent mark on my academic journey, and for that, I am profoundly grateful.

I would like to thank Dr. Gigi Luk. You have been an endless source of knowledge and encouragement. You are a role model and an inspiration. Thank you for being there when I lost my compass to move forward, for all your valuable feedback, contribution to my thesis, and your financial support. Thank you for all the meetings to make sure that I stay on track.

To my three precious gems, Kirsty, Julie, and Stephanie. I have no words. All three of you made me want to wake up in the morning and come to the lab. Thank you for the forever friendship that I will always be grateful for.

To my friend Dr. Daphne Silvestre, I want to express my heartfelt gratitude for being my go-to person every time I panicked. Thank you for your unwavering patience during moments of constant complaints, and for providing endless support, listening to me, and offering valuable advice. I feel truly fortunate to have had you in the lab during those times.

To Charlotte McCaffrey and Marieve Isabel. I do not think that both of you know how your constant support made me move forward. Charlotte, thank you for always responding to my emails, for always clarifying things to me at times when I felt lost. I hope you know how important you are too many of us. Thank you Marieve for your endless support in every writing session. Thank you for your optimism and always being there. I am grateful and lucky to have had you in all those writing sessions.

I would like to express my thanks to my dearest friends Mhamad Nasreddine, Dr. Alexia Ostrolenk, Dr. Ahlam Rahal, Flavio Muharara, Nesma Bashi, Dr. Jesse Michaels, Dr. Emma Clark-Skelton, Samantha Scholes, Sabrina Censi, Catherine Archambault, and Taryn Perilmeter. Thank you for being there when I needed it the most.

To my research assistants Giuliano, Laury-Ann, Joann, Olivia, Julia, Catherine, Glory, Miriam, Luna, Alexandra, Shira, Audrey, and Alexa. You are the best research assistant that one could ask for.

I want to express my heartfelt gratitude to my dearest friend, Pierre, for always standing by my side. Thank you for taking the time to read my work when you didn't have to. Thank you for reading all my French abstracts. Thank you for listening to my academic work all the time.

Your willingness to listen to both my joyous and sorrowful moments, to my complaints, and even my gossip means more to me than words can convey.

I would like to express my infinite gratitude to my sisters Ayat and Layla. I am lucky to have you in my journey and life. Thank you both for your unwavering support and for always being there whenever I needed you. Loulou, thank you for being my inspiration. Without you, Khalil, Angela, and Julia, I could not have done it. Ayat, I cannot thank you enough for pulling me up every time I doubted myself. You beside me is the world to me. I am only stronger with you by my side. I would also like to thank my brother Hadi and my mama. Hadi, thank you for your constant support and always being there for me. Thank you, mama, for all your support, for keeping me in your prayers and for convincing everyone that I am a medical doctor. I have learnt so much about physiological diagnosis and blood tests than I ever thought I would know.

To my beloved father, this is for you. I love you baba, and I wish you were here to witness this moment. You have been and will always be my driving force, inspiring me to reach new heights. I hope, wherever you are, you are proud of the person I've become. I love you, again and again and always.

To my cat, Set El Hessen: You cannot imagine the amount of happiness you brought into my life when you entered it. Thank you for being my cuddle buddy through all the ups and downs of this journey.

To Issam. I was never prepared for this moment. Never in my lifetime, I thought I would complete my thesis while crying for your loss all the time. I had envisioned us celebrating this accomplishment together. It breaks my heart knowing that I won't see you again. Those who knew you, knew your dedication to your work. Your photography and work ethic inspired me and many people. I dedicate all this to you and all the cherished memories we had together.

Contribution to Original Knowledge

The content of this dissertation contributes to the current literature on the association between neurophysiological processes, reading ability, and related language abilities, and to neurophysiological changes that take place following a successful reading intervention for struggling word readers. This research was conducted with students with reading difficulties and severe reading disabilities, as well as typically developing students with no reading disabilities. Working with these two populations furthered our understanding of how to identify the neurophysiological and behavioral differences between these two populations as well as using targeted reading interventions that support children with reading disabilities. The findings of Study 1 expanded on extensive literature that demonstrates that there is a neurophysiological difference in event-related potentials (ERPs), specifically the N400, between individuals with and without reading disabilities across age groups. The N400 is an ERP component that is reflected by a negative peak that takes place at a time window 300-500 ms. Two components are of interest in this thesis: (i) the N400 amplitude, which is the change of the EEG signal towards a negative peak, (ii) The N400 latency, which is the time delay in which the negative peak voltage happens. Study 1 is the first systematic review and meta-analysis that provided evidence of such neurophysiological differences and the key moderators that affect such differences. The results of study 1 provided a rationale for Study 2, which explored the neurophysiological differences between typically developing children and children with reading disabilities below the age of ten years. In addition, this study explored the correlation differences of reading and ERP measures between these two populations. The rationale proposed was that children with reading difficulties likely have different profiles of correlations between reading and neurophysiological measures compared to typically developing children. Additionally, the study explored a specific reading-related measure, known as Set-for-Variability (SfV), in terms of typically developing (TD) versus reading disabilities (RD) group differences and its correlation with the N400. The rationale proposed that differences in SfV tap into differences in the N400 when compared between the two populations. This is the first study to study the relationship between SfV and N400 in TD versus RD samples. The findings of Study 2 revealed that indeed there is a difference between TD versus RD groups in N400 and SfV. Study 2 also revealed different by-sub-sample correlation patterns with reading measures. While TD children exhibited a correlation between reading comprehension and the N400 amplitude, SfV was correlated with N400 latency in children with RD, different patterns possibly reflecting differences in word reading expertise. The results of Study 2 informed Study 3 that used an SfV-based reading intervention within an experimental (randomised control trial) intervention design to assess whether systematic SfV instruction improves children's reading performance and produces N400 changes in amplitude and/or latency. Both claims were evidenced by the results of Study 3. Study 3 also for the first time in the literature produced evidence of far generalisation of effects to untaught exception words and untaught SfV task items in RD samples. More generally, study 3 is the first study to use this experimental intervention approach and track neural changes in N400. Overall, this research contributes new findings to the literature on our understanding of neurophysiological differences and improvement in reading and brain-behavior associations for students with RD. There are no studies so far that have directly assessed correlations between reading measures, SfV and the N400, and there are no studies yet that have directly assessed the usefulness of SfV as a reading intervention in RD samples that aims at improving both reading and related neural measures.

Contributions of Authors

I, Badriah Basma, am the primary author of this dissertation and have conceptualized, collected, analyzed, and written this dissertation entirely. Dr. Armando Bertone, who served as my primary supervisor provided me with all the EEG equipment, access to technical support and workshops for EEG study setup and data analysis. He provided me with lab space and access to all its human and physical resources. Dr. Bertone also contributed to the development of the study research design and provided me with constructive feedback on my three manuscripts and thesis. He also supported me financially throughout my doctoral studies whenever possible. Dr. Robert Savage, my co-supervisor helped me tremendously with guidance and continuous feedback on my proposal, the projects, and the three manuscripts. Dr. Savage and I have shared interest on reading interventions. Given this interest, Dr. Savage facilitated my training on the reading intervention with the London, UK teacher cohort. He provided extensive feedback on the study design, formulating my hypothesis and research questions, choosing the right analysis, writing ethics application for two school boards (English Montreal School Board and Sir Wilfrid Laurier School Board), contacting every possible person to facilitate my participant recruitment. Dr. Savage also helped me in delivering training to all my research assistants. Dr. Gigi Luk, my committee member, and an interim supervisor played a critical role in the supervision of the two studies and provided oversight for all my projects. Dr. Luk had weekly meetings with me to sure my progress in terms of ethics application, recruitment, and data-management. Dr. Luk paid for all my gift cards and supported me financially. Dr. Luk continuously gave me feedback on all my manuscripts and thesis.

List of Tables and Figures

Literature Review

Manuscript 1

Manuscript 2

Manuscript 3

List of Appendices

List of Abbreviations

ADHD	attention deficit hyperactivity disorder
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
ASD	autism spectrum disorder
BW	blending words
CBP	current best practices
CC2	Castle and Coltheart Test 2
CTOPP-2	Comprehensive Test of Phonological Processing - 2nd Edition
CVC	consonant vowel consonant
CVCC	consonant vowel consonant-consonant
DM	direct mapping
DMSfV	direct mapping and Set for Variability
DRC-model	Dual Route Connectionist model
DSM 5	Diagnostic and Statistical Manual of Mental Disorders - 5th edition
ERP	event-related potentials
EEG	electro-cephalogram
EOG	electro-oculogram
fMRI	functional magnetic resonance imaging
GPC	grapheme-phoneme correspondence
ICA	individual component analysis
LD:	learning disability
LQH	Lexical Quality hypothesis
IR	irregular word reading
MANCOVA	Multiple Analysis of Covariance
NW	non word reading
PDP	parallel distributed process
PI	phoneme isolation
PPVT	Peabody Picture Vocabulary Test - 5th Edition (PPVT™-5)
RAN	rapid automatized naming
RCT	randomized control trial
RD	reading disability
ROI	region of interest
RT	reaction time
SfV	Set for Variability
SVR	Simple View of Reading
TD	typically developing
WOE	Weight of Evidence
WR	word reading
WIAT-III	Weschler Individual Achievement Test - Third Edition (WIAT™-3)

Chapter 1. Introduction

A reading learning disability is characterized by slow and difficult word reading, frequent spelling and punctuation mistakes, and/or the need to repeatedly read text to comprehend it (Scanlon, 2013). Learning disability (LD) in reading is one of the most common neurobiological diagnoses in school-aged children (Barbiero et al., 2012). Statistics Canada reported that 3.2% of children in Canada have LD (Statistics Canada, 2012), making LD the country's most prevalent type of childhood condition. Electroencephalography (EEG) is one of the most common tools to inquire about brain information processing and neural changes (Bednar et al., 2018). To understand how the brain processes language in real time, Event-Related Potentials (ERPs) studies have been commonly used to assess the neural underpinnings of reading disability (RD), especially in children. Of particular interest, the N400 is a negative ERP waveform peaking at 400 milliseconds sensitive to the lexico-semantic aspect of language. The N400 represents an important language-relevant measure that can potentially be used to investigate the neural basis of reading comprehension acquisition, a process critical to reading development in school-age populations. The N400 assesses lexical-semantic processing, showing activation for example, when a reader is introduced to a sentence with a pseudo word or a semantically incorrect word as in "*the pizza was too hot to cry*". The N400 is sensitive in TD readers, indicating they understand semantic incongruency. These brain-based explanations describe the neural processes underlying reading in typically developing readers (TD). Aberrant neural processes in the N400 explain reading difficulties, including reading comprehension deficits in individuals with reading difficulties (RD). Several studies demonstrate the potential relevance of using neuroimaging techniques in identifying a reading disability (Abboud & Cohen, 2019). Researchers have also demonstrated that neural changes occur after a successful reading intervention (Barquero et al., 2014; Partanen et al., 2019; Simos et al., 2007). However, it remains unknown whether a reading intervention modulates reading-relevant brain activity, such as the N400. Only one study has used reading intervention and EEG to track neural changes in children with LD in reading (Hasko et al., 2014). Hence, little is known about reading interventions that induce neural changes in young children with RD that become more like age-matched TD readers. This doctoral dissertation aims to explore N400- reading associations to implement a specific reading intervention in children and track changes in brain activity after intervention. While it is established that the N400 measures lexicosemantic processes associated with reading, the extent to which N400 amplitude or latency indexes SfV remains unknown. This study hypothesizes that N400 amplitude and latency can serve as indices of SfV. Additionally, it is hypothesized that both processes (N400 and SfV) are susceptible to change through reading interventions guided by our theoretical framework. The thesis elaborates and tests these claims. The results of this thesis will inform theory about brain-behaviour links for N400 and potentially be invaluable in improving the understanding of reading disabilities in younger children from a neurophysiological perspective, identifying better-targeted reading interventions, and allocating resources to health and educational services. Behavioral assessments provide valuable insights into the observable outcomes and patterns of reading difficulties, such as poor reading fluency, comprehension deficits, and frequent errors in word recognition and spelling. However, these assessments alone cannot fully capture the underlying neural mechanisms that contribute to these difficulties. Electrophysiological techniques, such as event-related potentials (ERPs), offer a unique advantage by providing real-time data on the brain's electrical activity during reading tasks. Specifically, the N400 component is known to reflect the brain's processing of semantic information and can highlight differences in how children with reading disabilities process language compared to typically developing readers. By examining the amplitude and latency of the N400, we can gain insights into the timing and efficiency of semantic processing, which is not easily detectable through behavior alone. For instance, if electrophysiological data indicate that a child has delayed or atypical semantic processing, interventions can be tailored to focus on enhancing semantic integration and vocabulary development. Additionally, electrophysiological measures can be used to monitor the effectiveness of interventions over time, providing a more precise measure of progress than behavioral measures alone.

The Rationale for the Doctoral Thesis

Research suggests that RD has a neurophysiological basis involving mechanisms operating in the brain, suggesting reduced neural activation associated with RDs (Caylak, 2009; Hasko et al. 2014; Shaywitz et al. 2017). In addition, the purported neural substrate of one specific aspect of reading, namely lexical-semantic processing may be operationalized using neurophysiological measures and ERP tasks such as those measuring N400. The purpose of this thesis is to use this brain behavior link to assess the effects of a targeted reading intervention on reading behaviour and neural activity underlying reading abilities in children with RD. Only one study (Hasko et al. 2014) has investigated associations between behavioural measures of reading and the N400 and the impact of educational interventions. Here, children received one of two interventions twice a week for 6 months in written German. The first intervention focused on teaching children's orthographic rules for long (e.g., /i/ in bee), and short vowel sounds (/a/ in apple or cat). The second intervention focused on teaching children grapheme-phoneme correspondence (GPC) rules, which are letter-sound associations. For example, the letter "a" represents the sound /a/ as in apple. Hasko et al. did not explain whether the N400 was associated with one intervention (teaching GPCs) over the other (teaching orthographic knowledge). Additionally, the study was delivered in German, a transparent spelling system where the GPC system is a reliable guide to word pronunciation and spelling, and where semantic resources may feature less as there is little or no ambiguity to resolve while reading individual words. This is unlike the English language that is opaque and includes many irregular spelling patterns and exceptionalities (Seymour et al., 2003). I propose using the N400 measure within a targeted reading intervention based on the promising results of studies teaching Set-for-Variability (SFV), (e.g., Dyson et al., 2017; Savage et al., 2018). SfV may represent one step beyond simply teaching GPC rules and exception words by sight as it instills in children a generative strategy to read words that breach the most common phonic rules by teaching them how to search for the correct pronunciation of such words. The intervention is appropriate for exploring lexical-semantic problems through N400 change, as it teaches children to use phonological and vocabulary resources to read. Together then, the doctoral dissertation will examine the impact of SfV intervention on lexical-semantic reading skills of children with reading disabilities, measured using the N400.

The primary objective of this doctoral dissertation is to examine the effect of an SfV intervention on the N400 amplitude in school-age children with RD. To achieve this objective, three separate studies, written as manuscripts, will comprise the author's doctoral dissertation for the Doctor of Philosophy in Educational Psychology. Chapter 1 presents a comprehensive literature review describing theories of reading acquisition, theories of reading disabilities, theories of the N400, the relationship between the N400 and reading disability, and finally, the possibility of the N400 being improved using the reading intervention SfV. The subsequent chapters present three empirical studies. All three articles included in the dissertation are independent manuscripts submitted or accepted for publication in peer-reviewed academic journals. The manuscripts are written and presented according to guidelines outlined by the Faculty of Graduate and Postdoctoral Studies at McGill University. A bridging section that links each of these articles is included between each manuscript.

The empirical studies comprise of three components: (1) a systematic review and a meta-analysis aimed at contextualizing the research questions of the two proposed empirical studies; (2) A correlation study to explore the relationship between reading behavioral measures (specifically SfV) and N400 in children with and without RD; (3) An intervention study to explore the effect of the SfV reading intervention on the N400 amplitude in school-age children. Manuscript 1 (Chapter 4) is a systematic review and a meta-analysis to assess whether the N400, a lexical-semantic ERP measure, has the same sensitivity in neurotypical versus readers with RD in reading. The study assesses whether individuals with reading disabilities have a lower N400 peak and amplitude than neurotypical readers. This will be achieved by (1) synthesizing the results of well-executed studies using the N400 in TD and RD readers and (2) by assessing the difference in mean (effect size) between TD and RD, as well as testing theorized moderators of variability in effect size. The meta-analysis results will give us a better understanding from a quantitative perspective of the N400 in TD and RD in reading. Manuscript 2 (Chapter 6) examines the semantic processes in children with RD in reading and neurotypical readers along with behavioral measures. This study also assesses whether negative N400 amplitude correlates with SfV, word reading, and comprehension in children with RD compared to typical readers. Manuscript 3 (Chapter 8) aims to use the N400 ERP in semantic incongruity tasks in a randomized controlled study to inform our understanding of the specific comprehension impairments by providing estimates of processing abilities before and after a reading intervention. The goal will be achieved by implementing a targeted reading intervention (SfV) and measuring the accuracy and speed of a semantic incongruity task. Based on theories of SFV (Tunmer & Chapman, 2012) we predict an SfV intervention to improve regular and especially exception word reading and SfV measures. From the results of Hasko et al. (2014) study, we hypothesize changes in the N400 amplitude and latency after a reading intervention and further predict a specific link with SfV treatment over control intervention conditions. Post-intervention assessment will take place to explore if the intervention influences both reading, specifically SfV, and the N400. Chapter 9 is a discussion of the main findings across the three manuscripts. The contribution of each manuscript will be highlighted and will include an extensive discussion of limitations and implications for future research based on this current work.

Chapter 2. Literature Review

Reading Acquisition

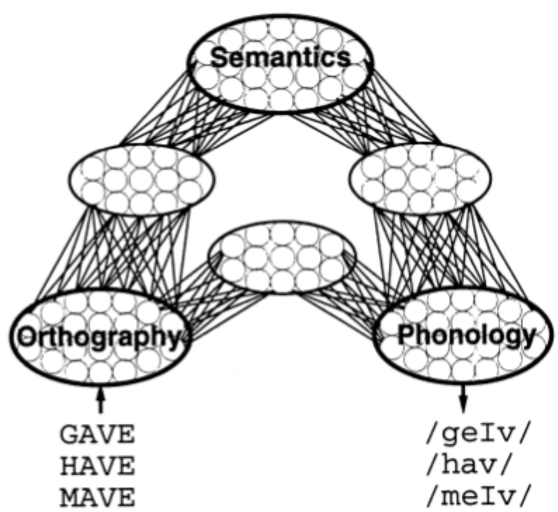
The written English language is an alphabetic system that has 26 letters. One letter or several letters can be used to form *graphemes* that are the print representations of the 44 *phonemes*, the smallest units of speech sound (Chen & Savage, 2014) that are used in spoken English. This representation is thus complex: In "night," for example, three letters "i- g -h" represent the phoneme /ai/ whereas only one grapheme represents the sound /n/ and /t/. These 44 phonemes are either consonant phonemes: */b/ bus*, */d/ dog*, */m/ man*; short vowel phonemes: /a,(æ)/ in c*a*t, /e,(e)/ in p*e*g, /i,(I)/ in p*i*n, /o,(ɒ)/ in h*o*t, /u,(ʌ)/ in b*u*s; or long vowel phonemes as in *brain* and *tree* diagraphs: ch- church, sh- shoe, -ph- phone, or in single vowel letters such as 'go', "me" (Ott, 2008). The 44 phonemes also include several schwa sounds. Schwa sounds are unstressed vowels. In written language contexts, we see *schwa* sounds written quite variably as a short vowel sound -*u (ʌ*) or -*uh(*ɒ), for example, ***a**gain* and *c**o**mmitment* (Weber, 2018). Skilled readers can transform graphemes (letters in print) into phonemes (speech sounds) with efficiency. Efficiency in this context involves rapid and correct text processing with minimal cognitive load, allowing more resources to be allocated to understanding and interpreting the material. Efficient reading implies that the reader can process and comprehend text swiftly and accurately, indicating a high degree of both automaticity and accuracy (LaBerge & Samuels, 1974. This process relies on a mental lexicon, which functions as a mental dictionary. It enables the listener to identify sounds and understand the meanings of words (Marslen-Wilson et al., 1994). Share (1995) explained that this Item-based word development suggests that word recognition development depends on the frequency of exposure to specific words and successful item identification. High-frequency words are recognized visually, while low-frequency words rely on phonological recoding. As reading proficiency grows, phonological recoding evolves to incorporate more complex orthographic rules, making the process appear more lexical, and thus the mental lexicon. Both phonological and orthographic components independently contribute to fluent word recognition, with phonological recoding being primary. Orthographic processing builds on phonological decoding, allowing for rapid acquisition of word-specific knowledge. Share (1995) argues that such ability to decode words builds the mental lexicon. He states that "each successful decoding encounter with an unfamiliar word provides an opportunity to acquire the word-specific orthographic information that is the foundation of skilled word recognition' (p. 155).

Muter et al. (2004) argue that readers need phonological awareness (whole sounds of words) and phonemic awareness (the smallest individual sounds in a word) to distinguish and manipulate phoneme sounds essential for assembling word pronunciations in reading. Consistent with this view, Cardoso-Martins et al. (2011) found that phonological awareness helps children learn grapheme-phoneme rules more easily, specifically middle letter sounds, rhyme, and alliteration. Yeung and Savage (2020) found that practicing 'direct mapping' (text-based practice of recently taught GPC rules in reading books) improved word reading and comprehension, but that pre-intervention phoneme awareness strongly moderated GPC learning. Phonological skills play a crucial role in the development of direct mapping in reading. As readers become more proficient in phonological processing, they are better able to identify and remember these word patterns and structures, thereby facilitating the process of direct mapping. Direct mapping, in turn, allows for the rapid recognition of familiar words, while phonological skills continue to support the initial learning and ongoing refinement of these word forms. For example, understanding that the word "cat" follows a common phonetic pattern helps readers quickly map this word into memory, illustrating how phonological skills and direct mapping work together as complementary processes in reading. Similarly, researchers have shown that readers must have letter knowledge and phonemic awareness to read fluently (Bowey, 2005; Melby-Levarg, et al., 2012). Hulme et al. (2012) ran a mediation analysis on a large-scale randomised control trial reading intervention. The intervention taught children letter-sound knowledge and phonemic awareness. Both skills mediated the improvement in children's word reading and spelling skills five months post-intervention. This indicates that both knowledge of grapheme-phoneme correspondence and phonemic awareness have an impact on children learning to read.

However, the English language has inconsistencies in GPC rules (Georgiou et al., 2008). So how do skilled readers learn words to recognise words that are exceptions to phonic rules? What is it that the children are learning while reading? One potential answer to the first question may be found in some connectionist models such as those of Harm & Seidenberg (2004); Plaut et al. (1996); and Seidenberg & McClelland, (1989). Seidenberg & McClelland (1989) explain that reading acquisition takes place through a network training of a set of orthographic inputs connected to the phonological network via *hidden uni*ts. Hidden units are representations that act as mediators of activity between input and output. This model does not use explicit GPC rules, but rather backpropagation that produces the correct phonological output to a given orthographic input. Backpropagation suggests that through exposure to words and trial and error, an individual learns to read all words including those that do not follow primary GPC rules. Input, in this case, is print (grapheme) and output is speech (phoneme). While backpropagation helps us understand how to read words with regular spelling patterns, alone it does not explain how we read exception words. Here, semantics might be needed to read such words. Gonnerman et al., (2007), elucidate that reading “involves mapping between sound, spelling, and meaning” (p 327). These mappings are connected through networks in the form of processing units. Each unit contains connected and overlapping orthographical, phonological, and semantic groups (Figure 1). The overlap between these three groups helps the reader learn complex words' consistent but distinctive characteristics. In other words, individuals learn to read through repeated exposure to printed words by comparing errors against patterns of sounds and adjusting, which result in a neural net learning system (Savage, 2019). Zeigler et al. (2020) have proposed such a triangle model as the instantiation of the self-teaching hypothesis proposed by Share (1995), (For review, Ziegler et al., 2014; Zeigler et al., 2020, see Figure 2).

Figure 1

Connectionist Framework for Reading reprinted from Seidenberg and McClelland (1989).



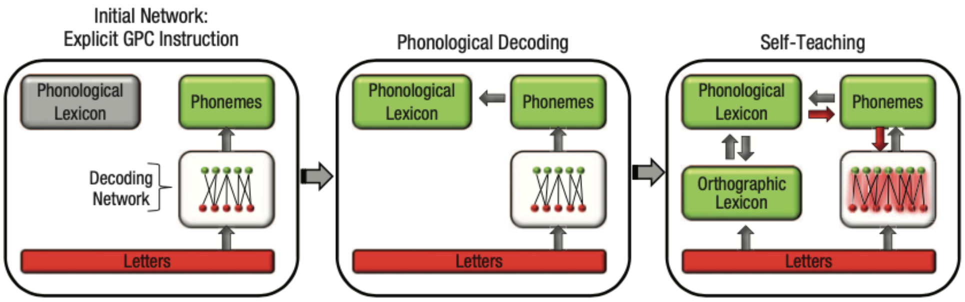
The arrows are the connection between the three groups of units. Knowledge of words depends on the weights of connection between the three groups of units. By weights, we mean the result of exposure to distributed connection strength.

A second approach to this problem of inconsistencies in grapheme-phoneme correspondence is to suggest that reading aloud involves two routes to pronunciation. This is most clearly represented in the Dual-route Connectionist (DRC) model of reading aloud

(Coltheart, 2006). The DRC theory states that a reader reads a whole word from print to speech through either a lexical or non-lexical route. A non-lexical route uses GPC rules to assemble word pronunciations similar to what explained earlier with "backpropagation" and self-teaching. The reader, in this case, reads words by applying knowledge of GPC rules and then blends the products. A lexical route process assumes more direct access through print, typically for familiar words already stored in the lexicon among skilled readers. This route has been implemented as a 'Cascade' connectionist model that represents features, letters, and words, that are processed in parallel in word reading (Coltheart, 2006). Coltheart (2006) explains that knowledge about the printed word is stored in three separate systems: the orthographic lexicon (spelling), the phonological lexicon (pronunciation), or the semantic lexicon (meaning). Regular words can be read through a non-lexical route. Irregular words, however, can only be read through the lexical route because these words do not obey the most frequent GPC rules (Figure 2). For example, regular words such as *beat* and *gain* can be read by applying the grapheme-phoneme rule of long vowel sounds. Yet, the reader cannot read *bear* and *said* by applying the non-lexical route because the reader will mistakenly rhyme these words with *beat* and *gain* (Coltheart, 2006; Coltheart et al., 1993; Coltheart & Rastle, 1994; Grainger & Zeigler, 2011; Mousikou et al., 2010; Pritchard et al., 2018).

Figure 2

Phonological decoding and self-teaching mechanisms in the context of the Connectionist model adapted from Ziegler et al. (2014)



Triangle connectionist approaches can be distinguished from the Coltheart DRC model in that the former parallel processing networks of neuron-like processing units learn by adjusting the weights[1] of the connections between them, and where no 'word' nodes[2] exist in the system. There are therefore no 'rules' to adhere to but rather a unique system in which items coexist based on representations that reflect a certain *consistency* in the mapping of different items. Consistency here is thus different from regularity. Regularity is the most common pronunciation of a grapheme. For example, the most common pronunciation of *ea* is /i/ (long vowel sound e) as in eat, beat, seat, and heat, which makes these words regular but words such as bear, and swear as irregular. Consistency reflects the shared pronunciation of grapheme and graphemic units across words e.g. -ink in pink, link, shrink, think…etc. The same system represents inconsistency in items such as -int in mint and pint. The model is adaptive by adjusting the weight on connections between units accordingly. One issue concerns the real-world plausibility of weights of inputs in parallel distributed process (PDP)[3] models of reading. Larger weights only develop through learning opportunities and regular system-wide feedback (as in backpropagation of error algorithms) across 'epochs'[4] of learning. Such extensive feedback might not be available for the child to learn at school. On the other hand, the DRC models of reading aloud via two [1]routes: the lexical route of three separate systems to read irregular words (orthographic,

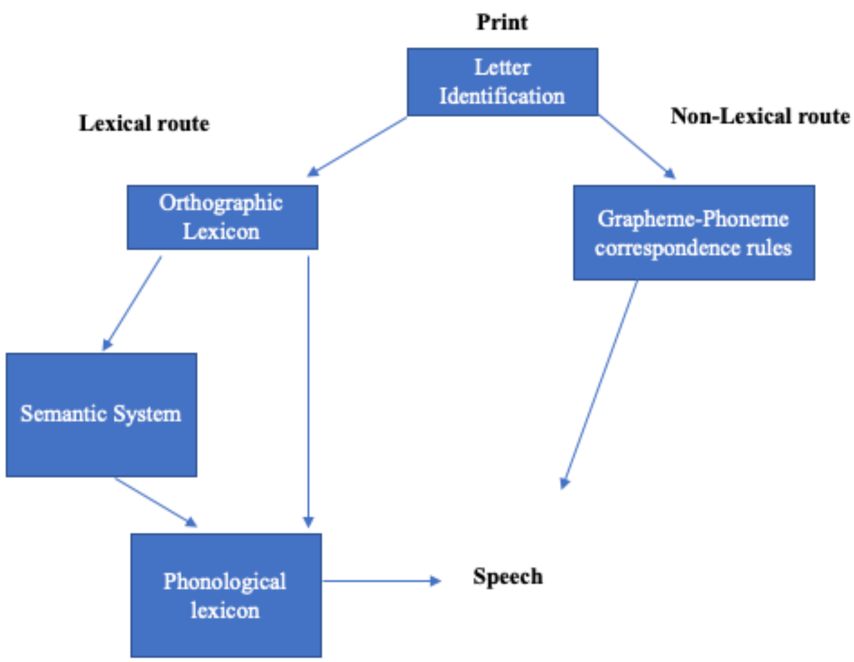
---

[1]weights: the force exerted on the connection between processing units.

phonological & semantic systems) and a non-lexical route to read words that apply GPC rules do not explain how the reading of polysyllabic words and exception word learning takes place. The DRC model also fails to explain the role of semantics in word reading in detail.

Figure 3

The Dual Route of Reading adapted from Coltheart (2006)



---

[2]nodes: point of location in a network
[3]PDP: is a framework that generates models of how the brain might perform the same task.
[4]epoch: one complete cycle of data training.

Reading Comprehension

Skilled word reading is essential to understanding the meaning of a text (Ecalle et al., 2013). Word reading, as an integral part of skilled reading, is necessary but not sufficient for developing the cognitive resources necessary to understand a written text, i.e., reading comprehension (García & Cain, 2014). Reading comprehension has been defined as "learning to understand writing as well as one understands spoken language" (Perfetti et al., 2005, p. 227). Then comprehension becomes more complex because other factors such as morphology, syntax, proposition, context, and background knowledge become required to fully comprehend a text (For review, see Landi et al, 2012; Nuss, 2018; Snow, 2002). Nuss (2018) argues that a young reader focuses first on word reading, later connecting the words, and then connecting the meaning of the words before comprehending the full sentence and the full text. Researchers agree that reading comprehension is a very complex skill, requiring two main processes; (i) word reading: which requires phonological and phonemic awareness, knowledge of GPCs, understanding of print concepts, semantic knowledge and executive controls, such as attention and verbal working memory (ii) linguistic comprehension: understanding the meaning of a text which requires vocabulary, which also requires understanding of language structure such as semantics and syntax (Aboud et al. 2016; Gough et al., 2013; Gough & Tunmer, 1986; Hudson et al., 2016). Gough and Tunmer (1986) argue that word recognition and linguistic comprehension are the main critical processes needed for reading comprehension. In other words, these are the foundational skills of reading and literacy and the definition of the Simple view of Reading (SVR) developed by Gough and Tunmer (1986). The SVR proposes that decoding and linguistic comprehension are two necessary skills and that neither alone is sufficient, to achieve successful reading comprehension. If the reader is poor in any of these skills, it would lead to poor reading comprehension (Hoover & Gough, 1990).

Other researchers claim that successful reading comprehension is not based only on successful word reading and linguistic comprehension. Reading comprehension requires an understanding of a situation. The "Situational Model" (Perfetti, 2005, P 228) is the mental representation of a written text (Dijk & Kintsch, 1983). The situational model means what the reader might understand after reading a sentence. Consider the following sentence: "Cathy was riding her bike in the park, dark clouds began to gather, and it started to storm" (Perfetti & Stafura, 2014; P 27). According to the "Situation Model," the reader will build a scheme based on the situation that there is a storm based on: *Cathy riding her bike*, *clouds getting dark* and then the event *storm*. Once the reader clears this situation that there will be a storm, the following sentence is presented in the text '*The rain ruined her sweater*". According to the situational model, the rain would be understood via the previous sentence: *dark clouds* and *storm*. However, if the reader encounters this sentence: *While Cathy was riding her bike in the park, dark clouds began to gather. The rain ruined her beautiful sweater."* In this text, there is no preceding situational referent for "rain" (Perfetti and Stafura, (2014). As Djalali (2008) explains, to understand this prose, a bridging inference is required. Bridging inference is a "specific type of inference in which the objects or events referred to in sentences are connected in a meaningful way so as to create coherent discourse." (p.1). For a bridging inference to take place, however, Perfetti explains that smooth access to lexical representations of words is necessary. This is known as the Lexical Quality Hypothesis (LQH). The lexical quality hypothesis posits that for successful reading comprehension to occur, orthographical, phonological, and meaning must be retrieved accurately and without effort (Perfetti, 2007;

Perfetti & Hart, 2002). Poor access to the lexical representation, as Perfetti (1985) explains, would ultimately result in poor reading comprehension (Richter et al., 2013). According to the LQH, there is an interplay between the lexical processes (word), syntactic processes (grammar), and written text (comprehension). The interplay among these three processes includes a range of inferential routes, such as decoding the word, analyzing the sentence, and building a semantic map (word meaning and word relation) to understand a written text as in the sample text above. Reading comprehension is a complex task because comprehension acquisition depends on these three components: the ability to infer, the ability to identify comprehension errors, and the ability to identify text structure (Landi et al., 2013). Perfetti's (2007) *LQH* supports the critical role of lexical semantic processing and retrieval in reading comprehension.

In summary, "the ultimate goal of reading, beyond its acquisition stage, is typically comprehension of a connected text" (Landi et al., 2013, p. 146). Skilled reading requires integrating phonological, orthographic, and semantic information to decode a written text (Lam et al., 2017). Lexical and semantic processing play an additional and critical role in developing reading comprehension skills (Nobre & de Salles, 2016). It is also noteworthy to comment that decoding, word reading, and word knowledge are crucial targets for intervention towards readers who face difficulties in reading, (Braze et al., 2007). These are thus discussed in detail in the following section.

Reading Disability.

Currently, there is no separate category of *Dyslexia* in the DSM-5 (APA, 2013). Rather dyslexia is included under the category of a specific learning disorder with a specification of impaired reading in either reading fluency, spelling, and/or reading comprehension. Therefore, the DSM-5 (APA, 2013) classifies *Dyslexia* as a reading disability under broader learning disability categories. The DSM-5, however, is a diagnostic tool that is based on theory, but it does not explain the theoretical foundations of reading disabilities.

The most common hypothesis for reading disabilities is the phonological deficit theory, which is difficulty in the representation and retrieval of letter-sound correspondence (Caylak, 2010)., Wolf and Bowers (1999), however, claimed that RD involves deficit beyond phonological awareness. Dyslexia could be a result of both phonological awareness and rapid naming deficits. Rapid automatized naming (RAN) tasks involve naming multiple letters, numbers, objects, or colours in rows as fast as possible (Wolf & Denckla, 2005). The reader, for example, reads a limited number of specific items (book, chair, dog, hand, star) repeated across rows, as fast and accurately as possible. Consistent with this view, Schatschneider et al (2002) in their longitudinal study tested 945 students from kindergarten to Grade 2 on phonological awareness and naming speed and showed a high correlation between rapid naming and phonological awareness. Participants with the poorest performance showed deficits in both naming speed and phonological awareness - the double deficit hypothesis. Consequently, children with double deficits in both phonological awareness and naming speed are more likely to have severe reading difficulties than those with phonological deficits only. Torppa et al. (2013) followed 1006 kindergarten students longitudinally. Students with difficulties in both RAN and phonological awareness, and students with difficulties in either RAN or phonological awareness were tested on measures of phonological awareness, RAN, word reading, vocabulary, and letter knowledge. Their results showed that phonological awareness deficits predicted spelling accuracy and reading fluency deficits. The authors also found that RAN predicted reading fluency. However, the group with a double deficit in RAN and phonemic awareness had larger and more generalized difficulties in reading and spelling. The authors concluded that the children with the greatest challenges are those with deficits in rapid naming and phonological awareness. Irrespective of the theory, these behavioral indicators suggest that a reading disability could result from a double deficit disorder and not only due to a phonological awareness deficit.

Beyond decoding it is quite possible to identify groups of children who exhibit average decoding skills (as evidenced by performance on a test of pseudoword decoding) but who show difficulties in reading and wider language skills (Griffith & Stuart, 2013; Savage et al., 2023). Some of these children have been labelled 'surface dyslexics' and their difficulties are attributed to problems in lexical and lexical-semantic processing situated in DRC accounts (e.g., Griffith & Stuart, 2013). It is also quite possible though that word reading difficulties evident for some average decoders reflect vocabulary and morphological difficulties (e.g. Seymour, 1997; Savage et al., 2023) which are perhaps less readily explained within DRC models.

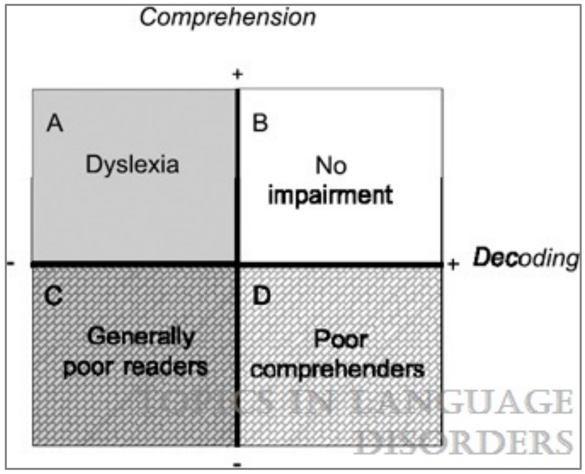
Another perspective on reading disability focuses on multiple factors and not just those related to phonological, lexical-semantic, and RAN deficits alone. As explained earlier the LQH posits that high-quality representation of words is a necessary feature of reading fluency. Failure in the representation of orthography, phonology, or semantics would result in difficulty in reading comprehension. The focus of LQH is on the representation (spelling) of words. An alternative view is offered by Zeigler et al (2014, 2020). Their theory focuses on disability in the acquisition of words in connected systems, rather than the representation of words. Zeigler et al. (2014) argue the specific role of vocabulary, rather than semantics, in general, in reading disabilities. In their 2014 paper, Ziegler et al. argue that poor access to vocabulary, and its influence also construed within the triangle model, could lead to phonological deficits and dyslexia. Zeigler et al. (2020) modelled their theory by creating a multi-component computational model. The model included five main components, letters, phonemes, a phonological lexicon, an orthographic lexicon, and a decoding network. The authors then simulated a deficit by switching off each one of the components. The authors then compared this multi-deficit computational model to alternative models: a phonological deficit model, a visual deficit model, and a global noise model that assumes an overall low level of performance. The multi-deficit model provided the best depiction of reading behaviour performance among all. The authors deduced that dyslexia is a result of a multi-deficit model and not due to a single trajectory of either an orthographic or phonological deficiency but could also be a result of difficulties in vocabulary.

Reading Disability and Reading Comprehension.

According to the DSM-5, reading comprehension difficulty is difficulty in understanding the meaning of a text). Reading comprehension difficulties can exist regardless of whether the word reading level is intact (Yuill et al., 1991). As noted earlier in the chapter, in the SVR, Gough and Tunmer (1986) state that word decoding and linguistic comprehension are two somewhat orthogonal elements of successful reading comprehension. These patterns of predictable strengths and weaknesses of decoding and linguistic comprehension within the SVR set out a frame of a possible pattern of individual differences in reading difficulties (Fig.4). Hence the SVR could be a rather useful practical tool in understanding the relationship between language and word reading in reading difficulties (Catts et al., 2003).

Figure 4

The Simple View of Reading as a Classification Tool reprint from Nation and Norbury (2005)



The SVR states that comprehension is a product of decoding and linguistic comprehension. Hence children with poor reading comprehension might have deficits in decoding (quadrant A), linguistic comprehension (quadrant D), or both (quadrant C).

Nation et al., (2010) conducted a longitudinal study to explore reading and language skills in children with poor reading comprehension. Fifteen poor comprehenders matched with controls were tested on language and reading skills at ages 5.5, 6, 7, and 8. The results showed that poor comprehenders had normal accuracy and fluency and their reading comprehension was poor at all time points. Poor comprehenders had intact phonological skills, yet they showed impairments in listening comprehension, grammar, and expressive and receptive language. This suggests that poor comprehenders would indeed have difficulty in comprehension regardless of their word accuracy and fluency performance and that oral language impairments could also be a consequence of reading comprehension difficulties.

Another factor that might be related to the nature of reading comprehension difficulties is lexical-semantic processes. This hypothesis was based on research studies that showed that individuals with reading comprehension difficulty perform very poorly on semantic judgment tasks (Landi & Perfetti, 2007).

In summary, reading disabilities definition is complex. The DSM-5 categorizes dyslexia within the broader spectrum of specific learning disorders, highlighting impaired reading in various aspects. Zeigler et al.'s research challenges a narrow focus on phonological deficits, proposing a triangle model that incorporates orthographic and semantic pathways. The multi-deficit computational model by Zeigler et al. (2020) reinforces the idea that dyslexia is a result of complex interactions across multiple components, including letters, phonemes, and semantics, rather than a singular deficiency. Wolf and Bowers (1999) introduce the concept of rapid naming deficits alongside phonological awareness, contributing to the double deficit hypothesis. This is supported by studies such as Torppa et al. (2013), demonstrating that children with deficits in both rapid naming and phonological awareness face more substantial reading challenges. Additionally, the notion of 'surface dyslexics' adds another layer, suggesting that average decoders may struggle with reading and language skills due to issues in lexical and lexical-semantic processing. As for reading comprehension, the SVR emphasizes the independent roles of decoding and linguistic comprehension. Nation et al.'s (2010) study on poor comprehenders underscores that difficulties in reading comprehension can persist despite normal accuracy and fluency, indicating broader language impairments.

In essence, the complex relationship of components across various levels of processing, as indicated by the research, highlights that reading disabilities, including dyslexia and comprehension difficulties, are not monolithic but rather multifactorial. The acknowledgement of these complexities is crucial for a comprehensive understanding and effective intervention in addressing reading-related challenges.

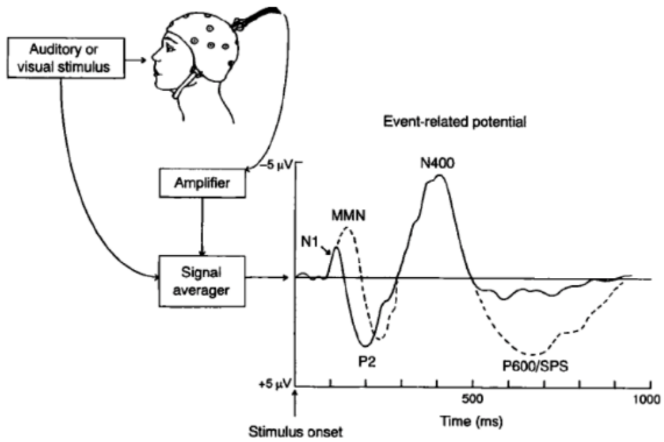
Reading Disability and EEG

A beginning reader needs to develop phonemic and phonological awareness to understand individual sounds in written words (Muter et al., 2004). Reading however is a complex skill, and there may be many reasons why some individuals fail to learn to read. As noted above, Wolf and Bowers (1999) have argued that deficiency in rapid automatized learning and phonological awareness causes reading disability. Zeigler (2020), as discussed earlier, says that RD results from multiple semantic and phonological-orthographic deficits, not just one deficit. In many cases, individuals with RD who have poor word identification have difficulty in reading comprehension (Gough & Tunmer, 1986; Perfetti, 2007).

Basic behavioral research of reading has informed us of the cognitive processes that are likely involved in reading acquisition, reading fluency, and reading comprehension as well as of the deficits in cognitive process in individuals with RD. To build on the study of behavioral reading processes, the fullest accounts will also describe underlying neurophysiological processes (e.g. Morton & Frith, 1995). Electroencephalography (EEG) is a brain imaging method that measures the electrical activity of the human brain by placing electrodes on the scalp and computing the negative and positive peaks of electrical activity resulting from a stimulation event (Luck, 2005). Brain activity can be assessed using Event-Related Potentials (ERPs) by time-locking the neural response of specific stimuli. ERPs can be used to investigate how the brain processes language in real-time with a high temporal resolution (Coles & Rugg, 1995), since they have a fine-grain temporal accuracy in milliseconds (Kaan, 2007) (Figure 5). ERP studies are a common imaging approach used in the RD literature, especially with children (Caylak, 2009).

Figure 5

Illustration of an ERP procedure reprint from Kaan (2007)



Electrodes are placed on a participant's scalp. The participant is presented with either auditory or visual stimuli task i.e., reading or listening to words. A signal will be obtained via placed electrodes. The signal is then amplified and averaged, time-locked to the stimulus of interest. Time in milliseconds is on the x-axis, and the voltage difference is on the y-axis.

Most relevant to this thesis proposal, the N400 is a negative ERP waveform peaking at 400 milliseconds (300-500-time window) that appears to be sensitive to the lexico-semantic aspects of language The N400 has been evaluated using semantically incongruent words in a non-linguistic context, such as in pictures (Kutas & Hillyard, 1980). Kutas and Hillyard (1980) conducted an experiment modifying the methodology of an earlier study by Schuberth and Elimas (1977) where participants read sentences with an unexpected stimulus, such as "I drink coffee with cream and SUGAR" where SUGAR is in upper case. A P3b is a positive peak occurring at 300 ms after the participants read the sentence with the unexpected stimulus at the end. Kutas and Hillyard incorporated an additional condition with sentences that include semantically incongruent endings such as "*The pizza was too cry*". The authors discovered that with these types of sentences, there was no P3b but rather a robust peak at negative 400 ms (For review, see Kutas & Fredermeier, 2011). Holcomb et al. (1992) also found that ERP activity peaked at negative 400 ms when participants read sentences with incongruent endings such as *Kids learn to read and write in* ***finger***. The same results were not seen in participants when they read semantically correct sentences as in children who *learn to read and write in* ***schools.*** Comparable results were reported by Henderson et al. (2011) with 8-10-year-olds, where participants had a sensitive N400 with semantically incongruent sentences. Kaan (2007) adds that the N400 is used in sentences in two conditions. In this case for example, a participant would read two sentences during an ERP: The first would include a semantic anomalous word as in *the fish swims in the* ***tree*** a sentence with a plausible word as in *the fish swims in the* ***sea***. Researchers use these two conditions in an ERP task to better understand how the N400 peaks when participants read both sentences. Kaan (2007) found that participants elicited an N400 peak in the anomalous condition, which indicates that the N400 is sensitive in sentences with incongruent meanings.

Furthermore, Kutas and Fredermeier (2011) found that the N400 elicits the same sensitivity in incongruent words and pictures as in sentences, e.g. looking a picture of a CONE and the label word is BONE. Pre-readers have also shown sensitive N400 with lexical-semantic tasks involving incongruent words. For example, researchers uncovered a sensitive N400 after showing young infants two priming words followed by an incongruent word like "*banana*" followed by "*yellow*" and then "*doctor*" (Duta et al., 2012; Friedrich and Friederici, 2008). Some researchers note that the N400 effect is also found in pseudowords like *cigpet* (Kutas & Federmeier, 2011; Laszlo & Federmeier, 2009).

The N400 has been used in the reading disability literature as an ERP component in attempts to identify children with reading disabilities. The N400 is aberrant in atypical readers in well-conducted studies (Hasko et al., 2013). Researchers have found that children with RD have orthographic and phonological deficits which affect their semantic processing by showing an aberrant N400. Bergmann, et al., (2005) as an example, showed that young readers with RD, who scored at the 15^{th} percentile on reading fluency and spelling tests, had a small N400 amplitude in comparison with typical readers on a task that required participants to read words and non-words. Plante et al., (2000) found that in a group of adults with reading disabilities, who scored below level on the Peabody Vocabulary test (PPVT), language processing tests and a reading span test, did not show an N400 amplitude on a task that requires participants to distinguish between semantically related and unrelated word pairs. Thus, the N400 results suggest that there is limited semantic processing in individuals with RD. Perfetti et al. (2008) argue in their study that semantic processing deficits resulted in reading comprehension difficulties. The authors found that participants with poor reading comprehension showed delayed event-related potential (ERP) processes when asked to complete lexical-semantic tasks compared to matched-aged controls that were good comprehenders, suggesting that those with comprehension difficulties have slower word processing. Hasko et al. (2014) ran a reading intervention and showed a more negative N400 amplitude post-reading intervention with young children between 8-10 years with dyslexia. This suggests that individuals with aberrant N400 effects might benefit from early reading intervention (Duncan et al., 2009; Friedrich & Friederici, 2006).

Reading Intervention

Reading difficulties and reading disabilities alike may be preventable if a reading intervention focusing on effective and evidenced teaching practices is provided at a very young age (Wanzek et al., 2018). For readers with RD, reading intervention and prevention has been widely used in research for its potential to improve students' reading (Wolff, 2011). Scammacca et al. (2014) showed in their meta-analysis of intervention studies in reading in grades 4-12 had a much lower effect size than intervention implemented before grade 3. Students who do not receive intervention before grade 3 continue to have difficulty in reading throughout school and run a raised risk of subsequently dropping out of school (Hernandez, 2011).

The National Reading Panel (2000) report has indicated, based on a meta-analysis, that explicit instruction in phonics, phonemic awareness, vocabulary, fluency and reading comprehension are critical for the development of reading skills. There are several meta-analyses that have been conducted to show the effect of phonics instruction on student reading attainment. Ehri et al., (2001) conducted a meta-analysis to evaluate the effect of systematic phonics instruction, the explicit instructions of grapheme-phoneme rules, from 38 treatment-control studies. The outcome measures included in these studies were decoding regular words, pseudowords, reading mispronounced words, reading text orally and comprehending a text. The results of the meta-analysis showed that phonics instruction had a positive effect on decoding, word reading, spelling, and text comprehension. Blachman et al. (2014) conducted a study with grade 2 and 3 readers with poor word decoding skills. The students were assigned to 8 months of intervention that focuses on phonics and orthography along direct mapping in books. The post-test showed greater improvements in word reading, non-word reading, spelling as well as passage reading that were sustained to at least some degree at a 10-year post-test. McArthur et al., (2018) conducted a meta-analysis to explore the effect of phonics intervention on elementary school students with RD. The intervention involved sight word training and phonics training (pairing letters, letter clusters, and syllables). Their results showed that phonics instruction modestly improved children's reading of words, non-words, and irregular words and slightly improved reading comprehension. This is to say that the earlier the intervention, the more effective the results are. were.

For elementary school students with RD, phonics instruction intervention has shown positive results in improving reading performance. For example, Noltemeyer et al. (2019) conducted a randomised control study in a group of kindergartners with poor reading skills. Their phonics instruction was comprised of teaching children to sound out individual consonant-vowel-consonant (CVC) such as *bat*, *bed*, consonant-vowel-consonant- consonant (CVCC) such as *lamp*, *pond*, and consonant-vowel-consonant-e (CVCe) such as *fate* and *late*, then blending these sounds and finally reading the whole word. Their results showed that the students had immediate improvement in word recognition. Other studies combined phonics instruction with other interventions such as sight word reading. For example, McArthur et al. (2015) used phonics instruction and sight word intervention to improve children's reading with children who are poor readers. The children were taught to blend grapheme-phoneme correspondences to synthesize word pronunciations for the phonics instruction component. The sight word intervention focused on learning whole words. Their results showed an improvement in reading both words and nonwords for both groups. Some research studies used a multiple-component approach to intervention. O'Connor et al. (2002) used an intervention, with poor readers, focused on phonemic awareness, word recognition, spelling, and comprehension. The results showed that students in the treatment condition outperformed the control in phonemic awareness, word reading, phonemic awareness, comprehension, and fluency on standardised measures. Wolff (2011) also conducted a multiple-component intervention study focusing on phonemic awareness, decoding, comprehension, and fluency with grade 3 students with reading difficulties. After 12 weeks of intervention the intervention group improved significantly in spelling, reading speed, reading comprehension and phonemic awareness. The results of both these studies suggest that students may benefit more if the intervention is focused on more than one element of reading.

Thus far, many English reading interventions focus on phonics intervention (McArthur et al., 2018), and many focus on word reading accuracy (Landerl & Wimmer, 2008). Since the English language does not simply follow GPC rules and contains many words that are exceptions to common phonic rules. The inconsistent orthography makes the language an opaque one. This may make readers with RD slower in reading and make more errors (Wolff, 2011). One common approach to this problem is to teach common exception words by sight (e.g. McArthur et al., 2018). An alternative approach some researchers have proposed is that in addition to teaching phonics, children are trained to “generate alternative pronunciations when they come to unknown words until they produce a pronunciation that is a real word, which makes sense in context” (Zipke, 2016, p72). This process is called Set-for-Variability (SfV) (Venezky, 1999). For example, a child is taught the regular pronunciation of “ch” in *match*, *hatch*, and *catch*. The child would then use this regularized pronunciation of “ch” to read the word *stomach, an* irregularly spelled word (Tunmer & Chapman 2012). If a child fails to read the irregularly spelled word *stomach* using the regularized pronunciation, then “the child has to change one or more sound associations and try again” (Venezky, 1999, p 232.). SfV is “the ability to determine the correct pronunciation of proximation to spoken English words” (Tunmer & Chapman, 2012, p123). This means having the ability to make a link from spelling pronunciations to conventional pronunciations. If pronouncing a word does not produce a meaningful word in context, the individual would need to try a different pronunciation (Venezky, 1999). Steacy et al. (2017) found that individual differences in such tasks in individuals with reading disabilities were predicted by non-word, decoding, and vocabulary. Steacy et al. (2019) conducted another study to understand the role of SfV on word reading. Their results suggested that SfV and student task performance strongly predict word reading. Their results also showed that phonological awareness did not significantly predict word reading when SfV was first entered into the model. This suggests that adequate phonological skills are needed but are not sufficient to lead to accurate word reading of non-words. Tunmer and Chapman (2012) used SfV to assess whether it contributes to other reading skills such as decoding, word recognition, exception word reading, vocabulary, and reading comprehension. Tunmer and Chapman (2012) hypothesized that SfV is a key mediating factor between vocabulary and word recognition skills. Children were tested three times: at the end of year one, middle of year 2 and middle of year 3 on measures of phonemic awareness, syntactic awareness, vocabulary knowledge, decoding, word recognition skills and SfV. SfV in year 1, indirectly influenced future reading comprehension through decoding and word recognition, and vocabulary knowledge by year three. The results also showed that phonemic awareness and vocabulary knowledge and SfV make unique contributions to reading. This suggests, as the authors state, that SfV may have a critical role in the growth of word recognition skills related to reading comprehension. Similarly, Elbro et al. (2012) found that even after controlling for word reading, phonemic awareness, rapid automatized naming (RAN), and vocabulary, SfV contributed to word recognition and reading of both regular and irregular words.

Dyson et al. (2017) ran an intervention involving explicit instruction of SfV to assess whether it improved children's reading of words. Typically developing students were trained on using SfV to read irregular words from Tunmer and Chapman's (1998) study. The students were taught "tricky words" that do not follow a grapheme-phoneme correspondence rule. The students practiced correcting the mispronounced and were taught their definitions. Once the participants were able to correctly read the word independently, the participants were asked to match words that rhyme and have the same first sound as the target word that was taught, and then were asked to write the target word on their own. To correct the mispronounced word, the participant would first (i) read the word aloud (ii) decide if the word has meaning (iii) if not, then the participant must think about what words sound like that word (iv) choose a word that has the most approximate sound (v) check if the word has correct meaning in context. After a four-week intervention program, the authors found that children in the intervention group made significant improvements in reading mispronounced words, in reading taught irregular words, and in defining the meaning of untaught words. The authors conclude that the SfV had generalized effects on an experimental untaught word set reading. This suggests that SfV training can have transfer effects. However, the effects of the intervention did not generalize to improvements in reading the Castles and Coltheart irregular word set (d = .12, p = .864), nor did the results generalize to any of the other single-word measures of reading, providing evidence of limited generalization.

Savage et al. (2018) implemented an SfV intervention with poor readers along with a Direct Mapping strategy, delivery of systematic synthetic phonics and shared book reading. As described earlier Direct Mapping (DM) involves teaching students GPC rules and immediately linking the taught correspondence rules to a given text where it is richly exemplified. The authors used two interventions; the first was the DM and SfV (DMSfV) program and the second was Current and Best Practices (CBP) reading program. The CBP program is an intervention with three main aspects: (i) a synthetic phonics model that focuses on blending and segmenting phonemes, (ii) teaching common sight words and, (iii) shared book reading, all in the absence of SfV content as well as the absence of teaching close linkage between the shared reading of read books and grapheme-phoneme rules. The results indicated that DMSfV produced a significant positive effect over CBP on word reading at the immediate post-test and on word reading and sentence comprehension at a delayed post-test 5 months after the intervention closed. However, it is unclear whether the improvement is due to the direct mapping of GPCs only, SfV only, or because it was combined. Evidence from a recent study by Yeung and Savage, (2020) that used Direct Mapping as an intervention for struggling readers with English as a second language is suggestive. They found that DM only produced an interaction effect rather than a main effect of the intervention on reading outcomes, and effects were limited to these struggling readers with better phonological skills. Savage et al., (2018) similarly showed in their study that children benefitted from a GPC intervention but those with poorer phonological skills needed additional support. This implies that the use of SfV is an effective approach in the Savage et al. (2018) study in teaching children with poor phonological skills.

In sum, findings from these studies suggest that SfV is a potential contributor to word decoding, word recognition, irregular word reading, and vocabulary, which are necessary skills for successful reading comprehension. However, all these studies were run with either typically developing children, children with weak reading skills, or at-risk readers. We have yet to know about the correlation of SfV with these reading skills in children with RD and the impact of SfV intervention alone on irregular word reading and SfV in children with RD.

Reading Intervention and Neurophysiological Changes

A very limited number of studies have sought to explore neural and electrophysiological changes that may result from reading interventions. For example, a functional magnetic resonance imaging (fMRI) study by Temple et al. (2003) used a reading intervention focused on identifying sounds of individual phonemes with 20 children (8-12 yrs. old) with dyslexia and 12 age-matched controls. The intervention took 100 minutes a day, presented 5 times per week, for a month. Post-reading intervention, children with RD had increased activity in the left tempo-parietal cortex, an increased activity associated with reading improvement. The authors collected fMRI images of participants during pre- and post-reading intervention. The results showed that children with dyslexia had increased activity in the left temporoparietal cortex, the left inferior frontal gyrus, and some parts of the right hemisphere. The results also showed a positive correlation between increased left hemisphere activation and improved phonological awareness. These results suggest that reading interventions improve reading ability and increase activation in the brain, mainly the left hemisphere.

Shaywitz et al. (2003) present data showing that children with RD who received a phonics-based intervention had increased activity in the left hemisphere post-intervention. Seventy-seven children between 6 and 9 years old participated in the study (49 with RD & 28 typical readers). The intervention focused on word-level instruction by reviewing sound-symbol associations and phoneme analysis. This phonologically based intervention increased activation in the anterior (inferior frontal gyrus) and posterior (middle temporal gyrus) reading systems. The increased activation in the left hemisphere suggests that phonologically based interventions increase activation in the neural systems that underlie reading skills and acquisition. These results also indicate that the type of reading intervention might play a critical role in individuals with RD.

These example studies from the literature show fMRI detected reading improvement post-reading intervention. ERP studies have also shown marked improvement in ERP responses associated with improvement in reading. For example, Lovio et al. (2012) used a three-hour intervention based on GPC training for kindergartners with reading difficulties. Thirty-one children scored one standard deviation below the age-expected performance participated. None were able to read a single word on a standardized Finnish test. They found that children post-intervention had increased attention to changes in recognizing sounds of letters measured by the ERP components, suggesting that reading improvement is reflected in functional changes in the brains of children as young as five. These results also suggest that even a brief intervention of three hours may result in positive behavioral and physiological changes in children with reading disabilities. However, again, the study has no control group.

Most relevant to this thesis, Hasko et al. (2014) found an increase in an N400 amplitude following a reading intervention that taught 8 years old with RD. The study included 28 children diagnosed with developmental dyslexia with 25 aged, matched control children. Pre-intervention measures showed that children with RD had a reduced N400 compared to controls. The authors allocated children with RD to two treatment groups: the first received orthographic knowledge training. The participants were taught to discriminate between long vowels followed by a silent /h/ and long vowel sounds followed by another consonant. The second treatment was GPC training that focused on learning letter-sound correspondence. Both interventions lasted for 6 months, delivered for 45 minutes per week. The authors showed that an increase in the N400 amplitude was associated with reading improvement, and an aberrant N400 was associated with children who did not respond to the reading intervention. The authors later investigated the difference in the N400 amplitude between those who responded to the reading intervention and those who did not. The N400 ERP was associated with reading improvement and a neural profile that accurately marked responders vs non-responders to the reading intervention. However, it was unclear whether the responders benefitted from the orthographic training intervention or the GPC intervention. We only have information between improvers over non-improvers in general. This was addressed as a limitation in the study as only 11 participants improved, which made it difficult to look further into intervention effects. It would perhaps still have given an insight into intervention effects had it been reported to which group the 11 improvers belonged.

Research Objectives

Event-related potentials are used to assess and understand the neural underpinnings of general language processes and reading disabilities (RD) in real -time. The N400 is a negative wave peaking at 400 ms (300-500 ms). Several researchers have reported that children with RD who have orthographic and phonological deficits, that affect their semantic processing and that is then evident in an aberrant N400. For example, Coch and Holcomb (2003) found that children with RD showed a slow shift rather than a peak at 400 ms in a lexical semantic task. Bergmann et al., (2005) compared children with RD with age-matched middle school controls who were typical readers and showed that readers with RD had a small N400 amplitude on a task that required participants to read words and pseudowords, while a larger N400 was elicited in the control group. These N400 results suggest limited semantic processing in young children with RD. Arújo et al. (2016) found a late N400 in adults with RD. A late N400 means that adults with RD had a smaller N400 amplitude that is ‘suppressed’ and less widespread. Jednoróg et al. (2010) showed that in a semantic priming task, children in the RD group showed a delayed N400 but were comparable to the TD control group on the semantic priming task. Children with RD had a reduced N400 in the phonologically incongruent condition compared to the TD group. This pattern suggests that some individuals with RD have semantic processing difficulties and phonological processing deficits. Sabisch et al. (2006), on the other hand, showed that children with RD elicited the same N400 as that of controls. Rüssler et al. (2007) found that adults with dyslexia had a delayed N400 on the rhyme-judgment and semantic judgment tasks. The delayed N400, in this study, was characterized by a late N400 onset which lasted longer than the N400 time window (300-500 ms). McPherson et al. (1998) found that adolescents with RD had a comparable N400 on a phonological task compared to an age-matched control group. In summary, these results show inconsistent patterns of findings in populations with RD.

Hasko et al. (2013) have raised the possibility that reported inconsistencies in results could be due to the modality of ERP. For example, Sabisch et al. used an auditory modality where participants listened to the task, whereas Bergmann et al. (2005) had participants read the task silently. Thus, the modality of stimulus presentation could be a candidate moderator. The age of participants and the stimuli task being used, such as incongruent sentences or reading words and pseudowords, could explain the reported discrepancies, as Hasko et al. (2013) elucidate. The inconsistencies could also result from orthographic or phonological processing deficits but intact semantic processing, as Plante et al.(2000) observed.

In the current literature, Kutas and Iragui (1998) conducted a semantic categorization task across typically developing adults aged 20 to 80. Their results showed that the semantic congruity effect gets smaller and slower as participants get older. This is consistent in many ERP studies with the N400 with typical readers. It needs to be clarified whether N400 declines with age in readers with RD since the studies discussed earlier in this section were from all age groups. For this reason, age could be another moderator that affects the N400 amplitude.

To explore whether the N400 is an electrophysiological measure of lexical semantic abilities that predicts word reading and to explore if there are moderators (such as modality and age as discussed above) that might affect the N400 amplitude, we first intend to synthesize and quantify the existing studies on the N400 in an attempt to establish whether it distinguishes between TD and RD across written languages using both a systematic review and meta-analysis (Chapter 4, Manuscript 1). Behavioral and ERP measures may be sensitive to different stages of information processing, and this is the first proposal in research so far that explores N400 and SfV in conjunction. Therefore, the study in Chapter 6 aims to compare ERP (N400) and behavioral data of SfV, word reading, phonological awareness, and comprehension to understand the dynamics of reading processes in young children with RD compared to age-matched TD. We, then, intend to use the N400 as a neural measure in a targeted reading intervention (SfV) that teaches children with RD strategies to develop a pronunciation repertoire for printed words and use context as well as their lexical knowledge (knowledge of words) in a randomized control study (Chapter 8). The N400 will be used in pre-post to determine if it predicts improvement.

Chapter 3. Bridging Literature Review to Manuscript 1

In Chapter 2, I discussed that successful reading acquisition is crucial for successful reading to take place. Reading acquisition depends on phonological, orthographic, and semantic access to a word exemplified by conceptual and implemented connectionist models of word reading and conceptual models of reading comprehension. Decoding and word reading are critical for successful reading to take place, underpinning comprehension. I also discussed evidence that RDs may be understood as reflecting difficulties in a number of these areas that are to some degree modifiable through theory-driven intervention such as the teaching of SfV. It is hypothesized that such difficulties and intervention effects are indexed neurologically by the N400 profile, though data on this question is currently lacking.

The first step in answering my research question, to explore whether 1) whether the N400 is an electrophysiological measure of lexical-semantic abilities that predicts word reading and comprehension, whether the N400 effect is different between TD and RD and to explore if there are moderators (such as modality language and age) that might affect the N400 amplitude 2). a SfV- intervention that focuses on teaching children with reading disabilities to manage grapheme-phoneme inconsistency in written English improves reading performance 3), and to explore whether the SfV reading intervention, having improved literacy outcomes, also changes the N400 amplitude and latency in children with reading disabilities. I first intend to synthesize and quantify the existing studies on the N400 to establish whether is different between TD and RD using both a systematic review and meta-analysis (Chapter 2).

Chapter 4. Manuscript 1

This chapter is an exact reproduction of the following article, published in the *International Journal of Psychophysiology*.

The N400 in Readers with Dyslexia: A Systematic Review and a Meta-Analysis

Badriah Basma[1], Robert Savage[2], Armando Bertone[1]

1. of Education & Counselling Psychology, *McGill University*, Montréal, Québec, Canada 3700 McTavish Street Montréal, QC, Canada H3A 1Y2

2. Department of Education, York University, Toronto, Ontario, Canada

Basma, B., Savage, R., & Bertone, A. (2023). The N400 in readers with dyslexia: A systematic review and meta-analysis. *International Journal of Psychophysiology*, 112283. https://doi.org/10.1016/j.ijpsycho.2023.112283

Abstract

This systematic review and meta-analysis aimed to assess whether (i) significant differences exist in the N400 response to lexico-semantic tasks between typically developing (TD) readers and readers with dyslexia, and (ii) whether these differences are moderated by the modality of task presentation (visual vs. auditory), the type of task, age, or opaque orthography (shallow and transparent alphabets vs Chinese morpho-syllabary). Twenty studies were included in the meta-analysis, and the analysis did not demonstrate strong evidence of publication bias. An overall effect size of Hedge's $g = 0.66$, $p < 0.001$, was found between typically developing readers and readers with dyslexia. All moderators were found to be significant; larger effects were associated with visual modality ($g = 0.692$, $p < 0.001$), semantically incongruent sentence tasks ($g = 0.948$, $p <.001$), pseudowords/characters tasks ($g = 0.971$, $p <.001$), and orthography [Chinese ($g = 1.015$, $p < .001$) vs. alphabets ($g = 0.539$, p < .001)]. Analysis of reaction time showed Hedge's $g = 1.613$, $p < .001$. Results suggest that the N400 reliably differentiated between typically developing readers and readers with dyslexia. Implications for future research and practice are discussed.

Keywords: *Systematic Review*, *Meta-analysis, N400, reading disability, dyslexia*

The N400 in Readers with Dyslexia: A Systematic Review and Meta-analysis

Reading fluency and comprehension are among the most critical components of academic success (Yong et al., 2015). Reading also plays a fundamental role in everyday knowledge acquisition and application (Hasko et al., 2013). This ability requires complex skills, including the identification and manipulation of words, phonemic awareness, and knowledge of letter clusters and their relationships to pronunciations. Additionally, broader linguistic comprehension skills are essential. (Castles et al., 2011; Melby-Lerväg et al., 2012). While most typical readers acquire these skills, they may not be as well-developed in some readers, particularly those with reading disabilities (Kemény et al., 2018).

A reading disability (RD) is characterized by the inability to read fluently and may result from a phonological deficit within the language system (Shaywitz & Shaywitz, 2005; Wimmer & Schurz, 2010). Early indicators of a reading disability can manifest themselves in slower naming speed and letter-sound correspondence use (Lovio et al., 2012). Some individuals with an RD usually perform poorly on phonological awareness activities such as rhyming, blending, and segmenting words (Ball & Blackman, 1991; Bradley & Bryant, 1983). Other difficulties include reading polysyllabic pseudo words such as *tegwop* or *vlinders* (Rack et al., 1992). An RD can also present as a deficit in reading comprehension (Aboud et al., 2016). In addition, reading fluency and reading comprehension are strongly correlated (Klauda & Gutherie, 2008). Therefore, individuals with RD may require assistance to simultaneously read fluently and comprehend a text. Poor performance, whether at the word or sentence level, ultimately results in poor reading comprehension (Landi et al., 2013).

Lexico-semantic processes involve deriving meaning from words and play a critical role in predicting reading comprehension. Perfetti et al. (2008) argue that deficits in semantic processing result in reading comprehension difficulty in adults and children. Hagtvet (1997, 2003) demonstrated that semantic processes predict poor reading comprehension performance in children with reading difficulties. Hulme and Snowling (2009) provide evidence that some individuals with RD can read aloud but exhibit poor text comprehension; these individuals also demonstrate low performance in semantic tasks. In addition, Kronbichler et al. (2006) suggested that individuals with RD and poor reading comprehension perform poorly on semantic processing tasks. Perfetti (2007) corroborated these findings, suggesting that individuals with poor reading comprehension perform much slower in retrieving word meanings. In addition to these findings, individuals with reading comprehension deficits find it challenging to perform simple lexical tasks such as differentiating between words and pseudowords (Kast et al., 2006). Therefore, a range of evidence suggests a role for semantics in reading (Taylor et al., 2015 for review). In conclusion, a large body of evidence suggests that lexico-semantics processes are essential predictors of aspects of reading comprehension, and an adequate lexical-semantic level is needed to efficiently complete reading comprehension tasks.

The Neural Underpinnings of Lexical-Semantic Processes

The present study aims to assess whether reliable differences exist in the N400 amplitude between typically developing readers and readers with RD and, if present, whether there are moderating factors that would explain these differences. One of the earliest language processing studies using an electrophysiology approach was conducted by Schuberth and Elimas (1977). In this study, typical readers were asked to read sentences containing unexpected stimuli, such as “I drink coffee with cream and SUGAR,” where SUGAR was presented in capital letters. The authors observed a positive peak occurring 300 ms after the participants read the sentence (P3b) when the unexpected stimulus was presented at the end of the sentence. Based on this paradigm, Kutas and Hillyard (1980) introduced an additional condition with semantically incongruent endings, such as "*The pizza was too hot to cry.*" They found that such semantically incongruent endings did not elicit an altered P3b but rather a robust peak at negative 400 ms or N400 (for review, see Kutas & Fredermeier, 2011). This finding provided the first evidence suggesting that the N400 can serve as a measure to investigate the neural basis of language comprehension acquisition in typically developing readers (Kutas & Fredermeier, 2011). The difference between the semantically congruent and incongruent conditions of a specific ERP task offers multi-dimensional information on the amplitude, latency, and topography of the N400. Thus, the N400 is a negative ERP waveform peaking at 400 milliseconds (300-500 ms time window) that appears sensitive to lexico-semantic aspects of language (Kutas & Hillyard, 1980). The N400 topography is most negative (large) in the centro-parietal region of the brain and spreads bilaterally over the left and the right hemispheres (Curran et al., 1993); a delay in the N400 latency is usually associated with a small N400 amplitude (Kutas & Iragui, 1998).

Although the N400 effect was initially demonstrated based on the lexical-semantic aspect of sentence processing tasks, it has been demonstrated that phonological processes also elicit the N400. For example, Dumay et al. (2001) reported an N400 effect in both rhyme judgment tasks and lexical decision tasks, and Radeau et al. (1998) observed such an effect using rhyme and semantic priming tasks. One possible explanation for the elicitation of an N400 during phonological tasks could be the participants' engagement in some aspect of semantic processing, particularly when their expectations are violated while reading non-rhyming or non-matching words. This phenomenon is referred to as a "semantic mismatch" by Rugg (1984). Another possible reason why phonological processes elicit the N400 is that individuals may engage in visual and orthographic analysis when presented with letter strings during phonological processing. This engagement results in bilateral neural activation, thus, an increased probability of detecting an N400 response.

In addition to phonological processes, the N400 effect is observed in orthographical matching words, non-words, and picture-based tasks (Barret & Rugg, 1990; Khateb et al., 2007). Studies have suggested an N400 effect on lexical decision tasks, where individuals decide whether a word is real or non-real, and in tasks involving word pairs, where individuals are asked to decide whether car-pen are related prime words (Holcomb & Neville, 1990; Kutas &Van-Petten, 1994).

The N400 in Typically developing (TD) Readers

Friedrich & Friederici (2005) demonstrated the N400 effect in children aged 19 and 24 months using a semantic sentence processing paradigm. In this study, participants viewed pictures on a screen accompanied by auditory words. While looking at the screen, participants listened passively to sentences with congruent and incongruent conditions. An N400 effect was elicited for both age groups (19 and 24 months) in incongruent conditions, with the most substantial negativity peak at parietal-central regions within the 400-700 ms window. These results suggest that TD infants can process sentences and semantically integrate words into a semantic context. In a similar study, Friedrich & Friederici (2008) reported that 14-month-old infants exhibited an N400 incongruency effect when listening to two priming words followed by an incongruent word like *banana, yellow, and* a *doctor* (Friedrich & Friederici, 2008). Specifically, their results indicate negativity for an incongruent condition over the parietal region within a 200-600 ms window for 14-month-old infants.

Duta et al. (2012) conducted a study with infants (aged 14 months) as well as TD adult readers (aged 21 years). Participants were presented with pictures while simultaneously listening to a word. The words presented were either real words, mispronounced words or pseudowords. For example, participants would be shown a visual image, such as "*fish,*", and simultaneously hear the word in one condition. In another condition, participants would see the same visual image of "fish" but hear it mispronounced as "fesh" or presented auditorily as a pseudoword like "soob." Both infants and adults exhibited an N400 peak for pseudowords within a window between 370- 670 ms. The results of this study suggest that there is an N400 effect in infants as well as adults.

Studies have also suggested the presence of an N400 effect in (TDs) school-aged children. For example, Henderson et al. (2011) used picture stimuli presented with either a correct (congruent condition) or an incorrect name (incongruent condition). Participants looked at the picture on a computer screen and simultaneously listened to the word through earphones. Henderson et al. report that 8-10-year-olds produced a similar N400 incongruency effect to semantically incongruent sentences. Holcomb (1993) also reported an N400 effect across a large age range (5 to 26 years) when participants were asked to read sentences with incongruent endings, such as "Kids learn to read and write in *finger*." Notably, participants did not exhibit the same results when reading semantically congruent sentences. In a semantic decision task for picture stimuli, McPherson and Holcomb (1999) similar results with a with a group of TD adult readers (mean age 21.5 years). The authors concluded that an N400 effect exists, characterized by negative amplitude peaking at 300-500 ms with incongruent conditions.

Researchers have also demonstrated that the N400 is elicited not only in response to semantic manipulations but also when orthographic and phonological manipulations are applied, particularly with pseudoword stimuli. For example, Khateb et al. (2010), conducted a study involving 22 typical French adult readers (mean age 23 years) in which participants performed 3 tasks: a semantic judgment task with semantically related (congruent condition) and unrelated pairs (incongruent condition), a phonological task with orthographically unrelated but phonological related word pairs (congruent words such as *tabac (tobacco) – bras (arms)* *tabac* = [*taba*] – *bras* = [*bRa*]) and unrelated pairs (incongruent), and a visually and semantically related image categorization task comprised of related (*cigar-pipe*) and unrelated (*table- elephant*) images. The participants were asked to read the word pairs silently and decide whether the words were semantically related (semantic task), whether the words were phonically related (phonological task), and whether the images were related (image task). Khateb et al. (2010) found an N400 peak amplitude difference in the semantic and phonological tasks but not in the image categorization task. This suggests that the N400 serves as an index of lexical-semantic processing.

In a study by Coch and Benoit (2015), typically developing third, fourth, and fifth graders along with college students, were asked to read words visually presented pseudowords, unpronounceable letter strings (e.g., *mbe*, *nrfgi*), and a false font. The N400 was elicited among all age groups for pseudowords. These results suggest that phonological and semantic processing have similar N400 effects. Kutas and Van Petten (1994) argue that there is an N400 effect with pseudowords because these types of words require more processing time for the reader to determine whether they are actual words and whether they fit the presented context. Friedrich and Friederici (2006) further explain that because pseudowords adhere to phonological rules, the reader needs to decide whether the words are meaningful or not. This is consistent with the interpretation of the N400 as a metric of the lexical-semantic process.

In conclusion, the reported data suggest the N400 is highly sensitive to individual differences in lexical-semantic skills across developmental periods, including TD infants, school-aged children, and adults. Moreover, the N400 effect is evident in lexical-semantic tasks involving phonological abilities, pseudoword reading and orthographic processing, with some evidence of developmental patterns for the latter stimuli. As Kutas and Van Petten (1994) suggest, the N400 effect is present across various tasks requiring some form of semantic processing common to sentence incongruity.

The N400 and Readers with Dyslexia

Hasko et al. (2013) investigated the N400 with eight-year-old dyslexics compared to age-matched controls to investigate children's phonological and orthographical processing. Participants completed a decision task, using a button press to indicate whether a word presented was a real word, a pseudoword or a pseudo homophone. The results indicated a weak N400 for participants with dyslexia, typically developing children showed a robust N400 effect. The authors concluded that children with dyslexia might have orthographic deficits and impairments in applying grapheme-phoneme correspondence rules that the N400 indexes. Bergmann et al. (2005) report slightly different results in for 13–14-year-old children with and without dyslexia. Results indicated that readers with dyslexia exhibited a smaller N400 amplitude than typical readers on tasks requiring participants to read words and pseudoword; a larger N400 was elicited in the control group. A similar finding was found Robichon et al., (2002), in which twenty-three-year-old participants with dyslexia and age-matched controls were asked to read congruent and incongruent sentences silently. The authors found that adults with dyslexia demonstrated a delayed N400 latency (time interval) for congruent and incongruent sentences. This was also reflected in a reduced N400 for the participants with dyslexia in both conditions (congruent and incongruent), signifying more time to process the sentences. Robichon et al. (2002) argue that this finding may suggest that individuals with dyslexia find it difficult to integrate the meaning of words into a sentence context.

Plante et al. (2002) conducted an N400 study with 16 adults with dyslexia and 16 age-matched controls. The participants were presented with nonverbal and verbal conditions. In the nonverbal condition, they were presented with environmental sounds while verbal condition, they were presented with printed and spoken words that were either semantically related or semantically unrelated. Unlike the controls, the dyslexia group showed no significant difference in N400 amplitude between related and non-related words. The authors proposed that adults with dyslexia show difficulty in semantic processing and thus could not differentiate between semantically related or non-related words. The semantic processing deficits were also reflected in the aberrant N400 amplitude.

Jednoróg et al. (2010) conducted a study with 18 children with dyslexia and 18 age-matched control. In a semantic priming task, participants listened to a list of seven words in which the seventh word was either semantically congruent or incongruent. In a phonological priming task, participants listened to seven words that shared two or three phonemes but were semantically unrelated or listened to seven words with different phonemes but were semantically related. Their results suggested that children with dyslexia had a reduced N400 in the phonologically incongruent condition compared to the control group. On the other hand, Sabisch et al. (2006) conducted a study where 16 participants with dyslexia and 16 age-matched controls were asked to listen to semantically congruent versus semantically incongruent sentences. The results indicated that children with dyslexia elicited the same N400 as controls. The authors concluded that the participants with dyslexia showed no difference in the semantic integration process (N400 amplitude) during an auditory sentence comprehension task. Additionally, McPherson et al. (1998) also found an N400 effect in individuals with RD comparable to an age-matched control group when participants were asked to listen to a set of words and complete a lexical decision task.

In conclusion, the research presented thus far suggests that the N400 effect remains consistent in participants without RD. However, participants with dyslexia exhibit varied N400 amplitude, with some studies demonstrating an aberrant N400 in individuals with dyslexia (Bergmann et al., 2005; Coch & Holcomb, 2003; Hasko et al., 2013) while others show a comparable N400 effect across both groups (McPherson et al., 1998; Sabisch et al., 2006). Considering these inconsistent results, Hasko et al. (2013) proposed the possibility that reported inconsistencies could be due to various factors including the modality of stimulus presentation (auditory versus visual mode) in ERP studies. For example, Sabisch et al. (2006) used an auditory modality where participants listened to stimuli, whereas Bergmann et al. (2005) asked participants to read the task words silently. Consequently, we explored the modality of stimulus presentation as a candidate moderator of the N400 effect in readers with dyslexia.

Several additional factors may explain the inconsistency in the N400 effect between typically developing readers and readers with dyslexia. The variability could stem from the stimuli or task employed in the study, such as incongruent sentences (Schulz et al., 2008), or reading words and pseudowords (Hasko et al., 2013). The inconsistency could also result from some dyslexic participants exhibiting orthographic or phonological processing deficits but intact semantic processing, as observed by Plante et al.(2000). Kutas and Iragui (1998) conducted a semantic categorization task across adults ranging from 20 to 80 years old. The authors found that the semantic congruity effect gets smaller and slower as participants get older. This is consistent in many ERP studies with the N400 with neurotypical readers. It is unclear whether

N400 declines with age in readers with dyslexia since the earlier studies were from all age groups. Therefore, age also could be another moderator that affects the N400 amplitude and contributes to the observed inconsistencies.

A final issue concerns the spelling system used, a factor recognized by Frost (1987) as influential in the speed of learning the alphabetic system of an alphabetic language, which depends on how opaque or shallow the grapheme-to-phonemes rules are in a given language. Serrano and Defior (2008) further explain that shallow or transparent orthographies have consistent or highly predictable grapheme-phoneme correspondences across all words in the spelling system. Opaque orthographies are more complicated in that these orthographies include inconsistent/irregular spelling patterns (e.g., ‘oo’ in ‘moon’ and ‘good’). Seymour et al. (2003) classified orthographies such as English, French, Dutch, and Portuguese as opaque, whereas Spanish and German were considered shallow. Wimmer (1993) argues that dyslexia in languages with opaque orthographies is characterized by inaccuracy and reduced reading rate. In contrast, dyslexia in shallow orthographies is best described as being based on reduced reading speed rather than reduced reading accuracy. Non-alphabetic morpho-syllabic spelling systems, such as Chinese, require a high degree of visual and semantic processing with minimal or no phonemic processing (e.g., Share, 2008). Recognizing the impact of language orthography on dyslexia, we have chosen to investigate it as a moderator to discern its potential role in influencing the N400 amplitude.

To firmly establish the reported patterns of the N400 effect, more consensus across studies is needed regarding the N400 effect in individuals with RD compared to age-matched controls. Despite the significance of this topic, there are currently no systematic reviews or meta-analyses that comprehensively investigate the differences in the N400 effect between typically developing readers and readers with dyslexia. In addition, it is also imperative to understand whether the inconsistency of N400 amplitude findings in dyslexia groups can be systematically attributed to task modality, age, or other moderators, such as the orthography studied. To address this void in the literature, this systematic review and meta-analysis aim to synthesize findings from group comparison studies examining the N400 in typically developing readers and readers with dyslexia.to assess the N400 profile differences between the two groups. The following research questions will be addressed; (i) What is the difference in mean N400 amplitude in typically developing readers and readers with dyslexia? (ii) Is the difference in the N400 amplitude between typically developing readers and readers with dyslexia moderated by ERP modality, stimulus task, age, and experiment language?

Method

The first author developed a list of keywords based on keywords in N400 research articles such as Coch and Holcomb (2003), Hasko et al. (2014), and Kutas and Federmeier (2011). These keywords are *N400*, *Event-related potentials*, *semantic congruity*, *semantic incongruity*, *semantic processing*, *listening comprehension*, *reading comprehension, lexico-semantic access, reading deficit, developmental dyslexia, word recognition, phonological, lexical decision, lexical priming, and phonological processing.* To identify relevant literature, a search was conducted using the aforementioned keywords using the following databases: *Prospero*, *The EPPI Center for Systematic Reviews*, *The Campbell Collaboration Library*, *The Cochrane Library, ERIC (ProQuest)*, *PsychInfo*, *MEDLINE*, *ProQuest Dissertations & Thesis, and Scopus*. The first author recorded the outcomes in a table, including the journal's name, associated keywords, and the number of hits obtained. The keywords were combined using the Boolean search functions of “AND”, “OR”, “NOT” and exploding keywords N400 “AND” semantic incongruity "AND" dyslexia OR dys* OR reading dis* or read* dis*. Additionally, keywords such as "Event-related potentials" were exploded using the Explode (Exp.) function, ensuring a comprehensive and systematic search. For example, "ERP" AND "N400" AND "semantic congruity" OR "semantic incongruity."

All articles were then transferred to the Rayyan application for further examination of titles and abstracts to determine the inclusion or exclusion of studies. The inclusion and exclusion process through Rayyan was conducted in a completely blinded manner using the "Blind ON" feature. The blind review was completed independently by the first and second authors, who assessed the abstract, keywords, and text to decide whether the article met inclusion and exclusion criteria. Subsequently, both authors met 4 weeks later with the "Blind Off" setting to review and discuss the results of their inclusion, and exclusion decisions. Both authors achieved a consensus with more than 90% agreement. However, the authors disagreed on 5 articles, which were resolved by consulting a senior academic expert on systematic reviews. The search methods for both phases were conducted at three time points: October 20th, 2020, and January 20th, 2021, and March 16th, 2021, June 2021, to determine any updates. Additionally, a search was also conducted within the reference sections of all articles for other potentially relevant studies and then within those relevant studies for other candidate studies, using a 'snowballing' technique. Conference abstracts and unpublished theses were also searched. The first author and a graduate research assistant then independently coded a subset of all the inspected records to establish the reliability of inclusion and exclusion criteria that reflect the main research question.

Figure 1 is a PRISMA flow diagram depicting all studies identified through the search, inspected, and either included or excluded in the review.

Inclusion Criteria

- Studies that use the N400 as a dependent variable
- Studies that focus on N400 and dyslexia OR reading disability (Table 1).
- Studies that included N400 amplitude and latency in ERP tasks that involved language processing.
- Studies that included N400 in all age groups
- Studies used the N400 task with sentence incongruity tasks, phonological-lexical decision tasks, reading words and non-real words, and picture-word tasks.
- Studies that included N400 in readers with dyslexia in comparison with typically developing participants

Exclusion Criteria

- Studies that focus on N400 and autism spectrum disorder (ASD) or Asperger's Syndrome
- Studies that focus on N400 and psychiatric disorders (e.g., schizophrenia)
- Studies that focus on N400 and memory
- Studies that focus on receptive and expressive language impairments
- Studies that included no comparison group
- Studies that included N400 with dyslexia co-morbid with other disorders such as "ADHD."
- Studies that included N400 with dyslexia and other learning difficulties such as "dysgraphia" or "dyscalculia.

- Studies that focused on N400 and emotional violation of faces and emotions
- Studies that used the N400 and emotional impulsivity.

Results

Study Selection

The keyword search initially resulted in 199 articles. Subsequently, the first author identified another 15 themes through a snowballing technique, i.e., reviewing reference lists of potentially included studies. Forty-five studies were identified as duplicates. After screening the abstracts of all 169 articles, 124 were excluded as they did not fit the inclusion criteria. The remaining 45 studies (*K*= 45) were thoroughly assessed. Of the 45 studies, an in-depth reading of the entire text revealed that 25 articles did not fit the inclusion criteria and were consequently excluded. Inter-rater reliability was established for the remaining 20 studies with a graduate student, yielding a kappa of .82, which is considered acceptable. Please refer to Figure 1 for details of selected articles.

Coding Articles for Quality

We first assessed the quality of the 20 included studies based on a modified list derived from Downs and Black's (1998) taxonomy for quality assessment for non-randomized control trials (Appendix A, Table 3). The assessment focused on key aspects, including whether studies reported their method of allocation, established a comparison group through randomization, and provided a detailed description of the randomization method. The criteria for sample justification involved determining the adequacy of sample size justification (denoted as "n") and assessing the power estimate. The "intention to treat" criterion assessed whether the study's groups were statistically analyzed based on their original assignment, irrespective of any subsequent attrition or dropouts. Each criterion received a score according to predefined categories: "Yes" indicated inclusion along with an explanation of the criterion, and "No" signified the absence of an explanation.

Results indicated that all 20 studies were strong in reporting the method of allocation and provided detailed information about the participants' characteristics. Furthermore, all analyses clearly described their primary outcomes. However, a critical issue across all 20 studies concerned internal validity; none used blinding of assessment or reported attrition and reported reliability or validity of the task. Additionally, none of the studies reported how they controlled for extraneous variables, which are critical missing methodological controls, especially given that most were quasi-experimental studies. These limitations thus affect the certainty one can attribute to the conclusions drawn from the included studies.

To further assess the quality of the included 20 articles beyond the primary inclusion criteria, we conducted a Weight of Evidence (WOE) analysis using a method adopted from the EPPI Center for systematic reviews (Appendix A, Table 4). WOE is a specific model used to decide the overall quality of included studies in a systematic review (Gough, 2007). The Weight of Evidence is based on three main questions: WOE A: Are the reported findings and the study question internally consistent? WOE B: Is the research design appropriate to the research question? WOE C: Is the study relevant to the review question? Cumulative answers to these questions are reported as an overall D ranking of either "High," "Medium," or "Low." Any study with a "Low" A is reported low in all other categories and thus excluded for further analysis. However, studies that say "High" or "Medium" on WOE A are evaluated in all categories. For example: if an investigation is assessed with "High" on WOE A, "Medium" on WOE B and "LOW" on WOE C, then overall WOE D would be Medium. Suppose the study is assessed with two Highs and one Medium. In that case, the study will have a general High D. The first author and a graduate student independently coded all 20 articles for WOE quality to establish reliability. Coding showed a reliability Kappa of .79, which is acceptable. Primary points of contention revolved around the internal consistency and appropriateness of research design in the described studies. However, disagreement between the two reviewers was minimal, at less than 3% concerning the details of the studies. These disagreements were addressed through discussions, referencing original study data and decision codes to reach a consensus. The WOE coding assessment showed that all studies included in this review are of High A, with all studies reported as medium B and Medium C, as they did not report any randomization in their research design, nor any evidence of validity or reliability in the tasks used nor as noted above, nor did they say how they controlled for extraneous variables.

Study Characteristics

The characteristics and the results of the included studies are presented in Appendix A, Table 2. The studies originated from China, Europe, North America, and South America. Reading tasks and measures in Europe were assessed in multiple languages: German (K= 6), Dutch (K= 1), Polish (K= 1), Swiss-German (K= 1) and Italian (k = 1). All Studies in North America were conducted in English (K = 3). Five studies were conducted in Chinese. In South America, two studies were in Portuguese. All studies examined the N400 in groups with dyslexia compared to typical readers groups.

Sample Characteristics

The total sample size across all included studies was 709 participants. All studies had relatively small sample sizes. Bonte and Blomert (2004) reported the smallest sample size, n = 18, with Hasko et al. (2013) reporting the largest sample size, n = 81. Six studies included elementary school-age participants (M_{age} = 7.5). Eight studies included middle school-age participants (M_{age} = 12). Five studies included adults (M_{age} = 23), and two studies included adolescents (M_{age}= 15.5). All included studies reported two groups: TD participants as an age-matched control group and participants with dyslexia or low reading. All studies included poor readers based on reading below the 20-25th percentile or on reading below grade level in standardized reading assessment batteries. All studies reported that none of the participants had any co-morbidity with other disorders such as attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), or anxiety. All studies reported normal-range IQ levels for both groups.

N400 Latency Across Studies

None of the authors of the included studies provided *p* values or *f* values to report differences in latency. All studies reporting latency only reported whether a latency difference existed or not. Thus, we could not take the latency analysis further into the meta-analysis. Three studies (Araújo et al., 2016; Bonte & Blomert, 2004; Sabisch et al., 2016) reported no N400 latency differences between the neurotypical and readers with dyslexia group. The same studies found a comparable N400 effect among the two groups. Desroches et al. (2013) found latency differences between the two groups only with unrelated mismatch tasks. Six studies found differences in the N400 latency between the two groups (Aiello et al., 2018: Brandeis et al., 1994; Chung et al., 2012; Desroches et al., 2013; Hasko et al., 2014; Robichon et al., 2002) with the dyslexic group being slower. These studies also found aberrant N400 amplitude in the group with dyslexia. Two studies (Hasko et al., 2013; Schulz et al., 2007) found no latency differences between the two groups but reported aberrant N400 amplitude in the group with dyslexia. The remaining studies did not report any results on latency.

N400 and Topography Results Across all Studies

Five studies reported topographical results comparing control and dyslexia groups. Araujo et al. (2016) found topographical differences in the phonologically primed task as a widespread N400 restricted to the right hemisphere in the dyslexic group. Brandeis et al. (1994) reported topographical differences in semantic incongruency between the two groups. Desroches et al. (2013) described topographical differences specific to phonological mismatch but not in other tasks. Jednoróg et al. (2010) showed topographical differences in the two groups manifested by a delayed N400 response in semantic and phonological priming tasks in the frontal regions compared to typical readers. Rüssler et al. (2007) found topographical differences reflected by strong left lateralization in the group with dyslexia. Tzeng et al. (2018) reported topographical differences where a more negative N400 was shown in the frontal regions in the dyslexic group, while controls showed more negative N400 in the frontal and central parietal regions of interest. These were the only five studies describing topography results, but no statistical values were given. Therefore, we could not move further with these results in the meta-analysis.

Behavioral Results: Reaction Time and Accuracy

Our primary focus here was on N400. We thus briefly report relevant behavioral data for selected papers. Three studies did not report information on reaction time and accuracy rates (Table 5). Two studies reported similar accuracy rates where both groups had a similar rate of correct responses in congruent and incongruent conditions. The remaining 15 studies reported that the dyslexia group responded slower to the task (reaction time) and made more errors (accuracy) than the control group, as might be expected. We conducted a meta-analysis of the behavioral data, and the results showed an overall effect size for reaction time of $ES = 1.623$ (CI = 1.141- 2.105), $Q = 60.068$, $df(Q) = 12$, $p < .001$. We could not run a meta-analysis for accuracy rate because only four studies reported the percentage of accuracy, and two studies reported the *p-value*. Some studies reported standard error, which is inconsistent with studies that reported standard deviation.

Candidate Moderators

The next step in this analysis was to examine candidate-moderators throughout all 20 studies to see if any of the moderators affect the N400. The candidate-moderators identified in the literature review were the ERP modality (auditory versus visual), stimuli task, participants' age, and the task's language.

Modality

The studies used two ERP modalities to complete the N400 tasks: auditory (K= 7) and visual (K = 13). Two studies had both stimuli (McPherson et al., 1998; Silva et al., 2016). Thus, the results of the studies were explored separately and were reported as separate effect sizes in the meta-analysis. In the auditory modality, participants typically listened to words or sentences while looking at the screen and then decided whether the word was semantically correct. Three studies using an auditory modality found typical N400 in both groups (Bonte & Blomert, 2004; Jednörog, 2010, Sabisch et al., 2006). Using the auditory modality, two other studies found typical N400 only in the dyslexia reader group that did not present phonological deficits (McPherson et al. 1998, 1999). One study (Rüssler et al., 2007) using an auditory modality found typical N400 in the group with dyslexia reading only with semantic priming tasks. One study in the auditory modality used a picture stimuli task (Bonte & Blombert, 2004). All studies reported reaction time and accuracy rate for the task except for five studies (Bonte & Blomert, 2004, Brandeis et al., 1994; Coch & Holcomb, 2003; Robichon et al., 2002).

Participants looked at the screen, read silently, and decided about semantic incongruity in the visual modality. All studies in the visual modality required the participants to read silently and judge whether the given stimulus was semantically congruent or not (Arújo et al., 2016; Brandeis et al., 1994; Chung et al., 2012; Hasko et al., 2014;2015; Meng et al., 2007; Rüssler et al., 2007; Schulz et al., 2008; Silva-Pereyra et al., 2003). Two studies (Araujo et al., 2016; Silva-Pereya, 2003) used pictures and figures as stimuli tasks in the visual modality. All studies in the visual modality included reaction time and accuracy for the task except for Kemény et al. (2018) and Robichon et al. (2002). All studies that used the visual modality found either dyslexia (more diminutive or late N400) in the RD condition compared to the typical readers except for Silva-Pereyra et al. (2003).

Literacy Task

Semantic congruent/incongruent sentences. Seven of the included studies used sentences that were either semantically congruent or incongruent with the stimuli (Table 1). These seven studies showed that participants with dyslexia reading had either a delayed N400 or no N400. Sabich et al. (2006) and Silva et al. (2016) used auditory modality with semantic congruent sentence comprehension and found that both groups were, in fact, sensitive to the N400. Silva et al. (2016) noted that the N400 was preserved. Only one study used both phonological (e.g., bull, beer) and semantic priming stimuli (e.g., the word "nurse" proceeded by "doctor" or "butter," Jednoróg et al., 2010). The study showed a delayed N400 in the group with dyslexia.

Phonological Stimuli and Lexical decision task. Seven studies focused on a lexical decision task as the main literacy task (Table 1). Three studies using this task found that the N400 was absent in the groups who had dyslexia, while the controls had the usually elicited amplitude of the N400. Bonte and Blomert (2004) found that the dyslexia group had N400 effects comparable with control readers. Three studies reported a delayed sensitive N400 occurring only at 650-800 ms. Hasko et al. (2014) was the only study that used both a phonological and a lexical decision (pseudowords) task, then provided a short reading intervention and measured the N400. Hasko et al. (2014) found an increase in an N400 amplitude following a reading intervention that taught eight-year-olds with RD. The authors placed children with RD in two groups: the first received orthographic knowledge training, which was based on transferring the correct graphemes into words. The second was grapheme-phoneme correspondence (GPC) training focused on letter pronunciation. Both interventions lasted for six months, delivered for 45 minutes per week. The authors showed that an increase in the N400 was associated with reading improvement, and an aberrant N400 was associated with children who did not respond to either reading intervention. The participants were required to read silently and decide whether a given word was an accurate word, a pseudoword, a pseudo homophone (e.g., JALE), or a false front (). Eleven participants improved post the reading intervention out of 28 participants. The reading improvement was associated with a typical N400.

Pseudowords/characters. In addition to Hasko et al. (2013; 2014) using pseudowords task, three other studies (Chung et al., 2012; Tong et al.; Wang et al., 2017) used pseudo characters in their task. All three studies found that participants with dyslexia showed no N400.

Meta-analysis

The first author calculated the effect size of the N400 amplitude of all 20 studies. Then, the first author converted the measures into Hedge's *g* for correction purposes by using an effect size calculator in Comprehensive Meta-analysis (www.meta-analysis.com). Hedges *g* in all studies was based on a random-effects model because random-effect models assume that the 'true' effect size can vary from study to study. McPherson et al. (1998) were treated as two different studies because the study had two different results: one for auditory and one for visual priming tasks. The same procedure was applied to Silva et al. (2016) study for the same reason. Overall, the analysis showed that the studies were moderately heterogeneous (Q = 32.008, df = 20, p = 0.10). The smallest positive effect size was 0.280, and the largest effect size was 1.284. Of the 20 studies, none reported a negative effect size (Figure 2).

Publication Bias

A funnel plot was created to assess the validity of possible publication bias in this analysis, see Figure 3. The result shows that studies with a small sample size are scattered around the bottom of the graph, and the studies with a relatively larger sample size are closely clustered together. The pattern in this analysis does not provide any substantial evidence for publication bias. However, there seems to be asymmetry around the effect size mean for the two more extensive studies. Beggs and Mazmudar's test for publication bias (1994) was used to analyze further if publication bias exists. The results showed p = 0.73 for this test. Also, Egger's test for regression intercept showed a non-significant effect, (p = 0.93). A non-significant p-value suggests there is no evidence of publication bias. Nevertheless, caution in interpretation is strongly advised because of the asymmetry of the two most significant studies.

The overall mean difference of the N400 amplitude between dyslexia and typical readers was g = 0.66, p < .001, 95% CI [0.391, 0.901], with an associated standard error of 0.083. A further analysis was added to explore which length modality, stimulus task, age, the language of the task, affect the mean difference in the N400 amplitude. We grouped studies with visual modality as "visual" and auditory modality as "auditory" to calculate the modality as a moderator. We grouped studies as "semantically congruent/incongruent sentences", "phonological stimuli," and "pseudoword/character stimuli" to calculate the stimuli task as a moderator. For age, studies were grouped with "children" vs "older participants." Older participants were 15 years and above. The language of the task used was divided into two groups, studies that had "opaque orthography" and "shallow orthography" based on Seymour et al.'s (2003) definition of opaque and shallow orthography. We grouped the studies conducted in Chinese characters as "morpho-syllabic."

A significant effect was evident in all moderators. Results showed that the stimulus modality moderated the overall difference of the N400 amplitude, but the analysis showed that $g = 0.749$, $p < .001$, 95% CI [0.453-1.045] for studies using a visual modality. No significant effect was evident in the auditory modality, $g = 0.466$, $p = 0.098$, 95% CI [0.087-1.019]. Age was also a significant moderator, and the difference in means declined but was still evident with age ($g = 0.8$, $p < .001$, 95% CI [0.533-1.067] for children, versus older participants $g = 0.53$, $p <.001$, 95%CI [0.148-0.915]. Stimulus task was also significant, where studies with the largest effect size for congruent/incongruent sentences $g = 0.948$, $p < .001$, 95% CI [0.618- 1.278] followed by pseudowords/characters task g = 0.692, $p < .001$, 95% 95% CI [0.320-1.0.64], while the phonological stimuli tasks were not significant with $p = 0.194$. The language used was significant, but both groups had a similar *g*: Opaque orthography was significant with $g = 0.765$, $p <.001$, 95% CI [0.545-0.985], and shallow orthography was significant with $g = 0.789$, $p < .001$, 95% CI [0.509-1.608]. Morpho-syllabic orthography (Chinese) was also significant with $g = 1.043$, $p <.001$, 95% CI [0.709-1.376].

Discussion

This systematic review and meta-analysis aimed to explore the N400 effect in typically developing readers and readers with dyslexia. The primary research question addressed was: What is the difference in mean in the N400 amplitude in typically developing readers and readers with dyslexia? This was investigated through a meta-analysis of well-selected studies identified through a detailed search. The second question explored here was: Do ERP modality, stimulus task, age, and experiment language moderate the difference in the N400 amplitude between typically developing readers and readers with dyslexia?

The inclusion criteria to the keyword search term identified 20 research papers included in this review, arguably a relatively small number given the growth of the N400 literature. However, the search strategies, electronic search, snowballing by a hand search of the reference list of included studies, and quality assessment of all included papers collectively provide some confidence in the results and conclusion of this systematic review and meta-analysis beyond the noted methodological weaknesses in some individual papers. All the included studies assessed the N400 amplitude in participants with RD by comparing them to age-matched controls. These included studies showed a difference in the mean in the N400 amplitude between these two groups.

The analysis of all the included studies showed an overall effect size of $g = 0.66$, which was significantly different from zero at $p < .05$. According to Cohen's (1988) interpretation of effect size, an effect size of 0.25 is small, 0.5 is medium, and 0.8 is large. As such, the effect size of this meta-analysis can be considered substantial and at least medium in size.

In terms of modality, researchers used either a visual or auditory modality for stimulus presentation. The results showed that the auditory modality studies ($k = 7$) showed typically and delayed N400 patterns with the dyslexia groups. In contrast, all the studies that used the visual modality ($k = 13$) demonstrated aberrant N400 patterns. The meta-analysis showed that considered together, studies that used auditory modality have a non-significant effect size $g = 0.466$, $p = 0.098$, versus a significant $g = 0.82$ in the visual modality. These findings suggest that the N400 is a variable that is strongly dependent on the stimulus modality. The non-significant N400 within the auditory modality presumably reflects participants recognizing the sounds of words by listening to the given oral instructions and with observed patterns reflecting cognitive-linguistic factors in processing experimental stimuli to resolve syntax and semantics.

On the other hand, individuals in the visual modality undergo the process of reading before making a lexical decision, likely indexing their core grapho-phonic difficulty in dyslexia and additional linguistic processing demands in processing experimental stimuli (Montgomery et al., 2016; Zwitserlood, 1989). A phonological processing deficit likely affects the ability to link sounds of letters and written letters (Authors, 2021; Wright & Conlon, 2009). As Richardson et al. (2004) argue, there is a possibility of an association between an auditory processing deficit and the well-documented phonological processing deficit or that more comprehensive language processing is compromised in readers with dyslexia. Table 2 shows that the studies that included auditory modality included a phonological stimulus, except for Silva et al. (2016). Another possibility is the argument that visual and orthographic inspection during a phonological stimuli task results in bilateral activation (Rugg, 1984b). Most of the studies that had phonological stimuli used an auditory modality. Perhaps because it was an auditory mode of analysis instead of a visual one, we did not see an aberrant N400 response.

Moreover, difficulty in lexical decision tasks could be indicative of difficulty in lower-level processing of decoding words (Coch & Holcomb, 2003). From this view, a core difficulty in phonological awareness contributes to challenges in Grapheme-Phoneme Correspondence (GPC) processing, which may then be linked to delayed semantic processing. Coch and Holcomb (2003) suggest that the N400 does not exclusively represent phonological or semantic processing, where sequential and interactive information of word-level processes (e.g., decoding, sight word recognition), followed by the activation and selection of meaning in the mental lexicon culminate in comprehension. Instead, Coch and Holcomb (2003) propose that the N400 reflects an interactive and dynamic system that supports meaning processing. Hence, difficulty in phonological stimuli task is linked to a smaller N400 amplitude. However, more studies exploring the associations between reading measures and aspects of N400 are needed to better understand the literacy-N400 relationship.

The behavioral results showed that participants with dyslexia in all studies were slower in reaction time and less accurate. Individuals with dyslexia have been shown to read more slowly than typically developing readers (Christodoulou et al., 2014). The timed nature of the response task might explain the difference in mean in the N400 amplitude. As Tamm et al. (2014) argued, reaction time variability may reflect impairments in information processing.

Regarding task stimuli, we coded the articles into studies using semantically congruent and incongruent sentences as moderators. The remaining stimuli tasks included phonological, lexical decision, and orthographic tasks coded as “pseudowords/characters and “phonological stimuli.” The meta-analysis showed that individuals with RD have more difficulty reading and understanding semantically incongruent sentences $g = 0.948$ and “pseudoword task” $g = 0.692$. As Kutas and Van Petten (1994) say, this suggests that the N400 is most robust during sentence processing. The N400 in this meta-analysis also showed that the N400 is large in the pseudowords/characters task. This might suggest that the N400 is strong in a task that requires some form of semantic processing, as discussed earlier in the introduction.

Age was a significant moderator, with effect sizes of $g = 0.8$ evident in children and $g = 0.531$in adults. The N400 pattern is aberrant in dyslexia populations, but while remaining significant, the effect size decreased with age. As the children get older, they might have learned ways to compensate for their difficulty or may have received remediation in reading sufficient to process the isolated words and short sentences generally used in N400 experimental tasks.

Opacity and shallowness of the language of study delivery were also moderators. Here, both groups showed significant effects but with no difference in the effect size between opaque ($g = 0.76$) and shallow ($g = 0.78$) languages, yet morpho-syllabic Chinese character reading tasks that are strongly associated with semantic processing (e.g., Liu et al., 2013; Perfetti & Liu, 2005), had the highest effect size in all the moderation analyses ($g = 1.043$). This result is interesting as it shows that the N400 could be a reliable index for semantic processing, strongly implicated in non-alphabetic written language systems.

The results showed that the studies that reported a difference in amplitude also reported a difference in latency, except for two studies (Hasko et al., 2017; Schulz et al., 2017) that did not find any difference in latency. Previous research highlighted that a latency delay could reflect a delay in information processing (Grillion et al., 1991). However, it is difficult to conclude latency in this paper because of the insufficient data we had to run the meta-analysis on latency differences.

Only five studies reported topographical results of their studies, and five showed topographical differences between typically developing readers and readers with dyslexia. In typically developing readers, topographical differences suggest involvement of distinct sources in congruent and incongruent conditions (Jost et al., 2014). The reported results of these five studies all showed topographic differences, including either stronger anterior negativity, widespread activity in the right hemisphere, or more negativity in the frontal regions. The reliance on the frontal areas suggests that readers with dyslexia are putting more effort into semantic retrieval. This is a common finding among individuals with reading comprehension difficulties (Kronschnabel et al., 2014; Wilson et al., 2012). However, we suggest interpreting this observation with caution as only a few studies reported topographical results in this paper. This is another variable we could not meta-analytically assess because of insufficient data.

Regarding reaction time and accuracy rate, the results showed that participants with dyslexia were slower and less accurate. The meta-analysis of the behavioral data suggests that individuals with dyslexia have difficulty processing semantic and pseudoword/characters, and integrating information in sentence contexts which are consistent with a speed processing deficit in dyslexia (Catts et al., 2002).

Limitations

This study has significant limitations, with some reflecting the state of the current literature in the N400 comparative studies between typically developing readers and readers with dyslexia. As noted earlier, a critical issue in all 20 selected studies concerned internal validity, wherein none used blinding of assessment nor reported attrition. None of the articles established the task reliability or validity, and none of the studies said how they controlled for extraneous variables. This suggests that caution needs to be attached to existing data and that better-controlled work is required to advance this field in the future. Another primary concern is the modest-sized literature in this field. One purpose of this systematic review and meta-analysis was to include unpublished studies that might otherwise be ‘grey literature’ (e.g., Ph.D. dissertations, conference papers, conference talks that fit the inclusion criteria) and published peer-reviewed papers. However, the results yielded only 20 peer-reviewed articles and no broader unpublished works, suggesting that more basic research is needed in this research area. All studies have a relatively small sample size, with some as small as $n = 7$ in each group. This raises the issue that interpreting our reported difference in means in the N400 amplitude between typical and dyslexia readers is sufficiently strong to conclude that the N400 is quite aberrant in dyslexia readers when considering the literature. Only five studies included some brief reports on topography analysis and latency differences. Eleven studies did not report any latency results. Hence, we could not run any meta-analysis on these two variables because of the lack of sufficient data. These facts may limit the generalizability of the results.

We initially wished to explore how the dyslexia population performed on both semantically congruent and semantically incongruent sentences. Studies here did not include the data for the congruency effect; only the relative performance on the incongruent ones was reported. Additionally, while analyses showed a result of age, this may need to be interpreted cautiously across all age phases: For example, only two studies focused on adolescents. Additionally, among participants from school- and university-aged populations, there is a lack of information about their academic performance or their current or previous exposure to reading interventions, so its impact on outcome generally is unknown. One included paper was an intervention study (Hasko et al., 2014). Here, individuals improved reading, and reading performance was associated with an enhanced N400. This suggests that the N400 is potentially a measure of brain malleability, sensitive to pre-post-intervention improvements in word reading. Only one study has evaluated N400 through intervention, and further work is warranted on targeted interventions to establish the dynamic role of N400. Another limitation is that all included studies involved participants matched on chronological age rather than on other reading level matches. We are thus unable to know whether the difficulties in the dyslexia group as a consequence of different reading levels between the two groups are simply or whether it is a central or potentially causal characteristic of their dyslexia group. Had the included studies used reading-matched control and chronological aged-match controls we would have better understood the dyslexia group's distinctive disability characteristic even when reading at the same word level as a typical reader (see Rack et al., 1992 for review). The current meta-analysis compared studies that generally differed in terms of the N400 stimuli task (i.e., semantic incongruency, phonological stimuli task, lexical decision task). This suggests that there is an N400 effect in language tasks that require a semantic process at some level. This also suggests a limitation in our results interpretation of the characteristics of N400 in readers with dyslexia. The field is still missing research that systematically tests the N400 using the same paradigm across all age groups and individual differences.

Conclusion and Future Directions

The current paper explored the research question, what is the difference in mean in the N400 amplitude in typically developing readers and readers with dyslexia? (ii) Do ERP modality, stimulus task, age, and experiment language moderate the difference in the N400 amplitude between typically developing readers and readers with dyslexia? The meta-analysis showed that the difference in mean of the N400 amplitude between neurotypicals and readers with dyslexia readers is 0.66, a large ES. The research in this paper on the N400 offers insights that may help us understand the neurobiological differences in how dyslexia readers process words and sentences. Our data from the significant difference in means between semantically congruent and incongruent sentence task stimuli for typical versus dyslexia readers (g = .948) and the somewhat larger effects shown in the morpho-syllabary of Chinese together suggests that the N400 is, at least to an extent, an index of semantic processing in the brain. From this evidence of a consistent difference in N400 amplitude across 20 studies, we may now have reliable neurobiological evidence that individuals with reading difficulties experience difficulty in semantic processing. Pedagogically such needs require attention in principled, effective interventions.

Future research is needed to understand whether the orthographic, phonological, or semantic deficits are associated with N400 differences in the brain or because of the neuro-diverse profile of those with dyslexia reading profiles. The findings also showed that the studies were of “Medium” quality, for there were some critical methodological issues, including the absence of reporting extraneous variables, blinding, selecting participants and iteration. In the future, researchers should address the methodological limitations found in the current results and include more studies with both reading-age- and chronological-age matches. Currently, no research focuses on studies that use the N400 as a dependent measure in a reading intervention or even describe the previous intervention type and duration history. An exciting finding from Hasko et al. (2014) is that the N400 was enhanced post-intervention with those who showed improvement in reading. This is another potential validation of the sensitivity of the N400 not only to the lexical-semantic context of language but also to the orthographic processes of language. More quality behavioural intervention studies linked to N400 profiles are now required.

Based on a literature review, we chose the main moderators that play a critical role in N400 amplitude in dyslexia readers. However, other moderators related to the neurodiverse profiles of participants included in the studies may also play a crucial role in N400 responses. We excluded studies with participants with comorbid disorders such as ADHD and Autism, considering confounding variables. Future studies could include participants with (and without) comorbid conditions to understand whether profiles other than those related to language abilities play a role in the N400 amplitude in dyslexia readers.

References

*References marked with an asterisk indicate studies included in the systematic review

Aboud, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental Science*, *19*(4), 632-656. https://doi.org/10.1111/desc.12422

*Aiello, G., Merati, S., & Toti, M. P. and Aiello I (2018) Predictive Value of the Auditory and Visual Evoked Potentials Using the Wavelet Method and Visual n400 in Subjects Affected by Dyslexia and in Normal Subjects. *Dual Diagn Open Acc*, *3*(3), 7. doi: 10.21767/2472-5048.100040

*Araújo, S., Faísca, L., Reis, A., Marques, J. F., & Petersson, K. M. (2016). Visual naming deficits in dyslexia: An ERP investigation of different processing domains. *Neuropsychologia*, *91*, 61-76. https://doi.org/10.1016/j.neuropsychologia.2016.07.007

Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, *26)*1, 49-66. doi:10.1598/rrq.26.1.3

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the phonological matching of picture names. *Brain and Language*, *38*(3), 424–437. https://doi.org/10.1016/0093-934X(90)90125-Z

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, https://doi.org/10.2307/2533446

Bergmann, J., Hutzler, F., Klimesch, W., & Wimmer, H. (2005). How is dysfluent reading reflected in the ERP? *Journal of Neurolinguistics*, *18*(2), 153−165. https://doi.org/10.1016/j.jneuroling.2004.11.004

*Bonte, M. L., & Blomert, L. (2004). Developmental dyslexia: ERP correlates of anomalous phonological processing during spoken word recognition. *Cognitive Brain Research, 21*(3), 360-376. https://doi.org/10.1016/j.cogbrainres.2004.06.010

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature*, *301*(5899), 419. doi:10.1038/301419a0

*Brandeis, D., Vitacco, D., & Steinhausen, H.-C. (1994). Mapping brain electric micro-states in dyslexic children during reading. *Acta Paedopsychiatrica: International Journal of Child & Adolescent Psychiatry*.

Castles, A., Wilson, K., & Coltheart, M. (2011). Early orthographic influences on phonemic awareness tasks: Evidence from a preschool training study. *Journal of Experimental Child Psychology*, *108*(1), 203-210. https://doi.org/10.1016/j.jecp.2010.07.006

Catts, H. W., Gillispie, M., Leonard, L. B., Kail, R. V., & Miller, C. A. (2002). The role of speed of processing, rapid naming, and phonological awareness in reading achievement. *Journal of learning disabilities*, *35*(6), 510-525. https://doi.org/10.1177/00222194020350060301

*Chung, K. K., Tong, X., & McBride-Chang, C. (2012). Evidence for a deficit in orthographic structure processing in Chinese developmental dyslexia: an event-related potential study. *Brain Research, 1472*, 20-31. https://doi.org/10.1016/j.brainres.2012.06.010

Coch, D., & Benoit, C. (2015). N400 Event-related potential and standardized measures of reading in late elementary school children: Correlated or independent? *Mind, Brain, and Education*, *9*(3), 145–153. https://doi.org/10.1111/mbe.12083

Coch, D., & Holcomb, P. J. (2003). The N400 in beginning readers. *The Journal of the International Society for Developmental Psychobiology*, *43*(2), 146-166. https://doi.org/10.1002/dev.10129

Cohen, J. (1988). Statistical power analysis for the behavioral sciences [Internet]. Hillsdale NJ: L.

Coles, M. G. H., & Rugg, M. D. (1995). Event-related brain potentials: An introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Oxford psychology series, No. 25. Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 1-26). New York, NY, US: Oxford University Press.

Christodoulou, J. A., Tufo, S. N. D., Lymberis, J., Saxler, P. K., Ghosh, S. S., Triantafyllou, C., Whitfield-Gabrieli, S., & Gabrieli, J. D. E. (2014). Brain Bases of Reading Fluency in Typical Reading and Impaired Fluency in Dyslexia. *PLOS ONE*, *9*(7), e100552. https://doi.org/10.1371/journal.pone.0100552

Curran, T., Tucker, D. M., Kutas, M., & Posner, M. I. (1993). Topography of the N400: Brain electrical activity reflecting semantic expectancy. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(3), 188–209. https://doi.org/10.1016/0168-5597(93)90004-9

Desroches, A. S., Newman, R. L., Robertson, E. K., & Joanisse, M. F. (2013). Electrophysiological indices of phonological impairments in dyslexia. *Journal of Speech,* *Language, and Hearing Research*, *56*(1), 250-264. doi:10.1044/1092-4388(2012/10-0351)

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, *52*(6), 377-384. doi:10.1136/jech.52.6.377

Dumay, N., Benraiss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, *13*(1), 121-143. https://doi.org/10.1162/089892901564117

Duta, M. D., Styles, S. J., & Plunkett, K. (2012). ERP correlates of unexpected word forms in a picture-word study of infants and adults. *Developmental Cognitive Neuroscience*, *2*(2), 223–234. https://doi.org/10.1016/j.dcn.2012.01.003

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757-1761. doi:10.1097/wnr.0b013e328318f014

Friedrich, M., & Friederici, A. D. (2005). Semantic sentence processing reflected in the event-related potentials of one- and two-year-old children. *NeuroReport*, *16*(16), 1801–1804. https://doi.org/10.1097/01.wnr.0000185013.98821.62

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 104–115. https://doi.org/10.1037/0096-1523.13.1.104

Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, *22*(2), 213-228. https://doi.org/10.1080/02671520701296189

Hagtvet, B.E. (1997). Phonological and linguistic-cognitive precursors of reading abilities. *Dyslexia 3*(3), 163-177. https://doi.org/10.1002/(SICI)1099-0909(199709)3:3%3C163::AID-DYS83%3E3.0.CO;2-0

Hagtvet, B. E. (2003). In poor decoders, listening comprehension and reading comprehension: Evidence for the importance of syntactic and semantic skills and phonological skills. *Reading and Writing*, *16*(6), 505-539. Doi:10.1023/a:1025521722900

*Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2013). The time course of reading processes in children with and without dyslexia: An ERP study. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00570

*Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2014). What does the brain of children with developmental dyslexia tell us about reading improvement? ERP evidence from an intervention study. *Frontiers in Human Neuroscience, 8*, 441. Doi:10.3389/fnhum.2014.00441

Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language*, *117*(2), 88-99. Doi:10.1016/j.bandl.2010.12.003

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*(1), 47-61. https://doi.org/10.1111/j.1469-8986.1993.tb03204.x

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain Research*, *1180*, 39–58. https://doi.org/10.1016/j.brainres.2007.06.110

Holcomb, P. J., & Neville, H. J. (1990). Auditory and Visual Semantic Priming in Lexical Decision: A Comparison Using Event-related Brain Potentials. *Language and Cognitive Processes*, *5*(4), 281–312. https://doi.org/10.1080/01690969008407065

Hulme, C., & Snowling, M. (2009). Developmental disorders of language, learning and cognition. Chichester, England: Wiley-Blackwell.

Grillon, C., Ameli, R., & Glazer, W. M. (1991). N400 and semantic categorization in schizophrenia. *Biological Psychiatry*, *29*(5), 467–480. https://doi.org/10.1016/0006-3223(91)90269-R

Jednoróg, K., Marchewka, A., Tacikowski, P., & Grabowska, A. (2010). Implicit phonological and semantic processing in children with developmental dyslexia: Evidence from event-related potentials. *Neuropsychologia*, *48*(9), 2447-2457. https://doi.org/10.1016/j.neuropsychologia.2010.04.017

Jost, L. B., Eberhard-Moscicka, A. K., Frisch, C., Dellwo, V., & Maurer, U. (2014). Integration of spoken and written words in beginning readers: a topographic ERP study. *Brain topography*, *27*, 786-800. https://doi.org/10.1007/s10548-013-0336-4

Kast, M., Elmer, S., Jancke, L., & Meyer, M. (2010). ERP differences of pre-lexical processing between dyslexic and non-dyslexic children. *International Journal of Psychophysiology*, *77*, 59-69. doi:10.1016/j.ijpsycho.2010.04.003

Kemény, F., Banfi, C., Gangl, M., Perchtold, C. M., Papousek, I., Moll, K., & Landerl, K. (2018). Print-, sublexical and lexical processing in children with reading and/or spelling deficits: An ERP study. *International Journal of Psychophysiology, 130*, 53-62. https://doi.org/10.1016/j.ijpsycho.2018.05.009

Khateb, A., Pegna, A. J., Landis, T., Mouthon, M. S., & Annoni, J. M. (2010). On the origin of the N400 effects: an ERP waveform and source localization analysis in three matching tasks. *Brain Topography*, *23*(3), 311-320.

Khateb, A., Pegna, A. J., Landis, T., Michel, C. M., Brunet, D., Seghier, M. L., & Annoni, J. M. (2007). Rhyme processing in the brain: an ERP mapping study. *International Journal of Psychophysiology*, *63*(3), 240-250. https://doi.org/10.1016/j.ijpsycho.2006.11.001

Kronbichler, M., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., & Wimmer, H. (2006). Evidence for a dysfunction of left posterior reading areas in German dyslexic readers. *Neuropsychologia*, *44*(10), 1822-1832. doi:10.1016/j.neuropsychologia.2006.03.010

Kronschnabel, J., Brem, S., Maurer, U., & Brandeis, D. (2014). The level of audiovisual print–speech integration deficits in dyslexia. *Neuropsychologia*, *62*, 245-261. https://doi.org/10.1016/j.neuropsychologia.2014.07.024

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*(2), 310. doi:10.1037/0022-0663.100.2.310

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kutas, M., & Iragui, V. (1998). The N400 in a semantic categorization task across 6 decades. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *108*(5), 456–471. https://doi.org/10.1016/S0168-5597(98)00023-9

Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics electrified: Event-related potential investigations In Gernsbacher MA (Ed.), Handbook of psycholinguistics.

Landi, N., Frost, S. J., Mencl, W. E., Sandak, R., & Pugh, K. R. (2013). Neurobiological bases of reading comprehension: Insights from neuroimaging studies of word-level and text-level processing in skilled and impaired readers. *Reading & Writing Quarterly*, *29*(2), 145-167. doi:10.1080/10573569.2013.758566

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

Liu, P. D., McBride-Chang, C., Wong, T. T. Y., Shu, H., & Wong, A. M. Y. (2013). Morphological awareness in Chinese: Unique associations of homophone awareness and lexical compounding to word reading and vocabulary knowledge in Chinese children. *Applied psycholinguistics, 34*(4), 755-775. doi:10.1017/s014271641200001x

Luck, S. J. (2014). *An introduction to event related potential technique*. MIT Press: London, England

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14. https://doi.org/10.1007/s11881-003-0001-9

*McPherson, W., Ackerman, P., Holcomb, P., & Dykman, R. (1998). Event-related brain potentials elicited during phonological processing differentiate subgroups of reading disabled adolescents. *Brain and Language, 62*(2), 163-185. doi:10.1006/brln.1997.1893

*McPherson, W. B., & Holcomb, P. J. (1999). McPherson, W. B., & Ackerman, P. T. (1999). A study of reading disability using event-related brain potentials elicited during auditory alliteration judgments. *Developmental Neuropsychology*, *15*(3), 359-378. https://doi.org/10.1080/87565649909540755

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological bulletin*, *138*(2), 322. https://psycnet.apa.org/doi/10.1037/a0026744

*Meng, X., Tian, X., Jian, J., & Zhou, X. (2007). Orthographic and phonological processing in Chinese dyslexic children: An ERP study on sentence reading. *Brain Research, 1179*, 119-130. doi:10.1162/jocn.1997.9.3.318

Montgomery, J. W., Gillam, R. B., & Evans, J. L. (2016). Syntactic versus memory accounts of the sentence comprehension deficits of specific language impairment: Looking back, looking ahead. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1491-1504. https://doi.org/10.1044/2016_JSLHR-L-15-0325

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of memory and language*, *39*(1), 85-101. https://doi.org/10.1006/jmla.1998.2564

Newman, R. L., Connolly, J. F., Service, E., & McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potentials. *Psychophysiology*, *40*(4), 640-647. https://doi.org/10.1111/1469-8986.00065

Nicolson, R. I., & Fawcett, A. J. (1994). Reaction times and dyslexia. *The Quarterly Journal of Experimental Psychology Section A*, *47*(1), 29-48. https://doi.org/10.1080/14640749408401142

Perfetti, C. A., & Liu, Y. (2005). Orthography to phonology and meaning: Comparisons across and within writing systems. *Reading and Writing, 18*(3), 193-210.

Perfetti, C., Yang, C. L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(3), 303-318. https://doi.org/10.1002/acp.1419

Plante, E., Petten, C. V., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia*, *38*(13), 1669–1684. https://doi.org/10.1016/S0028-3932(00)00083-X

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). Developmental dyslexia: A review. *Reading Research Quarterly*, *27*(1), 29-53. http://www.jstor.org/stable/747832?origin=JSTOR-pdf

Radeau, M., Besson, M., Fonteneau, E., & Castro, S. L. (1998). Semantic, repetition and rime priming between spoken words: behavioral and electrophysiological evidence. *Biological psychology*, *48*(2), 183-204. https://doi.org/10.1016/S0301-0511(98)00012-X

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

*Robichon, F., Besson, M., & Habib, M. (2002). An electrophysiological study of dyslexic and control adults in a sentence reading task. *Biological psychology*, *59*(1), 29-53. https://doi-org.proxy3.library.mcgill.ca/10.1016/S0301-0511(01)00118-1

*Rüsseler, J., Becker, P., Johannes, S., & Münte, T. F. (2007). Semantic, syntactic, and phonological processing of written words in adult developmental dyslexic readers: an event-related brain potential study. *BMC neuroscience*, *8*(1), 52. https://doi.org/10.1186/1471-2202-8-52

Rüsseler, J., Probst, S., Johannes, S., & Münte, T. F. (2003). Recognition memory for high-and low-frequency words in adult normal and dyslexic readers: an event-related brain potential study. *Journal of Clinical and Experimental Neuropsychology*, *25*(6), 815-829.
doi:10.1076/jcen.25.6.815.16469

Rugg, M. D. (1984). Event-related potentials and the phonological processing of words and non-words. *Neuropsychologia*, *22*(4), 435-443. https://doi.org/10.1016/0028-3932(84)90038-1

Rugg, M. D. (1984b). Event-related potentials in phonological matching tasks. *Brain and language*, *23*(2), 225-240.https://doi.org/10.1016/0093-934X(84)90065-8

*Sabisch, B., Hahne, A., Glass, E., von Suchodoletz, W., & Friederici, A. (2006). Auditory language comprehension in children with developmental dyslexia: Evidence from Event-related Brain Potentials. *Journal of Cognitive Neuroscience, 18*(10), 1676-1695. doi:10.1162/jocn.2006.18.10.1676

Savage, R., Georgiou, G., Parrila, R., Maiorino, K., Dunn, K., & Burgos, G. (2020). The effects of teaching complex grapheme-phoneme correspondences: Evidence from a dual site cluster trial with at-risk grade 2 students. *Scientific Studies of Reading*, *24*(4), 321-337. https://doi.org/10.1080/10888438.2019.1669607

Schuberth, R.E., Eleimas, P.D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Preparation and Performance, 3*, 27-36.

*Schulz, E., Maurer, U., van der Mark, S., Bucher, K., Brem, S., Martin, E., & Brandeis, D. (2008). Impaired semantic processing during sentence reading in children with dyslexia: Combined fMRI and ERP evidence. *Neuroimage, 41*(1), 153-168. doi:10.1016/j.neuroimage.2008.02.012

Serrano, F., & Defior, S. (2008). Dyslexia speed problems in a transparent orthography. *Annals of dyslexia*, *58*(1), 81. https://doi.org/10.1007/s11881-008-0013-6

Seymour, P. H., Aro, M., Erskine, J. M., & Collaboration with COST Action A8 Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of psychology*, *94*(2), 143-174. https://doi.org/10.1348/000712603321661859

Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584–615.

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry*, *57*(11), 1301-1309. doi:10.1016/j.biopsych.2005.01.043

*Silva, P. B., Ueki, K., Oliveira, D. G., Boggio, P. S., & Macedo, E. C. (2016). Early Stages of Sensory Processing, but Not Semantic Integration, Are Altered in Dyslexic Adults. *Frontiers in psychology*, *7*, 430. https://doi.org/10.3389/fpsyg.2016.00430

Silva-Pereyra, J., Rivera-Gaxiola, M., Fernandez, T., Diaz-Comas, L., Harmony, T., Fernandez-Bouzas, A., ... Marosi, E. (2003). Are poor readers semantically challenged? An event-related brain potential assessment. *International Journal of Psychophysiology, 49*(3), 187-199. doi:10.1016/S0167-8760(03)00116-8

Snowling, M. J. (1996). Annotation: Contemporary Approaches to the Teaching of Reading. *Journal of Child Psychology and Psychiatry*, *37*(2), 139–148. https://doi.org/10.1111/j.1469-7610.1996.tb01385.x

Tamm, L., Epstein, J. N., Denton, C. A., Vaughn, A. J., Peugh, J., & Willcutt, E. G. (2014). Reaction Time Variability Associated with Reading Skills in Poor Readers with ADHD. *Journal of the International Neuropsychological Society : JINS*, *20*(3), 292–301. https://doi.org/10.1017/S1355617713001495

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How word meaning influences word reading. *Current Directions in Psychological Science*, *24*(4), 322-328. https://doi.org/10.1177/0963721415574980

*Tong, X., Chung, K. K. H., & McBride, C. (2014). Two-character Chinese compound word processing in Chinese children with and without dyslexia: ERP evidence. *Developmental Neuropsychology, 39*(4), 285-301. doi:10.1080/87565641.2014.907720

*Tzeng, Y.-L., Hsu, C.-H., Lin, W.-H., & Lee, C.-Y. (2018). Impaired Orthographic Processing in Chinese Dyslexic Children: Evidence From the Lexicality Effect on N400. *Scientific Studies of Reading*, *22*(1), 85–100. https://doi.org/10.1080/10888438.2017.1353996

*Wang, E., Zhou, G., Wang, Y., Huang, H., Hou, Y., & Liu, J. (2017). Word recognition semantic processing defects in Chinese children with developmental dyslexia: An event-related potential study. *NeuroQuantology*, *15*(2). DOI:10.14704/nq.2017.15.2.1066

Wright, C. M., & Conlon, D. E. G. (2009). Auditory and Visual Processing in Children With Dyslexia. *Developmental Neuropsychology*, *34*(3), 330–355. https://doi.org/10.1080/87565640902801882

Wilson, K. R., O'Rourke, H., Wozniak, L. A., Kostopoulos, E., Marchand, Y., & Newman, A. J. (2012). Changes in N400 topography following intensive speech language therapy for individuals with aphasia. *Brain and Language*, *123*(2), 94–103. https://doi.org/10.1016/j.bandl.2012.06.005

Wimmer, H., & Schurz, M. (2010). Dyslexia in regular orthographies: Manifestation and causation. *Dyslexia*, *16*(4), 283-299. doi:10.1002/dys.411

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied psycholinguistics*, *14*(1), 1-33. https://doi.org/10.1017/S0142716400010122

Yeung, S. S. S., & Savage, R. (2020). Teaching grapheme–phoneme correspondences using a direct mapping approach for at-risk second language learners: A randomized controlled trial. *Journal of Learning Disabilities*, *53*(2), 131-144. https://doi.org/10.1177/0022219419894563

Young, C., Mohr, K. A., & Rasinski, T. (2015). Reading together: A successful reading fluency intervention. *Literacy Research and Instruction*, *54*(1), 67-81. https://doi.org/10.1080/19388071.2014.976678

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*(1), 25-64. doi:10.1016/0010-0277(89)90013-9

Appendix A

Table 1.1

Diagnostic criteria in the included studies

Study	Diagnostic criteria for RD
Aiello et al. (2018)	RD group had a formal diagnosis of Dyslexia
Araújo et al. (2016)	RD group had a formal diagnosis of Dyslexia - Read below level in frequency and pseudowords test done by the researchers
Bonte & Blomert (2004)	RD group had a formal diagnosis of Dyslexia – Read below level in one-minute reading test and non-word reading test done by the researchers
Brandeis et al. (1993)	RD group were referred by school psychologists and reading specialists based on their formal diagnosis.
Chung et al. 2012	RD group had a formal diagnosis of Dyslexia
Desroches et a. (2011)	RD group were identified with Dyslexia based on scoring below 15th percentile on Woodcock Reading Mastery Test and on Woodcock Word Identification Test
Hasko et al. (2013)	RD group had a formal diagnosis of Dyslexia
Hasko et al. (2014)	RD group had a formal diagnosis of dyslexia. RD group scored below the 25th percentile on 1-minute fluent reading test, common and pseudo word, as well as reading comprehension test.
Jednorog et al. (2010)	RD had a formal diagnosis of Dyslexia based on Polish Language Battery Tests.
McPherson et al. (1998)	RD group had a formal diagnosis of Dyslexia and read two years below grade level based on Woodcock Reading Mastery Test
McPherson & Ackerman (1999)	RD group had a formal diagnosis of Dyslexia and read two years below grade level based on Woodcock Reading Mastery Test
Meng et al. (2007)	RD group had a diagnosis of Dyslexia based on reading below grade level on Reading fluency test and vocabulary test.
Rüssler et al. (2007)	RD group had a formal diagnosis of Dyslexia
Robichon et al. (2002)	RD group had a formal diagnosis of Dyslexia and read below level in comparison with TD on Reading Mastery Test
Sabisch et al. (2006)	RD group had a formal diagnosis of Dyslexia
Schulz et al. (2008)	Formal diagnosis of Dyslexia
Silva et al. (2016)	RD group had a formal diagnosis with Dyslexia and performed below level on Phonological Awareness Test and Word Reading Competence Test
Tong et al. (2014)	RD group had a formal diagnosis of Dyslexia
Tzeng et al. (2017)	RD group had a formal diagnosis of Dyslexia
Wang et al. (2017)	RD group had a formal diagnosis of Dyslexia

Table 1.2

Characteristics of the 20 Included Studies

Author (date)	Study	Participants n (TD)	Participants n (RD)	Age	Language	ERP modality	Stimuli	Findings
Aiello et al. (2018)	Quasi-experiment	13	13	10-14	Italian	Visual	Semantically congruent and incongruent sentences	TD = typical RD = atypicalN400 (latency)
Araujo et al. (2016)	Quasi-experimental	17	18	23	Portuguese	Visual	Phonological and semantic stimuli	TD = typical N400 RD= [2]atypical N400
Bonte & Blomert (2004)	Quasi-experimental	8	10	8	Dutch	Auditory	Phonological stimuli	TD = typical N400 RD = typical N400
Brandeis, et al. (1993)	Quasi-experimental	12	12	10-12	German	Visual	Semantically congruent and incongruent sentences	TD = typical N400 RD = *atypical N400
Chung, et al. (2012)	Quasi-experimental	11	12	8-10	Chinese	Visual	Pseudo & real characters	TD = typical N400 RD = lack of N400
Desroches et al. (2013)	Quasi-experimental	15	14	8-11	English Canadian	Auditory	Matched trials of spoken words with pictures	TD = typical N400 RD = atypical N400 But typical in phonologically unrelated matches
Hasko et al. (2014)	Quasi-experimental Intervention was	25	28	10-13	German	Visual	Phonological stimuli	TD = typical N400 RD = increased N400 post intervention for improvers. atypical for non-improvers.

[2] Atypical: small N400

	randomly assigned							
Hasko et al. (2013)	Quasi-experimental	29	52	10-13	German	Visual	Phonological stimuli	TD = typicalN400 RD = atypical N400
Jednörog et al. (2010)	Quasi-experimental	18	18	10	Polish	Auditory	Semantic and phonological priming task	TD = typical N400 RD = typical N400 ONLY in semantic priming and atypical phonological priming
McPherson et al. (1998)	Quasi-experimental	16	16	13-18	English US	Auditory	Phonological Stimuli	TD = typicalN400 RD Dysphonetic = atypicalN400 RD Phonetic = typical N400
McPherson & Ackerman (1999)	Quasi-experimental	16	16	13-18	English US	Auditory	Phonological Stimuli	TD = typical N400 RD Dysphonetic = atypical N400 RD Phonetic = typical N400
Meng et al. (2007)	Quasi-experimental	13	14	10 yrs.	Chinese	Visual	Semantically congruent and incongruent sentences	TD = typicalN400 RD = Lack of N400
Robichon (2002)	Quasi-experimental	12	12	23	German	Visual	Semantically congruent and incongruent sentences	TD = typical N400 RD= atypical N400
Rüsseler et al. (2007)	Quasi-experimental	11	11	24.9	German	Visual	Rhyme judgement task (Phonological stimuli) semantic judgement task, and gender judgement task	TD = Typical N400 RD = atypical N400 in rhyme judgment and gender judgment

Sabisch et al. (2006)	Quasi-experimental	16	16	9-12	German	Auditory	Semantically congruent and incongruent sentences	TD = typical N400 RD = typical N400
Silva et al. (2016)	Quasi-Experimental	18	14	19-41	Portuguese	Auditory	Semantically congruent and incongruent sentences	TD = typical N400 RD= there was an N400 effect, but preserved
Schulz et al. (2008)	Quasi-experimental design	31	16	11.5	Switz/not specified	Visual	Semantically congruent and incongruent sentences	TD = typical N400 RD = atypicalN400
Tong et al. (2014)	Quasi-experimental	10	15		Chinese	Visual	Pseudo & real characters	TD = typical N400 RD = lack of N400
Tzeng et al. (2018)	Quasi-experimental design	23	23	9-10	Chinese	Visual	semantic categorization task	TD= typical N400 RD =lack of N400
Wang et al. (2017)	Quasi-experimental design	18	18	12-14	Chinese	Visual	Pseudo & real characters	TD = Typical N400 RD =lack of N400

TD = Typically developing; RD= Dyslexia

Table 1.3

Modified Downs and Black (1998) Checklist for assessing the quality of studies.

Author/date	Reported method of allocation[3]	Detailed characteristics of subjects included in the study	Intention to treat analysis	Sample size justification	Main outcomes clearly described	Evidence of establishing validity and reliability
Aiello et al. (2018)	Yes	Yes	No	No	Yes	No
Araujo et al. (2016)	Yes	Yes	No	No	Yes	No
Bonte & Blomert (2004)	Yes	Yes	No	No	Yes	No
Brandeis et al. (1993)	Yes	Yes	No	No	Yes	No
Chung et al. (2012)	Yes	Yes	No	No	Yes	No
Desroches et al. (2013)	Yes	Yes	No	No	Yes	No
Hasko et al. (2014)	Yes	Yes	No	No	Yes	No
Hasko et al. (2013)	Yes	Yes	No	No	Yes	No
Jednörog et al. (2010)	Yes	Yes	No	No	Yes	No
McPherson et al. (1998)	Yes	Yes	No	No	Yes	No
McPherson et al. (1999)	Yes	Yes	No	No	Yes	No
Meng et al. (2007)	Yes	Yes	No	No	Yes	No
Russler et al. (2007)	Yes	Yes	No	No	Yes	No
Robichon et al. (2002)	Yes	Yes	No	No	Yes	No
Sabisch et al. (2006)	Yes	Yes	No	No	Yes	No
Silva et al. (2016)	Yes	Yes	No	No	Yes	No
Schulz et al. (2008)	Yes	Yes	No	No	Yes	No
Tong, et al. (2014)	Yes	Yes	No	No	Yes	No
Tzeng et al. (2018)	Yes	Yes	No	No	Yes	No
Wang et al. (2017)	Yes	Yes	No	No	Yes	No

[3] Studies reported method of allocating subjects, but no randomization took place since none of the studies are randomized control except for Hasko et al. (2013)

Table 1.4

WOE (Weight of Evidence)

Author/Date	WOE A	WOE B	WOE C	WOE D
Aiello et al. (2018)	High	Medium	Medium	Medium
Arjuo et al. (2016)	High	Medium	Medium	Medium
Bonte & Blomert (2004)	High	Medium	Medium	Medium
Brandeis et al. (1993)	High	Medium	Medium	Medium
Chung et al. (2012)	High	Medium	Medium	Medium
Desroches et al. (2013)	High	Medium	Medium	Medium
Hasko et al. (2014)	High	Medium	Medium	Medium
Hasko et al. (2013)	High	Medium	Medium	Medium
Jednörog et al. (2010)	High	Medium	Medium	Medium
McPherson et al. (1998)	High	Medium	Medium	Medium
McPherson et al. (1999)	High	Medium	Medium	Medium
Meng et al. (2007)	High	Medium	Medium	Medium
Rüssler et al. (2007)	High	Medium	Medium	Medium
Robichon et al. (2002)	High	Medium	Medium	Medium
Sabisch et al. (2006)	High	Medium	Medium	Medium
Silva et al. (2016)	High	Medium	Medium	Medium
Schulz et al. (2008)	High	Medium	Medium	Medium
Tong, et al. (2014)	High	Medium	Medium	Medium
Tzeng et al. (2018)	High	Medium	Medium	Medium
Wang et al. (2017)	High	Medium	Medium	Medium

Table 1.5

Behavioral Data (reaction time and accuracy) in the 20 included studies

Study	**Reaction time**	**Accuracy**
Aiello et al. (2018)	NA	NA
Arjuo et al. (2016)	No difference between groups	Higher error rates for the dyslexia group
Bonte & Blomert (2004)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group

Brandeis et al. (1993)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Chung et al. (2012)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Desrochers et al. (2013)	Slower rate for the dyslexia group	Similar accuracy rates for both groups
Hasko et al. (2014)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Hasko et al. (2013)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Jednörog et al. (2010)	NA	NA
McPherson et al. (1998)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
McPherson et al. (1999)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Meng et al. (2007)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Rüssler et al. (2007)	Slower rate for the dyslexia group	Similar accuracy rate for both groups
Robichon et al. (2002)	NA	NA
Sabisch et al. (2006)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Silva et al. (2016)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Schulz et al. (2008)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Tong, et al. (2014)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group
Tzeng et al. (2018)	NA	NA
Wang et al. (2017)	Slower rate for the dyslexia group	Higher error rates for the dyslexia group

Figure 1.1

Prisma Chart

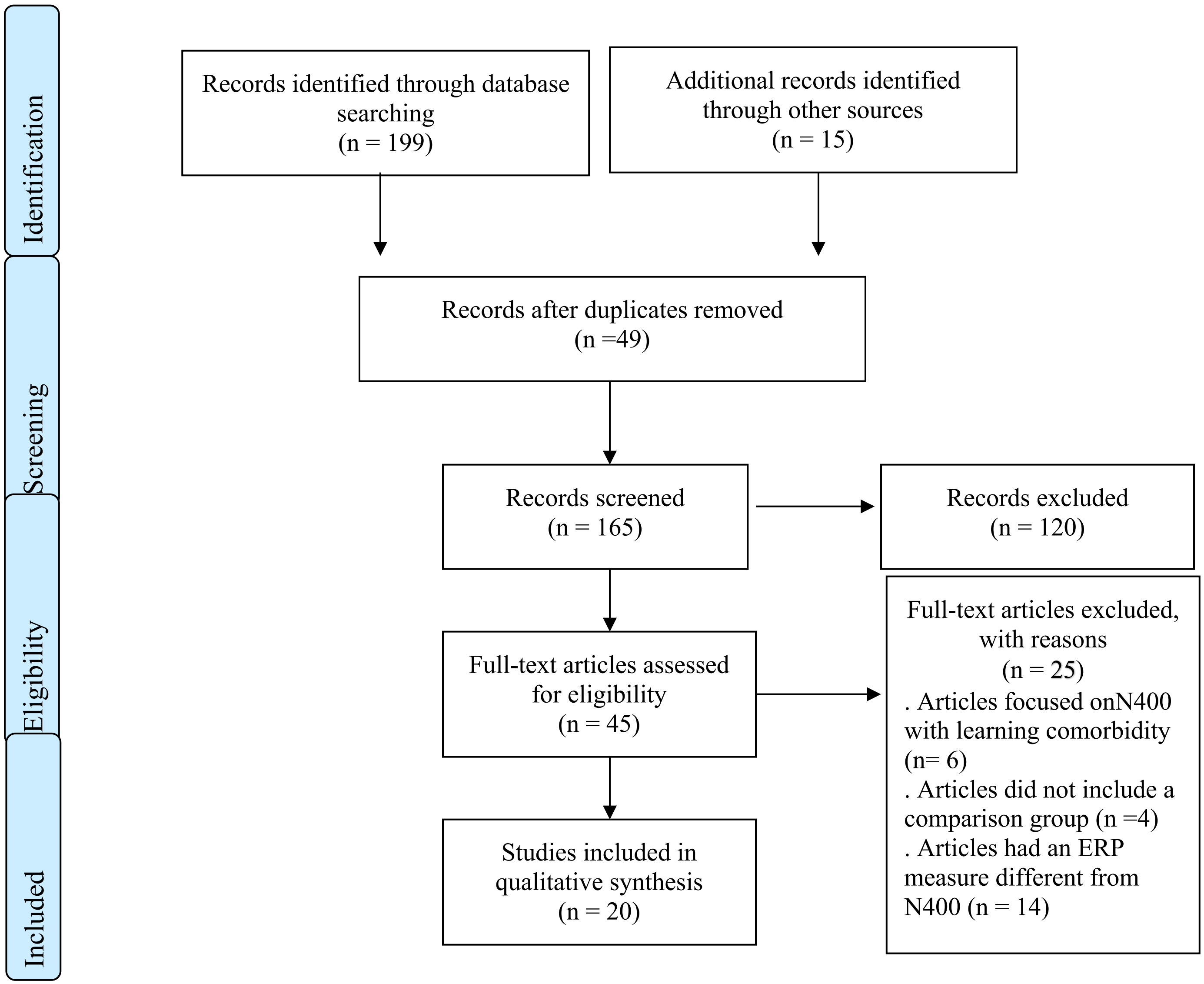
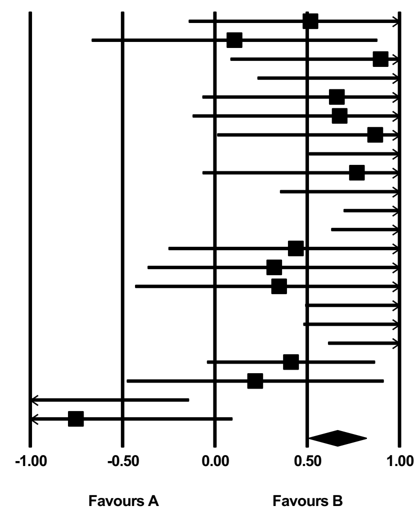


Figure 1.2

Meta-analysis Statistics for Each Study

Model	Study name	Statistics for each study						
		Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Araujio et al. (2016)	0.518	0.336	0.113	-0.141	1.177	1.540	0.123
	Bonte & blomert (2004)	0.106	0.394	0.156	-0.667	0.879	0.269	0.788
	Brandeis et al. (1993)	0.898	0.415	0.172	0.085	1.711	2.164	0.030
	Chung et al. 2012	1.080	0.433	0.187	0.231	1.928	2.495	0.013
	Desrochers et a. (2011)	0.661	0.371	0.138	-0.067	1.389	1.779	0.075
	Jednorog et al. (2010)	0.676	0.406	0.165	-0.120	1.472	1.664	0.096
	McPherson & Ackerman (1999)	0.869	0.436	0.191	0.013	1.724	1.991	0.047
	Meng et al. (2007)	1.284	0.395	0.156	0.510	2.058	3.252	0.001
	Russeler (2007)	0.770	0.426	0.182	-0.066	1.605	1.805	0.071
	Robichon (2002)	1.197	0.430	0.185	0.354	2.041	2.782	0.005
	Sabisch et al.	1.462	0.390	0.152	0.698	2.227	3.748	0.000
	Schulz et al. (2008)	1.278	0.330	0.109	0.631	1.925	3.872	0.000
	Silva et al. (2016) Visual priming	0.439	0.352	0.124	-0.251	1.128	1.247	0.212
	Silva et al. (2016) Auditory priming	0.322	0.350	0.122	-0.364	1.007	0.920	0.358
	Tong et al. (2014)	0.348	0.398	0.158	-0.432	1.128	0.875	0.382
	Tzeng et al. (2017)	1.101	0.312	0.097	0.490	1.712	3.531	0.000
	Wang et al. (2017)	1.173	0.354	0.125	0.479	1.867	3.314	0.001
	Aiello et al. (2018)	1.495	0.449	0.202	0.614	2.375	3.326	0.001
	Hasko et al. (2013)	0.412	0.232	0.054	-0.042	0.867	1.779	0.075
	Hasko et al. (2014)	0.218	0.355	0.126	-0.477	0.914	0.616	0.538
	McPherson et al. (1998) Visual	-1.009	0.443	0.196	-1.876	-0.141	-2.278	0.023
	McPherson et al. (998) Auditory	-0.753	0.432	0.187	-1.599	0.094	-1.743	0.081
Fixed	Pooled	0.663	0.079	0.006	0.509	0.818	8.410	0.000

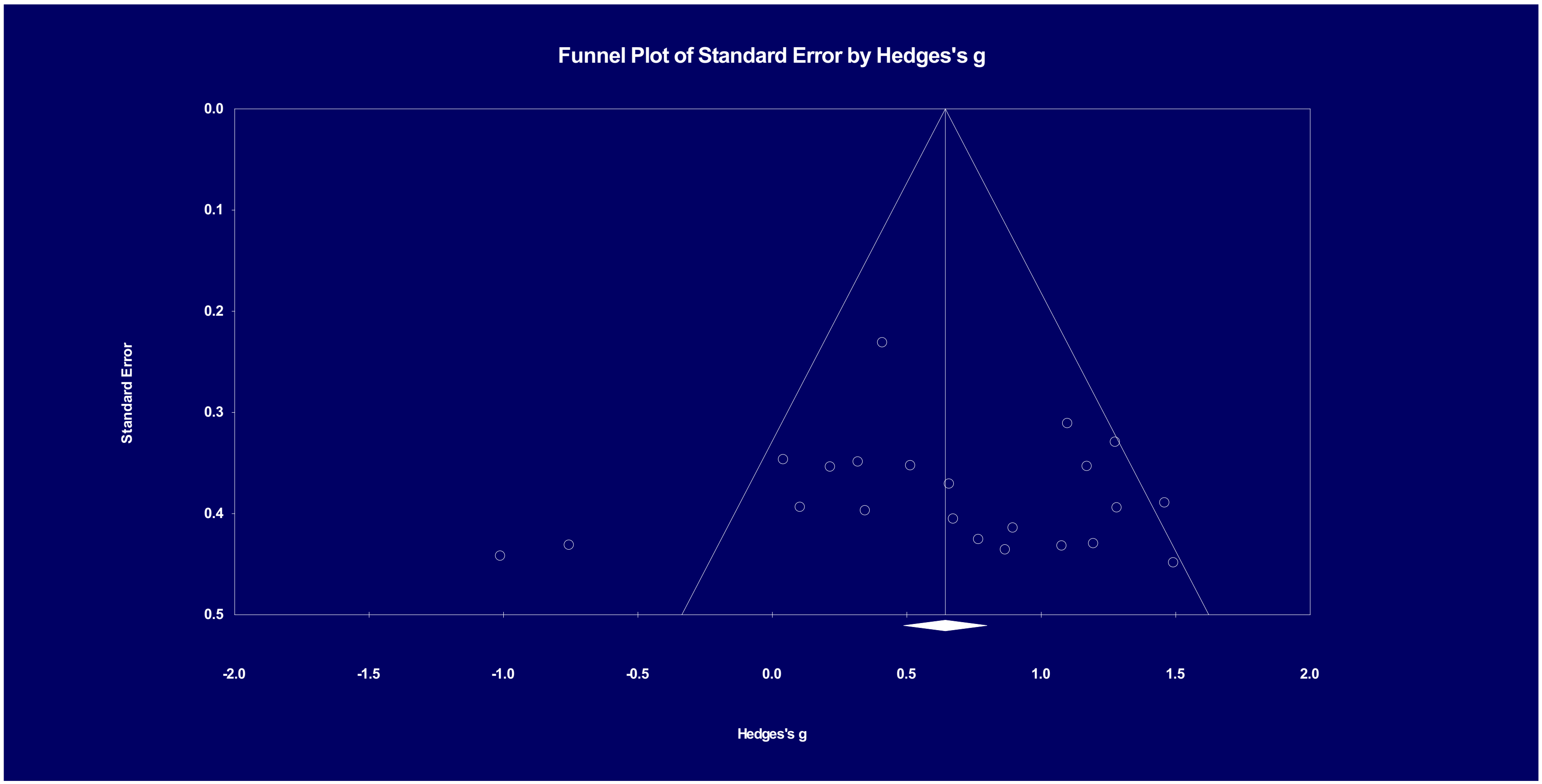
Hedges's g and 95% CI



Meta Analysis

Figure 1.3

Funnel Plot of 20 Studies in the Meta-analysis



Chapter 5. Bridging Manuscript 1 and Manuscript 2

In Chapter 4, Manuscript 1, I synthesized and ran a statistical analysis to explore whether the N400 is an index of semantic incongruity and reflects a reading disability. I ran a systematic review and a meta-analysis. The results showed that the N400 is different between TD readers and RD. Individuals with RD showed an aberrant N400 amplitude (small or no N400) with late latency and with different topographical activity reflecting the difficulty. The N400 effect was also most different with sentence incongruity and visual stimuli and language effects were evident with stronger patterns in Chinese than in alphabets. These patterns support claims that N400 indexes lexical-semantic processing. One limitation of the studies identified for the review is that the samples are fourteen years of age or older, specifically studies conducted in English. Such studies speak less clearly to acquisition where SfV might be expected to be maximally influential. To advance understanding a study of children below the age of ten years is needed, so is reported in the next section. This study also explores matched RD and TD group differences in SfV.

The conclusions of the meta-analysis and review also helped me in moving forward to the second part of my thesis and exploring whether behavioral and ERP measures may be sensitive to different stages of information processing in sentence reading and exploring N400 and SfV in conjunction. Three studies were located so far that addressed the correlation between reading measures and N400. There first was by Henderson et al. (2011) and the second was by Coch and Benoit (2015). However, both papers explored the correlation in a typical population only. A third study, Sun et al. (2023) included a group of children with RD to their study, but their results did not indicate whether there is a difference in correlations between TD and RD in the N400 and reading measures. Therefore, the purpose of the study in Chapter 6 (Manuscript 2) is to compare ERP (N400) and behavioral data of SfV, word reading, phonological awareness, vocabulary, and comprehension to understand the dynamics of reading processes in young children with dyslexia compared to age-matched typical readers. The next step is to further explore the correlation between the N400 and SfV, word reading and reading comprehension in younger children with dyslexia in comparison with typical readers to have a better understanding of the relationship between ERP and behavioral measures and to establish a rationale for running an SfV reading intervention linked to N400 measures subsequently reported in Chapter 8 (Manuscript 3).

Chapter 6. Manuscript 2

This chapter is an exact reproduction of the following article, submitted to *Developmental Neuropsychology*

Reading disability in children:

Exploring the N400 and its correlation with Set-for-Variability

[1] Badriah Basma, [2] Robert Savage, [1] Gigi Luk, [1] Armando Bertone.

[1]Department of Educational and Counselling Psychology
McGill University, Montreal, Quebec, Canada

[2]Faculty of Education,
York University, Toronto, Ontario, Canada

Correspondence to:

Badriah Basma, Department of Education & Counselling Psychology, McGill University, 3700 McTavish St, Montréal, QC, Canada, H3A 1Y2.

Tel: +1 514 398 6655; e-mail: badriah.basma@mail.mcgill.ca

Abstract

Introduction. The N400 event-related potential (ERP), established as an indicator of lexical-semantic processing, can be used to assess the neural underpinnings of reading difficulties. Existing literature has revealed various patterns of correlations between the N400 potentials and both word decoding and broader language comprehension. This study aims to: 1) examine cognitive and neurobiological theories about strategic semantic processes in English reading acquisition, operationalized by N400 components, as well as Set-for-Variability (SfV) profiles in elementary school readers with and without a reading disability (RD); 2) examine individual differences in correlations between N400 components, and Set-for-Variability in these samples. **Method.** Fifty-one 9-year-old children (20 typical developing, TD, readers and 31 children with RD) read semantically congruent versus incongruent sentences while undergoing EEG collection. SfV and standardized literacy tests were administered. **Results.** Children with RD did not exhibit an N400 effect in response to sentence incongruency compared to TD readers. Children with RD also exhibited delays in SfV, compared to TD. As predicted, individual differences in SfV were negatively correlated with N400 latency only in the RD group. Word reading measures consistently demonstrated a positive association, on average, with N400 amplitude and a negative association with peak latency in both TD and RD groups. **Conclusion.** Children with RD exhibit distinct delays in SfV alongside unique N400 component profiles. The N400 effect and latency differences between the two groups indicate variability in SfV, particularly for children with RD.

Keywords: *N400, reading disability, Set-for-Variability (SfV), reading ability, semantics*

Introduction

When learning to read in an alphabetic writing system, novice readers must understand that each word consists of speech sounds (phonemes) that are routinely linked to a specific letter or letter combination (graphemes) to decode words (Ehri 1998; Melhase et al., 2020; Snowling, 1980). However, the English language uses an opaque orthography (Seymour, 2003), which results in many exceptions to the primary grapheme-phoneme correspondence (GPC) rules, exemplified by irregular words such as *bear* and *deaf* where ‘ea’ were pronounced differently. Irregular words are words that cannot be pronounced accurately when using only the most frequent GPC rules (Colenbrander et al. 2020). For example, the most frequent regular pronunciation of the grapheme *ch* is the [tʃ] phoneme in *hatch*, *much*, *cherry*, *beach*, and *church*. However, the words *stomach*, *chaos*, and *machine* are mispronounced if read using this same GPC rule. This orthographic complexity thus requires additional cognitive resources for children to successfully acquire and process such printed words. Furthermore, it has long been recognized that comprehensive causal explanations of cognitive and behavioral processes in reading must extend to their biological underpinnings (Morton & Frith, 1995; Snowling, Hulme, & Nation, 2022). The present study aims to investigate variations in the acquisition of reading skills among typically developing (TD) children and late elementary school children with reading disability (RD). We specifically aim to assess one promising candidate resource for managing GPC complexity, known as Set-for-Variability, and examine its potential neurophysiological underpinnings that contribute to the capacity to read all English words accurately.

Set-for-Variability and reading acquisition. Developmentally, children who have mastered phonetic decoding can potentially determine the pronunciation of unfamiliar regular words as a form of self-teaching, a process that also results in word-specific orthographic representations (Share, 1995,1999). The replicability of evidence for this self-teaching process was reported by Li and Wang (2023) in a systematic review of some 45 relevant published experimental self-teaching studies. Theoretically, it is possible that this self-teaching also includes words with irregular spelling patterns. Children might achieve this by strategically deploying an increasing understanding of grapheme-phoneme correspondence variations alongside relevant word-specific vocabulary knowledge in decoding. More specifically, irregular word reading accuracy can be achieved by applying variant GPCs to create and evaluate candidate phoneme strings until one is produced that both (i) aligns with a plausible candidate word stored in the child's lexical memory, and (ii) makes sense in the context in which it appears. This systematic problem-solving process is referred to as Set-for-Variability (SfV; Gibson, 1965; Levin & Watson 1963; Tunmer & Chapman, 2012; Venezky, 1999), defined as "the ability to determine the correct pronunciation of approximations to spoken English words" (Tunmer & Chapman, 2012, p 123).

SfV is often conceptualized as a cyclical lexical problem-solving process: if a child struggles to read the irregularly spelled word *pint* by using the regularized pronunciation, then *"*the child has to change one or more sound associations and try again*."* (Venezky, 1999, p 232). SfV is usually measured using a mispronunciation correction task from Tunmer and Chapman's (1998) list. In this task, participants listen to a recording of a word that is pronounced incorrectly and then verbally provide the correct pronunciation. For example, "her granny is very kind (kind rhyming with wind). The mispronunciations are regularized pronunciations of irregular words, created by sounding them out using common decoding rules and the most frequent grapheme-phoneme correspondences and syllable patterns.

Dyson et al. (2017) used the SfV mispronunciation correction task to assess the impact of behavioral SfV strategy instruction given to TD readers over four weeks. The group that received the intervention showed improvements in their ability to correct mispronunciations and to define the words they were taught. Additionally, these improvements extended to a similar set of words that were not explicitly taught, indicating generalization of the acquired skills. Edwards et al. (2022) provide evidence of a strong correlation between SfV and important reading components including word identification, word attack, and phonological awareness, and a moderate correlation with rapid letter naming in 489 typically developing readers in English in grades 2-5. Tunmer and Chapman (2012) conducted a 3-year longitudinal study that revealed that SfV in 152 TD readers in grade 1 had an indirect impact on reading comprehension through its influence on both decoding skills and word recognition; SfV also mediated the impact of vocabulary on reading.

Steacy et al. (2019) explored whether a student's performance in the SfV task predicted their ability to read target words with irregular spellings. In their study, which included 103 students in grades 2 to 5, with an oversampling of poor readers, participants were asked to complete reading tasks that included irregular word reading, SfV, vocabulary, and phonological awareness. Their results demonstrated that students who correctly completed a specific SfV item had a 79% probability of reading that same word correctly. Importantly, a subset of students at risk of reading disability had a reduced likelihood of reading the corresponding word correctly if they failed the set for SfV task on that word. This finding indicates that SfV serves as a predictor of irregular word reading, particularly for readers with challenges in reading proficiency.

Kearns et al. (2016) demonstrated a correlation between mispronunciation correction (SfV) and reading ability, suggesting that SfV may reflect the role of both phonological and semantic processes in word reading. Additionally, Steacy et al. (2022) aimed to explore how SfV relates to word reading, along with other factors, across different levels of reading skill and demonstrated a strong correlation between SfV and word reading. Specifically, they found that SfV predicted word reading even after considering other important factors like phonological awareness rapid automatic naming, and vocabulary with SfV being the most significant predictor. Importantly, by demonstrating that SfV has the potential to predict word reading difficulties (especially at lower reading skill levels), the Steacy et al. (2022) study supports the proposition that SfV can be used for identifying reading difficulties. Understanding the specific cognitive processes involved in SfV tasks could provide valuable insights into the underlying mechanisms of word reading and reading difficulties.

Theoretical cognitive frameworks of reading, such as the Lexical-Quality Hypothesis (Perfetti & Hart, 2002) assert that detailed word-specific representations are comprised of interconnected sources of orthographic, phonological, and semantic information. A computational framework of word reading acquisition, the ‘Triangle Model’ of word reading (Seidenberg & McClelland, 1989; Plaut,1994; Seidenberg, 2005) posits a distributed parallel processing network of somewhat neuron-like simple units and hidden units that come to represent semantic, phonological, and orthographic information. Formally implemented versions of the models can learn to ‘read’ regular and irregular words (Seidenberg, 2005) following system-wide feedback on errors, with semantic information resolving orthography-phonology inconsistencies (Plaut, 1994). To this end, SfV has sometimes been related to this semantic ‘clean up’ in triangle models (Edwards et al., 2022), with Steacy et al. (2019) arguing that SfV is a “process that cleans up the mismatch between orthography to phonology conversion and word pronunciation.” (p. 523). Consequently, SfV is a potentially very important cognitive ability implicated in the reading acquisition process.

In the present study, we aim to further understand the specific role of SfV among readers with and without RD. To date, it remains unknown if SfV is delayed between well-matched proficient and poor readers. This study assesses this potential difference for the first time. Furthermore, it is widely accepted in developmental research, both generally (e.g., Lovio et al., 2012) and specifically within the context of reading (e.g., Snowling, Hulme, & Nation, 2022), that the most comprehensive reading models must account for both cognitive and associated biological processes in development. Consequently, to gain a full understanding of SfV, we aim to understand the neural processing implicated in its utilization during reading. The following section expands on the potentially relevant neural and brain-behaviour associations for SfV.

Temporal-parietal effects in semantic comprehension and meaning acquisition. Neurophysiological evidence highlighting the relationship between semantic processes and reading is evidenced by the N400, an event-related potential (ERP) waveform that exhibits a negative voltage deflection and peaks at around 400 ms after stimulus onset (Kutas & Hillyard, 1980). In semantically congruent conditions, where the final word of a sentence is semantically expected ('dad is eating the … pizza'), the N400 amplitude is reduced compared to incongruent conditions involving semantically unexpected words ('dad is eating the … sock') which elicit a larger negative brain response in the N400 amplitude (Kutas & Hillyard, 1980). Widely recognized as an index of semantic incongruency in word processing (Kutas & Federmeier, 2011), the N400 effect (difference in waves between congruent and incongruent conditions) has also been observed in infants who had typical expressive language skills, but not in those who had a higher risk of Specific Language Impairment (Friedrich & Friederici, 2006). The N400 effect is also sensitive to the lexical aspect of language (Sun et al., 2023). Hasko et al. (2013) conducted a phonological-lexical decision task (PLD) in eight-year-old children with and without RD. Children read orthographically familiar real words, orthographically unfamiliar pseudo homophones, pseudowords, and false fonts and decide which sounds like real German words. The results demonstrated that TD readers exhibited an N400 effect, while responses were aberrant in children with RD. Responses were considered "aberrant" or atypical in children with RD because they showed diminished or attenuated (smaller) N400 amplitudes when processing pseudo homophones and pseudo-words. The specific pattern of the reduced N400 amplitudes in children with RD, particularly in response to orthographically familiar and unfamiliar word forms, suggests a difficulty in lexical access rather than just phonological processing. This reduction suggests difficulties in accessing the orthographic lexicon and grapheme-phoneme conversion, resulting in an increased difficulty in processing and understanding words (lexical difficulty). Sun et al. (2023) extended these findings by demonstrating that TD readers exhibited an increased N400 effect during non-word reading, while children with RD did not. The task in this study was also a PLD task, where participants were asked to decide if a visually presented stimulus sounded like a real word or not. The stimuli included real words, pseudo homophones, pseudo words, and false fonts. The task aimed to explore orthographic and phonological processing in children with and without RD. Control children showed more pronounced N400 amplitudes for orthographic unfamiliar (PH, PW) words compared to familiar (W), indicating better lexical processing. This suggests that the N400 indexes difficulty in GPC processing in RD, although Sun et al. (2023) note that weak orthography-semantic connections suggested in the literature may be implicated in N400 deficits in RD. Coch and Benoit (2015) suggested that the N400 indexes interactive word processing at multiple, cascaded levels of word representation. Meng et al. (2007) conducted a study on sentence incongruity in 10-year-old Chinese children with and without dyslexia. Their results showed that the N400 amplitude increased in the incongruent over the congruent condition for the control group, whereas there was no difference in the N400 amplitude for the dyslexia group, suggesting that children with dyslexia may struggle with processing semantic mismatches in sentences. Additionally, Schulz (2008) found that 11-year-old children with dyslexia showed a decreased N400 effect during sentence reading when compared to controls, further supporting the idea that dyslexia is associated with impaired semantic processing, reflected by the aberrant N400 response. These findings highlight the importance of semantic processing deficits in dyslexia, as reflected by the N400 component responses, which appear to impact the ability to integrate context and detect incongruent information during reading.

In addition to the N400 effect, N400 latency is another important aspect of neurophysiological linguistic information processing (Lazzaro et al., 2000). Latency, measured in milliseconds (*ms*), is the time from stimulus onset to the point of maximum amplitude within the given latency time window (Kropotov, 2016; Liesefeld, 2018). Helenius et al. (1999) report a delayed N400 peak latency in adults with RD reading incongruent conditions, suggesting a delay in semantic activation word processing of the final word in a sentence. This finding agreed with Brandeis et al. (1994) who found peak delays post 100 ms of onset words of sentence reading in children with dyslexia. Jednoróg et al. (2010) found a delayed N400 peak latency in children with RD compared to age-matched TD readers. Their task involved participants listening to a set of primed words in congruent and incongruent conditions, revealing a group difference only in the incongruent task. This suggests that individuals with RD may have difficulty detecting semantically incongruent information (semantic anomalies) but not integrating semantic information in congruent sentences.

A recent meta-analysis of all relevant well-executed studies reported an N400 difference in amplitude and latency between individuals with and without RD, in the incongruent condition, (Author et al., 2023). However, the overall reported effect size for N400 between TD and RD varied: ERP tasks involving reading printed tasks silently (visual stimuli) had a larger effect size ($g = 0.74$) compared to tasks involving listening to the spoken equivalents. Semantically incongruent sentence tasks demonstrated the most robust difference, with an effect size of $g = 0.94$, over non-sentence tasks (real words, pseudowords). These patterns suggest a consistent N400 difference between TD and RD individuals. It is important to note that the meta-analysis (Author et al., 2023) included studies that explored the N400 in participants aged 10 years and older. Of these studies, only four studies used a printed congruent versus incongruent sentence task in alphabetic systems (Aiello et al., 2018, Brandeis et al., 1993; Robichon et al. 2002; Schulze; 2008); the average participant age across these studies was fourteen years. Importantly, research has yet to explore the differences in the N400 effect and peak latency in sentence incongruency tasks in populations younger than 10 years old, where reading is still being acquired. This is particularly relevant for readers with RD, where associations between the N400 and reading acquisition might be more evident.

Furthermore, the four selected studies were conducted in transparent spelling systems like German and Italian, resulting in a gap in understanding effects in opaque orthographies such as English. This study aims to address these gaps by exploring the differences in the N400 effect and latency during sentence incongruency tasks between younger children, particularly those under 10 years old, and younger readers. By examining these differences, especially in children with reading difficulties (RD), the research seeks to understand how the N400 effect is linked to the process of reading acquisition and comprehension.

Individual differences in Set-for-Variability, reading proficiency, and the N400 profile. Beyond the reader group contrasts considered thus far, further evidence of brain-behaviour links for N400 and early reading comes from correlational studies. Henderson et al. (2011) examined the relationship between the N400 and behavioural measures of listening comprehension, word recognition, non-word decoding, and receptive vocabulary knowledge in eighteen typically developing readers aged between 8 and 10 years. A significant correlation was found between the N400 amplitude and listening comprehension ($r = .57$) and with non-word decoding (incongruent condition), $r = .63$. Khalifian et al. (2016) used linear mixed-effect regression to explore the correlation between ERP components, including the N400, and phonological awareness, print exposure, vocabulary, and school report cards. The study involved 65 TD children ranging in age from 4-12 years who were asked to press a button once their name appeared among three types of items: regular words, pseudowords, and illegal letter strings. Results demonstrated a significant relationship between N400 amplitude and phoneme blending, as well as with a non-standardized measure of oral vocabulary. Coch and Benoit (2015) examined the correlation with measures of spelling, phonological processing, vocabulary, comprehension, naming speed and memory among 72 typically developing children aged eight to ten years. Results demonstrated a negative correlation between the standard scores on the PPVT and the average N400 amplitudes for both words ($r = -.272$) and pseudowords ($r = -.235$). In addition, Hasko et al. (2013) reported on N400 amplitude and literacy among 25 TD children and 28 (8-10yrs) children with RD, finding a negative correlation between the N400 amplitude and spelling ability when combining the two groups for analyses irrespective of diagnosis. Finally, Sun et al. (2023) explored the correlation of the N400 amplitude with reading comprehension, sight word, phonemic decoding, and receptive vocabulary in 32 TD children and

20 children with RD (aged 7 to 11.5 years). They reported no correlation between N400 amplitude and any of the reading measures. However, when combining the two samples, sight word efficiency and phonemic decoding efficiency jointly accounted for significant variance in the N400 effect in stepwise regressions.

With regards to N400 and word reading, well-established and influential theories (Bishop & Adams, 1990; La Berge & Samuels, 1974; Stanovich, 1980; Perfetti & Hart, 2001) strongly suggest that TD readers very likely differ from those with RD in the degree to which they have automated reading low-level word reading processes. Furthermore, Yang et al. (2005) argue that RD readers are slower in word integration processes, which is then manifested in a smaller, less negative N400 amplitude. This pattern was observed in Landi and Perfetti (2007) when comparing N400 amplitude between TD and RD readers. Thus, the delayed profile of the word integration process and smaller N400 amplitudes and later latencies expected for RD readers suggest it would be more insightful to explore individual differences in N400 reading measures separately for age-matched RD and TD groups. In the case of SfV, a cyclical problem-solving strategy for regular and exception word reading *acquisition,* different brain behavior pattern relationship is predicted between the RD groups who are still working to acquire and automate accessing the meaning of most words, and the TD readers, who have much more substantially mastered and automated word reading. Finally, in terms of measurements, point estimates for congruent versus incongruent sentences explored here will predictably be based on different performances for sentence reading accuracy and speed in TD versus RD. This necessitates the analysis of correlations in separate groups to avoid introducing systematic measurement errors. Conducting separate analyses for TD and RD participants is crucial for several reasons. First, it allows for a more accurate understanding of the distinct challenges faced by RD individuals, as their phonological processing and decoding skills often differ significantly from their TD peers. Second, separate analyses can help identify specific patterns and strategies that may be effective for one group but not the other, thereby informing more tailored and effective intervention approaches. Finally, understanding the differences between TD and RD groups can contribute to more precise diagnostic criteria and better-targeted educational practices, ultimately supporting improved reading outcomes for RD individuals.

In summary, the existing literature on individual differences in N400 and reading reveals associations reported at the level of word decoding, linguistic comprehension, or both. Interpretation of such patterns is further complicated by combining RD and TD samples in some studies (Hasko et al., 2013; Sun et al., 2023). Consequently, we aim to revisit patterns among TD children to establish more replicable patterns in TD readers to explore their generality of effects in children with RD. This study extends the literature to investigate the relationship between N400 and SfV, as both are plausibly implicated in lexico-semantic reading tasks. Notably, no correlation studies to date have included semantically congruent and incongruent sentences as the N400 task. Jednoróg et al. (2010) discussed that while sentence incongruency involves monitoring and assessing unexpected violations of meaning, sentence congruent tasks require monitoring of syntactic legality. It is yet to be determined whether there is a comparable correlation between the N400 in incongruent sentence tasks compared to congruent sentence tasks.

Aims of the Current Study

The current study has two primary research goals. The first aim is to evaluate group differences in the Set for Variability (SfV) and N400 components between elementary school readers with and without RD. This first aim is defined by the following research questions (RQs):

RQ 1A: Is there a difference in SfV measures between elementary school readers with and without RD? Readers with RD have difficulty reading irregular words (Castles & Coltheart, 1993; Colenbrander et al., 2020) and have poor decoding skills (Grizzle, 2007; Snowling et al., 2020). Decoding skills are necessary in SfV to correct mispronounced words (Tunmer & Chapman, 2012). Steacy et al. (2019) showed that individual differences in reading predict SfV. We, therefore, hypothesize a difference in SfV between the two groups since poor readers who struggle with decoding and irregular word reading also encounter difficulty with SfV tasks.

RQ 1B: Is there a difference in the N400 effect between elementary school readers with and without RD in processing the visual stimuli of a sentence incongruent task? We hypothesize that the N400 effect will differ between children with RD in a sentence reading task due to their difficulties in lexical-semantic processing (Schulz et al., 2008).

RQ1C: Is there a difference in the N400 peak latency and fractional peak latency between elementary school readers with and without RD? We predict a later peak latency in readers with RD since their predicted early GPC-related difficulties are indexed by N400 latency (Helenius et al. 1999). Consequently, in the N400 behavioral task (accuracy and reaction time), we expect to see lower accuracy and slower reaction time (RT) based on the hypothesis that children with RD might have difficulty in lexical processing, which leads to difficulty in semantic processing (Von Koss Torkildsen et al., 2007).

The second aim of this study is to assess correlations between the N400 effect and SfV in children with and without RD.

RQ 2A: Is there an association between the N400 effect and SfV in the TD and RD groups? RQ2B.Is there an association between SfV and the N400 peak latency in the TD and RD groups? We hypothesize a positive correlation between the SfV and the N400 effect.

Neurophysiological evidence suggests that the N400 is sensitive to the semantic context of language comprehension and word meaning (Henderson et al., 2011). Therefore, if the N400 effect reflects semantic processing, it should correlate more strongly with SfV, a meta-cognitive process that involves not only decoding but also accessing word meaning.

Additionally, we predict negative correlations between SfV and N400 latency and, with a particularly robust negative correlation expected with SfV. This prediction is based on the theorized role of SfV in early word reading acquisition, where individual differences in SfV have been shown to predict better word reading (Steacy et al., 2019). We hypothesize that the speed at which words are recognized and understood via SfV is correlated with the latency of the semantic activation measured by the N400 during final word reading.

Method

Participants. After obtaining ethics approval from the university research ethics board, participants were recruited through two local school boards and speech and language pathology clinics as per their respective protocols. Recruitment was facilitated by teachers and clinicians, who distributed information flyers and consent letters to possible participants. The children of caregivers who signed a consent form were then contacted and asked to participate. Fifty-four (N=54) children aged between 6.5 and 11.5 years (M_{age} = 9.2years) participated in this study (29 males and 25 females) recruited from Montreal, Canada. The TD group included 15 males and 7 females. The RD group included 14 males and 18 females. English was the dominant language at home for all participants. Inclusion criteria for participants with RD included either a confirmed diagnosis of RD, reading two grades below grade level on the WIAT-III (Wechsler Individual Achievement Test-III; REF) Word Reading and Reading Comprehension subtest standard scores. Given that phonological dyslexia represent a predominant feature of individuals with RD (Adel & Saleh, 2021), the Comprehensive Test of Phonological Processing (CTOPP) was used to evaluate the phonological processing skills of participants, thereby assessing the presence of RD (Park & Lombardino, 2013). Therefore, participants scoring Below Average on the Phonological Awareness composite score of the CTOPP, in addition to WIAT Word Reading and Reading Comprehension subtests were included in the study.

Sixty-three percent of the sample had a confirmed diagnosis of specific learning disability in reading according to the DSM-5 (APA, 2013) criteria, 17% had a report from parents and teachers of reading one or two grades below age-expected grade level, and 10% were transferred from speech and language clinics as reading one or two grades below age-expected grade level. None of the caregivers of participants in the TD group reported a reading, learning, attentional or any other neurological disorders condition of any kind during a pre-recruitment phone interview. The data from two participants from the RD group and one participant from the TD group were excluded from the electroencephalography (EEG) data due to the presence of artifacts while recording. The final sample used for data analysis included a total of fifty-one (N=51) participants: 21 in the TD group and 30 participants in the RD group.

General procedure.

A trained research assistant and the primary author provided participants and their caregivers with a brief laboratory tour and an overview of the day's procedures. Subsequently, children were asked to provide assent by signing a form, while caregivers provided consent by signing a separate form. Participants were then asked to complete a series of tests assessing various aspects of reading skills (see the Reading Measures section below for a complete description). After completing the reading tests, participants were seated in a separate, sound-attenuated room, positioned 70 cm away from a visual display (ViewPixx©), and fitted with an electrode cap. To limit boredom, participants were given the choice to watch a video of their choice while the electrodes were fitted. Once the cap was fitted, participants were asked to blink their eyes, move their eyes to the right and then to the left, and clench their jaws while the research assistant took a picture of the brain wave patterns generated by these movements and showed it to the participant to illustrate how physical movement is reflected in brain waves. Participants were then instructed to sit, relax, and avoid eye blinks as much as possible unless they saw a "+" on the display. The experimental task was explained to the participants in detail, followed by a brief practice session consisting of four trials. The task would only proceed once the participant indicated readiness by pressing on the spacebar. To limit fatigue, a "Big Pause" appeared on the screen every 25 trials, providing the participant with a break. The task resumed only when the participant was ready, as indicated by pressing the spacebar. All reading and EEG measures, including planned and unplanned breaks, were completed within 2.5 hours.

Readings Measures.

We used WIAT reading comprehension, word reading, and expressive vocabulary assessments, along with the CTOPP phonological test composite scores, as screening tools to ensure that the RD group performed below average based on standardized scores, to be included in the study. The Castle & Coltheart and SfV measures do not include standardized scores; therefore, we used the raw scores for these assessments. The SfV measure was included because it is the primary focus of this study. We also incorporated the Castle & Coltheart measure to examine its correlation with SfV, as demonstrated in previous studies (Castles et al., 2009; Dyson et al., 2017).

<u>Wechsler Individual Achievement Test-III</u> (WIAT-III; Wechsler, 2009). Three subtests of the WIAT-III, a standardized measure of academic achievement in the areas of reading, writing, math, and oral language, were used in the study. (1) *Reading Comprehension*. The purpose of this subtest is to measure the reader's inferential and literal reading comprehension skills by

reading passages and answering questions. The reader starts at their grade level based on their chronological age. The participant has the choice to read aloud or silently. After each passage, the examiner poses comprehension questions. If the reader scores two points or lower on the passage comprehension questions, the examiner returns to one grade level lower than the current level. The published internal reliability coefficient is $r = 0.90$. (2) *Word Reading*. The purpose of this test is to measure the reader's ability to identify isolated words with accuracy. The reader is required to read aloud from a list of words. The words increase in their level of difficulty as the participant continues to read. For each correct response, the reader gets 1 point. The test stops if the reader makes four consecutive errors. The published internal reliability coefficient is $r = 0.98$. (3) *Expressive Vocabulary*. This is a subtest of WIAT oral expression that measures oral vocabulary and word retrieval. The examiner shows the reader a picture with a definition and asks the reader to say the word that corresponds to the given picture and definition. For example, the examiner shows the participant a picture of a toothbrush and asks "Tell me the word that means a brush for cleaning teeth. The subtest is discontinued after 4 consecutive mistakes. The published split-half reliability is $r = .71$. The standard scores were included for analysis.

Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 2013). This is a standardized test that measures phonological processability. The CTOPP phonological awareness composite score is based on three subtests included in this study: Elision, Blending words, and Phoneme Isolation. *Elision*. The test measures the ability to remove sounds and syllables in spoken words. There are 34 items in this subtest. The examiner asks the participants to say compound words and then asks the participants to say the word after dropping the compound word. For example, the examiner asks the participant to say "cowboy." The examiner then asks to say cowboy without the "boy". The task then increases in difficulty when the participant is asked to delete phonemes (e.g., say “slight” without the “l”). The test is discontinued after three consecutive mistakes. The published reliability is *r* = .82. *Blending* words. To measure the participant’s ability to combine sounds to form correct words. There are 33 items in this subtest. The participant listens to recorded audio that for example, asks to combine “t”, and “oi” to have the word “toy”. The subtest is discontinued after three consecutive measures. The published reliability is *r* = .75. *Phoneme Isolation*. The purpose of this subtest is to measure the participant’s ability to identify individual sounds in given words. There are 32 items in this subtest. The examiner asks the participant to identify the first and last sounds in vowel consonant vowel (CVC) words. Subsequently, the participant is asked to identify middle sounds in words with four and five sounds. The published reliability is *r* = .83. The standard scores of phonological awareness were included in the analysis.

<u>Set-for-Variability</u> (SfV; Tunmer & Chapman, 2012). This test assesses children’s ability to determine the correct pronunciation of regularized exception words. Twenty items that were used in Tunmer and Chapman’s (1998) mispronunciation task were selected and recorded for consistent delivery. The examiner read the following script: “I have a friend who will read out some sentences, but my friend will read out the wrong words at the end of each sentence. Would you please assist my friend in using the right word?” Once the examiner read the script, the children listened to a recorded voice of sentences of mispronounced words at the end of each sentence. The following are some examples of the task.

1. He got mud on his **shoe** (pronounce: show)
2. The dog had to have a **wash** (pronounce rhyme with ash).
3. The cake was shaped like a **heart** (pronounce rhyme with hear-t).

The responses were scored as 1 (correct) or 0 (incorrect). The published inter-rater reliability is *r* =.86 (Tunmer & Chapman, 2012). The raw scores were included in the analysis since this test has no standard scores.

Castles and Coltheart reading test (CC2; Castles et al., 2009). This consists of three sets of printed items: 40 regular words, 40 irregular words, and 40 nonwords presented one word at a time and increasing in difficulty. A subscale test was discontinued when five consecutive errors were made. Reliability for this test is $r = 0.92$ for regular words, $r = 0.94$ for irregular words, and $r = 0.85$ for nonwords (McArthur et al., 2015). The raw scores were included in the analysis as no standard scores are provided for this test.

Electrophysiological measures. Electrophysiological data was collected while participants completed a sentence-judgment task adapted from Brandeis et al. (1995) and Sabisch et al. (2006), programmed using PsychoPy© software. The task included visually presented semantically congruent or incongruent sentences, made up of words from common storybooks widely accepted as known to children (Robert McCloskey book collection, Robert Munsch story books). The syntactic structure of the sentence was either subject-verb-object or subject-verb-prepositional object; the final word of the sentence was the critical word, as it determined whether the sentence was semantically congruent or incongruent. To be acquainted with the task, a trained research assistant explained the task and example trials were shown to the participants.

One hundred sentences, 50 with congruent and 50 with incongruent semantic endings were presented to each participant. The task was presented in four blocks, with each block containing 25 sentences. Congruous and incongruous sentences were presented randomly to each participant in each block. The length of each sentence varied between five and seven words; the final word of each sentence varied between four and eight letters. The following are examples of sentences in the congruent condition; "My father is eating an apple", and "The fish is swimming in the water". Incongruent sentences were based on congruent sentences with the critical, final word of the sentence replaced with another word, exemplified by: "My father is eating a *blanket*." and "The fish is swimming in the *lamp*.".

Each trial started with a yellow fixation cross presented on the screen for 600 *ms*. A sentence was then presented as a temporal sequence of individual words in lowercase letters, each word appearing for 700 *ms*., at the fixation point. The final (critical) word of each sentence string, which was either congruent or incongruent, was followed by a period. Participants were asked to read each sentence silently and were presented with a blue fixation + for 2500 *ms*, indicating to them that a judgement was required via a button press; green if the sentence was meaningful to them and red if it was not. Participants were given a break between each sentence. stimulus (approximately 1500 *ms*) for eye blinking. The participants were given a longer pause between blocks. The pause was self-paced, and they only resumed to the next block when they were ready. They were encouraged to breathe in and out and stretch their hands and legs before proceeding to the next block.

***Event Related Potential (ERP) data acquisition and processing*.** ERPs were recorded during task completion using Brain Vision Quickamp- 64 channels. The ERP components were measured using electroencephalography (EEG) recordings from 32 active Ag/AgCl electrodes placed in an elastic cap according to the 10/20 system (Morely, 2016) with a left and right mastoid as a reference, and AFz as the ground. To avoid large artifact potentials, the impedance was kept at 5 kΩ. Filtering was at a low cut-off: of 0.1 Hz and a high cut-off: of 30 Hz. Artifacts such as eyeblinks were removed with Independent Component Analysis (ICA) before average and other artifacts such as a gradient of more than 15µ V in a 200 ms window, low activity of less than 0.5

µ V in a 200 ms window, and absolute amplitude of 200 µ V in a 50 ms window. The vertical VEOG was recorded from electrodes placed above and below the right eye, and the horizontal EOG (HEOG) was recorded from electrodes placed 1.5 cm lateral to the left and right external canthi. The continuous EEG recordings were epoched offline (-200 to 1200 *ms)* with the onset of the final word in each sentence as 0 *ms*). Trials containing electrooculogram (EOG) artifacts, characterized by high-amplitude patterns in the brain signal due to blinking or low-frequency patterns caused by eye movements (such as rolling) exceeding ±75 µV, were excluded from subsequent analysis. For the initial inspection, we viewed a 64-channel waveform plot for all the participants. Time points (300-500) and electrodes were selected for further inspection based on previous research of the N400 (Duncan et al., 2009). Nineteen channels were selected for the final analysis: F3, FZ, F4, C3, CZ, C4, CP5, CP1, CP2, PZ, P3, P4, FC1, FC2, OZ, O1, O2, T7, T8. These channels were selected based on the guidelines reported by Duncan et al. (2009) and the commonly reported N400 electrodes in the review by Šoškic et al. (2022). Thirty-five clean trials were necessary to be included in the analysis, for each participant. Any trials that included less than 35 were excluded from further analysis. Three participants were excluded from the analysis for having less than 35 clean trials.

Results

Data analyses**.** The participants were matched on age and gender, and *t*-tests conducted by group (TD versus RD) for these variables showed no statistically significant differences in age ($t = -0.91$, $p = 0.364$) or gender ($t = -1.52$, $p = 0.132$). Demographic information is presented in Table 1. The descriptive statistics for the reading measures are presented in Table 2, with N400 behavioural measures (accuracy and reaction time) given in Table 3, and N400 amplitude and latency in Table 4. The brain-behaviour correlations of interest are presented in Table 5. The correlation between the N400 effect and SfV is presented in Tables 6 and 7. An overview of all reading measures is summarized in Table 2, with WIAT and CTOPP standardized scores allowing for between-group differences and comparison to the norm. All TD children scored in the average range for WIAT Word reading (standard score), WIAT reading comprehension (standard score), WIAT expressive vocabulary (standard score), and the Phonological awareness composite score of the CTOPP; all children with RD scored in the Below Average Range on these same measures. Overall, a MANOVA analysis showed the expected main effect of the group (TD versus RD) on all reading and expressive vocabulary measures (see Table 2).

Before conducting inferential analyses, the Shapiro-Wilk test of normality was used to ensure that all behavioral variables were normally distributed ($p > .05$). Skewness scores ranged from -3 to +3, and kurtosis scores ranged from -3 to +5, which are both acceptable values according to Kallner (2017). Levene's test for equality of variances resulted in a *p*-value greater than 0.05, indicating that assumptions of homogeneity of variance were met.

Research Question 1A. The first part of Research Question One (RQ1A) aimed to determine whether there is a Set for Variability (SfV) difference between elementary school readers with and without RD. An ANOVA revealed a significant main effect of Group F (1, 52) = 9.401, p = .003, with the TD group (*Mean* = 13.76, *SD* = 3.87) outperforming the RD group (*Mean* = 10.64, *SD* = 3.50). To control for WIAT expressive vocabulary differences across groups, MANCOVA was conducted since vocabulary explains a unique variance to SfV (Tunmer & Chapman, 2012). This analysis revealed that the main effect of the Group remained significant F (1, 54) = 4.026, p = .05, with the TD group (*Mean* = 13.14, SD = 5.85) outperforming the RD group (*Mean* = 11.02, SD = 4.58). These analyses confirm the hypothesized distinct reader performance difference on SfV between the TD and RD groups assessed.

Research Question 1B. *N400 effect.* To examine the N400 effect, which is the difference in wave between congruent and incongruent conditions, ANOVA was conducted with the N400 effect as the dependent variable and group (TD, RD) as between subject effect. The results showed a main effect of group on the N400 effect, F (1, 49) = 11.53, p = .001, η_p^2 = .191 using the Bonferroni correction for multiple comparisons, with the TD group demonstrating an increased N400 amplitude for the TD group (Table 4). Additionally, the plotted waves (Cp1, Cz, CP2, CP5) show the increased negative amplitude of the N400 effect for the TD group as shown in Figure 1.

N400 mean amplitude. Additional analysis was conducted to examine the N400 average mean amplitude. Analysis of the mean average of the N400 amplitude using Repeated measures ANOVA (RM-ANOVA) with group (TD; RD) as the between-subject facture and Condition (congruent; incongruent) as within-subject factor showed that there is a main effect of group F (1, 36) = 57.72, p < .001, η_p^2 = .616. There was an interaction effect of group x condition F (1, 36) = 61.09, p <.001, η_p^2 = .629, which shows that the TD group had a more negative peak in the incongruent condition than the RD group (See figure 2).

Research Question 1C. *N400 Peak latency.* There was no main effect of Group, F (1, 49) = 1.013, p =.319 found. There was no significant interaction between Group and Condition (F (1,49) = .698, p = .407, indicating that there were no latency differences between the two groups in either condition (congruent; incongruent).

N400 Fractional peak latency. To measure onset latencies, we recorded the fractional peak latency, which is the time when the ERP waveform reaches 50% of its maximum amplitude (Luck, 2014). Analysis showed the main effect of group F (1,72) = 46.89, p <.001, η_p^2 = .396, where the TD group had an earlier onset latency (*Mean* = 312, *SD* = 75.45) than the RD (*Mean* = 442, *SD* = 212.97).

N400 behavioral task. ANOVA was conducted to analyze the accuracy and reaction time during task completion (see Table 3), with factors being Condition (congruent, incongruent) and Group (TD, RD). For RTs, the main effect of the Group was significant, F (1,50) = 7.931, $p <$ .001, with the RD group exhibiting significantly slower RTs (*Mean* = 1173 ms) compared to the TD group (*Mean* = 920 ms). There was no main effect of Condition, F (1, 50) = 0.312, p = .578, nor a significant Group X Condition interaction effect, F (1, 50) =.10, p =.921. The RD group was consistently slower than the TD group across both congruent and incongruent conditions. For accuracy, there was a main effect of Group, F (1, 50) = 89.88, p <.001, with the TD demonstrating a higher accuracy rate in the congruent condition (*Mean* = 88.35%, *SD* = 17.09) than the RD group (*Mean* = 52.56%, SD = 21.56) and a higher accuracy rate in the incongruent condition (*Mean* = 87.84, *SD* = 18.74) than the RD group (*Mean* = 54.18*,* *SD* = 25.94). The interaction between Condition and Group was not significant, F (1, 50) = .030, p = .864, with a higher accuracy rate in the TD group for both congruent and incongruent conditions compared to the RD group (See Table 3).

Research Question 2

The second main research question aimed to assess the relationship between reading measures and the N400 effect and peak latency average across the two groups. For the TD group, a Pearson correlation for the reading measures revealed that reading comprehension was correlated with expressive vocabulary (r = .60, p = .004), and phonological awareness (r = .497, p <.05). SfV was positively correlated with all reading measures (values between r = .50 and .66, all p < .001, see Table 5). Pearson correlations were run to address the correlation between SfV and the N400 effect for the TD group (RQ2) and revealed a significant correlation, $r = .46$, $p = 0.03$ (See Table 6). A similar analysis was used to explore the existence of a correlation between N400 latency and SfV for the TD Group (RQ2B). There was no correlation between the N400 latency and SfV (See Table 3).

As depicted in Table 7, Pearson correlations between the N400 effect and SfV for the RD group revealed no correlation. Finally, we explored the correlation between the N400 latency and SfV for the RD group, analysis demonstrated a significant negative correlation between N400 latency and SfV ($r = -.43$, $p = 0.02$).

Discussion

This study aimed to evaluate if there is a difference in SfV and N400 effect and latency between elementary school readers with and without RD. In addition, we aimed to assess correlations between SfV and N400 effect and peak latency in children with and without RD. For our first research question, we found a significant difference in SfV between TD and RD groups. Significant between-group differences in SfV remained after additional statistical control for expressive language differences between the groups. We did not use listening comprehension measures; hence it was not controlled. This expressive language control is a conservative one given the close theorized association between SfV and vocabulary knowledge (Tunmer & Chapman, 2012). To our knowledge, this is the first time an SfV difference between TD and RD has been investigated and reported in a formal chronological age- and extraneous variable-matched control study. The results indicate a difference in SfV processing between the RD and TD groups, consistent with its theorized important role in reading development (Venezky, 1999; Tunmer & Chapman, 2012).

Also as predicted, TD versus RD group differences in the N400 effect were identified. Specifically, the N400 amplitude was found to be more negative to an incongruent condition in TD children, replicating other studies in older readers (Coch & Holcomb, 2003; Friedrich & Friederici, 2006; Henderson et al., 2011; Kutas & Federmeier, 2011). The N400 effect of the event-related potential (ERP) reflects semantic processing during language comprehension (Kutas & Hillyard, 1980). The differences we found in how the N400 effect differs between TD and RD groups imply that persons with reading difficulties might process words and sentences differently. Specifically, they may show unique patterns in how they respond to the meaning of words during sentence reading tasks (Tzeng et al., 2017). This suggests that the brain mechanisms underlying language comprehension could vary between individuals with and without reading difficulties. The absence of the N400 effect in the RD group could suggest that RD children have difficulty in the semantic integration process, specifically in the incongruent condition (Russler et al., 2007).

In the second research question, we hypothesized a correlation between the SfV and the N400 effect. We found a correlation between the N400 effect and SfV in the TD group. Correlation studies to date involving TD samples show that the N400 is often considered an index of word reading (Sun et al., 2015). This pattern suggests better word readers (TD group) showed larger N400 amplitudes while reading than struggling readers (RD group). Results here are also consistent with the anticipated disruption of N400-reading patterns in the context of additional task demands for incongruent sentences among RD samples. The difference observed in the SfV measure among the reading difficulties (RD) group compared to the TD group suggests underlying differences in lexical-semantic processing abilities. SfV requires individuals to rapidly associate sounds with corresponding visual representations, reflecting the integration of phonological and semantic information during language processing (Barnes et al., 2022). Similarly, the N400 component of the event-related potential (ERP) serves as an index of lexical-semantic processing, reflecting the brain's response to semantic incongruities (Kutas & Hillyard, 1980). The significant N400 effect observed in the TD group indicates efficient processing of semantic information and sensitivity to semantic incongruities among typically developing children. This aligns with previous research demonstrating robust N400 effects in response to semantic violations in proficient readers. In contrast, the absence of a significant N400 effect in the RD group suggests altered or impaired lexical-semantic processing mechanisms (Wang et al., 2017). This finding implies that individuals with RD may exhibit deficits in semantic integration or access, resulting in diminished sensitivity to semantic incongruities. The lack of correlation between the N400 effect and SfV performance in the RD group further underscores the dissociation between lexical-semantic processing and reading fluency in this population.

Overall, these results highlight the complex interplay between lexical-semantic processing, reading, and reading difficulties. The observed delay in SfV and absence of N400 effects in the RD group underscore the multifaceted nature of reading difficulties, implicating deficits in both phonological and semantic processing.

Finally, here, SfV was significantly negatively correlated with N400 latency in RD but not TD readers, as predicted based on its theorized role in word reading acquisition. This finding suggests that the N400 latency might also be a potential estimate of individual differences in SfV ability at the individual student level in the RD group. SfV involves both GPCs and meaning processing (Elbro et al., 2012), thus quality access to semantics is needed to decode mispronounced words (Tunmer & Chapman, 2012). Therefore, the observed correlation between SfV and N400 latency suggests that individual differences in N400 latency may reflect variations in semantic processing efficiency, which, in turn, influence SfV performance. Specifically, individuals with better SfV scores in the RD group may exhibit earlier N400 latencies, indicating a more efficient semantic activation processing. Conversely, individuals with poorer SfV scores may demonstrate delayed N400 latencies, indicative of slower or less effective semantic processing. This highlights the potential utility of N400 latency as a marker of semantic processing efficiency and its relevance for understanding individual differences in reading fluency, particularly in populations with reading difficulties. Previous research by Shaywtiz and Shaywitz (2008) suggests that individuals with reading difficulties often engage in more effortful and less automatic reading strategies. In contrast, typically developing readers are more likely to process words automatically. In this study, participants in the RD group exhibited longer latencies, which may reflect a greater reliance on phonological decoding strategies. Conversely, participants in the TD group demonstrated a strong N400 effect, indicative of more efficient semantic processing and automatic word recognition. Overall, the somewhat different pattern of associations for RD and TD groups for SfV and N400 here suggest that both groups exhibit qualitatively different processes when reading words and reflecting their relative reading expertise and word reading automaticity.

Limitations and Future Studies

Several limitations need to be considered. First, the present study had a small sample size in our TD group. Sample sizes here are however very similar to other published studies (e.g., Sun et al., 2023). The above-average word reading performance of the TD group, combined with the sample size, might have contributed to the reported pattern of associations. The study recruitment procedures were affected by COVID-19 as they involved direct contact with participants, and this was not possible because of COVID-19 restrictions. The COVID-19 restriction precluded us from performing a power-calculated study. Future studies are recommended to include power calculation estimates for better results. Nevertheless, the small sample size was sufficiently powered to detect significant differences between the groups and many predicted correlation patterns.

Although our groups were matched for chronological age, a limitation is the relatively broad age range of participants within each group, during which different levels of reading acquisition may be present, with such variability potentially introducing confounding factors that may affect the interpretation of our results s of the study. Therefore, future research should consider narrower age ranges or include age as a covariate in the analysis to account for these developmental differences. Additionally, while our data confirmed that TD children showed improved SfV scores with age, the scores for children with reading difficulties RD did not exhibit significant changes. In fact, an exploratory inspection of the data showed that the mean scores of the SfV increased with age for the TD group (7 yrs. old *Mean* = 11.5, 8 yrs. old *Mean* = 14, and 9 yrs. old *Mean* = 15). The mean scores for the RD group, however, were (7 yrs. old Mean = 5.5, 8 years old Mean = 11.4, 9 yrs. old = 11.2). This might suggest that TD children show a clear improvement in SfV performance with age, which aligns with existing research on the development of decoding skills. In contrast, the relatively stable scores in the RD group suggest that these children do not experience the same developmental trajectory. However, the results are interpreted with caution because of the small sample size in each age group. Future longitudinal studies with the RD population are important to account for developmental changes in SfV.

In addition, another limitation we had in this study is the accuracy rate of the RD group, which raises questions about the difficulty and validity of a sentence processing task to elicit

N400. We emphasize that the sentences were chosen from grade 1 and 2 storybooks, where words are expected to be frequent, that reflect how individuals read and understand text in natural settings like Brandeis et al. (1994) and Jednoróg et al. (2010) studies. We hypothesize that the N400-related difficulties in lexical and semantic processing affect behavioral measures, leading to low accuracy rather than the validity of the task. Additionally, it is important to consider that our observations may be a consequence rather than a cause. For instance, there may be another underlying cause of reading difficulty that leads to a smaller N400 amplitude, resulting in reading difficulties. Given that this is a correlational study, we do not assume causality. There may be other underlying factors at play, and further research is needed to clarify these relationships. Furthermore, our study only includes a chronological age-matched group comparison. Using a reading age-matched design would provide more specific evidence of deficits.

To our knowledge, this is the first study to investigate the correlation between the N400 and SfV. Replication studies are thus needed. N400 might be a reliable index of SfV knowledge. If this is the case, consideration of casual associations might be strengthened by future intervention studies teaching SfV to RD samples and measuring N400 components at pre- and post-test over a counterfactual intervention condition. Such work is in progress.

Conclusion

In conclusion, we report SfV processing using a sentence incongruity ERP paradigm to explore N400 patterns and SfV in TD children and children with RD. We also report N400-linked effects in congruent versus incongruent printed sentence tasks in English with children under the age of 10, still likely developing English literacy skills. We report an N400 amplitude effect that was more negative in the TD group in the congruent sentences, suggesting reduced lexical and semantic processing in the RD sample. We further explored the relationship between the N400 and reading-related measures. Differential patterns of correlations of N400 components in children with and without RD illuminated distinct processing of words and sentences in TD and RD groups. A novel finding is that SfV correlated negatively with the N400 latency. Future studies might focus on robust replication of these patterns and on using SfV intervention and predicted improvement in reading in young children with RD to also explore possible intervention-driven plasticity in underlying neurobiological N400 components.

References

Adel, A., & Saleh, M. M. (2021). Phonological deficit traits in verbal language of dyslexics. In *Learning Disabilities-Neurobiology, Assessment, Clinical Features and Treatments*. IntechOpen.

Aiello, G., Merati, S., & Toti, M. P. and Aiello I (2018) Predictive value of the auditory and visual evoked potentials using the wavelet method and visual n400 in subjects affected by dyslexia and in normal subjects. Dual Diagn Open Acc, 3(3), 7. DOI: 10.21767/2472-5048.100040

Araújo, S., Faísca, L., Reis, A., Marques, J. F., & Petersson, K. M. (2016). Visual naming deficits in dyslexia: An ERP investigation of different processing domains. *Neuropsychologia*, *91*, 61–76 https://doi.org/10.1016/j.neuropsychologia.2016.07.007.

Barnes, M. A., Clemens, N. H., & Miller, A. H. (2022). Contributions of cognitive science to special education research and practice: Historical context, current influences, and future directions. In *Handbook of Special Education Research, Volume I* (pp. 24-39). Routledge.

Battle, L., Crouser, R. J., Nakeshimana, A., Montoly, A., Chang, R., & Stonebraker, M. (2019). The role of latency and task complexity in predicting visual search behavior. *IEEE transactions on visualization and computer graphics*, *26*(1), 1246-1255. https://doi.org/10.1109/TVCG.2019.2934556

Bishop, D. V., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of child* *psychology and psychiatry*, *31*(7), 1027-1050 https://doi.org/10.1111/j.1469-7610.1990.tb00844.x

Brandeis, D., Vitacco, D., & Steinhausen, H. C. (1994). Mapping brain electric micro-states in dyslexic children during reading. *Acta Paedopsychiatrica*, *56*(3), 239–247.

Berninger, V. W., & Abbott, R. D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of educational psychology*, *102*(3), 635. https://psycnet.apa.org/doi/10.1037/a0019319

Castles, A., Coltheart, M., Larsen, L., Jones, P., Saunders, S., & McArthur, G. (2009). Assessing the basic components of reading: A revision of the Castles and Coltheart test with new norms. *Australian Journal of Learning Difficulties*, *14*(1), 67–88. https://doi.org/10.1080/19404150902783435

Coch, D. (2015). The N400 and the fourth grade shift. *Developmental science*, *18*(2), 254-269. https://doi.org/10.1111/desc.12212

Coch, D., & Benoit, C. (2015). N400 Event-Related Potential and Standardized Measures of Reading in Late Elementary School Children: Correlated or Independent? *Mind, Brain, and Education*, *9*(3), 145–153. https://doi.org/10.1111/mbe.12083

Coch, D., & Holcomb, P. J. (2003). The N400 in beginning readers. *Developmental Psychobiology*, *43*(2), 146–166. https://doi.org/10.1002/dev.10129

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... & Van Petten, C. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical neurophysiology*, *120*(11), 1883-1908. https://doi.org/10.1016/j.clinph.2009.07.045

Edwards, A. A., Steacy, L. M., Siegelman, N., Rigobon, V. M., Kearns, D. M., Rueckl, J. G., & Compton, D. L. (2022). Unpacking the unique relationship between set for variability and word reading development: Examining word- and child-level predictors of performance. *Journal of Educational Psychology*, *114*(6), 1242–1256. https://doi.org/10.1037/edu0000696

Ehri, L. C. (1998). Grapheme–Phoneme Knowledge Is Essential for Learning to Read Words in English. In *Word Recognition in Beginning Literacy*. Routledge.

Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A.-M. (2012). From Spelling Pronunciation to Lexical Access: A Second Step in Word Decoding? *Scientific Studies of Reading*, *16*(4), 341–359. https://doi.org/10.1080/10888438.2011.568556

Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*(1), 1–12. https://doi.org/10.1111/j.1469-8986.2006.00381.x

Gibson, E. J. (1965). Learning to Read: Experimental psychologists examine the process by which a fundamental intellectual skill is acquired. *Science*, *148*(3673), 1066-1072. https://doi.org/10.1126/science.148.3673.1066

Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2013). The time course of reading processes in children with and without dyslexia: An ERP study. *Frontiers in Human Neuroscience*, *7*. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00570

Helenius, P., Parviainen, T., Paetau, R., & Salmelin, R. (1999). Neural processing of spoken words in specific language impairment and dyslexia. *Brain*, *132*(7), 1918-1927. https://doi.org/10.1093/brain/awp134

Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language*, *117*(2), 88–99. https://doi.org/10.1016/j.bandl.2010.12.003

Jednoróg, K., Marchewka, A., Tacikowski, P., & Grabowska, A. (2010). Implicit phonological and semantic processing in children with developmental dyslexia: Evidence from event-related potentials. *Neuropsychologia*, *48*(9), 2447-2457. https://doi.org/10.1016/j.neuropsychologia.2010.04.017

Khalifian, N., Stites, M. C., & Laszlo, S. (2016). Relationships between event-related potentials and behavioral and scholastic measures of reading ability: A large-scale, cross-sectional study. *Developmental Science*, *19*(5), Article 5. https://doi.org/10.1111/desc.12329

Kropotov, J. D. (2016). *Functional neuromarkers for psychiatry: Applications for diagnosis and treatment*. Academic Press.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less skilled comprehenders. *Brain and Language*, *102*(1), 30–45. https://doi.org/10.1016/j.bandl.2006.11.001

Lazzaro, I., Gordon, E., Whitmont, S., Meares, R., & Clarke, S. (2001). The modulation of late component event related potentials by pre-stimulus EEG theta activity in ADHD.

International Journal of Neuroscience, *107*(3-4), 247-264. https://doi.org/10.3109/00207450109150688

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293-323. https://doi.org/10.1016/0010-0285(74)90015-2

Levin, H., & Watson, J. (1963). The learning of variable grapheme-to-phoneme correspondences. *A basic research program on reading*, 1-23.

Li, Y., & Wang, M. (2023) A systematic review of orthographic learning via self-teaching, *Educational Psychologist, 58, 1*, 35-56, doi: 10.1080/00461520.2022.2137673

Liesefeld, H. R. (2018). Estimating the timing of cognitive operations with MEG/EEG latency measures: a primer, a brief tutorial, and an implementation of various methods. *Frontiers in neuroscience*, *12*, 765. https://doi.org/10.3389/fnins.2018.00765

McArthur, G., Castles, A., Kohnen, S., Larsen, L., Jones, K., Anandakumar, T., & Banales, E. (2015). Sight Word and Phonics Training in Children With Dyslexia. *Journal of Learning Disabilities*, *48*(4), 391–407. https://doi.org/10.1177/0022219413504996

Meng, X., Tian, X., Jian, J., & Zhou, X. (2007). Orthographic and phonological processing in Chinese dyslexic children: an ERP study on sentence reading. *Brain Research*, *1179*, 119-130. https://doi.org/10.1016/j.brainres.2007.08.046

Morley, A., Hill, L., & Kaditis, A. G. (2016). 10-20 system EEG Placement. *European Respiratory Society, European Respiratory Society*.

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol. 1. Theory and methods* (pp. 357–390). John Wiley & Sons.

Park, H., & Lombardino, L. J. (2013). Relationships among cognitive deficits and component skills of reading in younger and older students with developmental dyslexia. *Research in Developmental Disabilities*, *34*(9), 2946-2958. https://doi.org/10.1016/j.ridd.2013.06.002

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). American Psychological Association. https://doi.org/10.1037/10459-004

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy*, *11*, 67-86.

Plaut, D. C. (1994). Parallel distributed processing challenges the strong modularity hypothesis, not the locality assumption. *Behavioral and Brain Sciences*, *17*(1), 77-78. https://doi.org/10.1017/S0140525X00033483

Robichon, F., Besson, M., & Habib, M. (2002). An electrophysiological study of dyslexic and control adults in a sentence reading task. *Biological Psychology*, *59*(1), 29–53. https://doi.org/10.1016/S0301-0511(01)00118-1

Schulz, E., Maurer, U., van der Mark, S., Bucher, K., Brem, S., Martin, E., & Brandeis, D. (2008). Impaired semantic processing during sentence reading in children with dyslexia: combined fMRI and ERP evidence. *Neuroimage*, *41*(1), 153-168. https://doi.org/10.1016/j.neuroimage.2008.02.012

Seymour, P. H. K., Aro, M., Erskine, J. M., & Network, collaboration with C. A. A. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151-218. https://doi.org/10.1016/0010-0277(94)00645-2

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of experimental child psychology*, *72*(2), 95-129. https://doi.org/10.1006/jecp.1998.2481

Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: The neurobiology of reading and dyslexia. *Development and psychopathology*, *20*(4), 1329-1349. https://doi.org/10.1017/S0954579408000631

Sun, H., Shi, Q., Pazoki, S., Jia, Y., & Woltering, S. (2023). Neurophysiological Correlates of Reading Difficulties in Elementary School Children. *Developmental Neuropsychology*, *0*(0), 1–21. https://doi.org/10.1080/87565641.2023.2225664

Von Koss Torkildsen, J., Syversen, G., Simonsen, H. G., Moen, I., & Lindgren, M. (2007). Brain responses to lexical-semantic priming in children at-risk for dyslexia. *Brain and language*, *102*(3), 243-261 https://doi.org/10.1016/j.bandl.2006.11.010 .

Venezky, R.L. (1999*) The American Way of Spelling: The Structure and Origins of American English ... - Richard L. Venezky—Google Books*. (n.d.). Retrieved July 25, 2023, from https://books.google.com.lb/books?hl=en&lr=&id=jO-_5aZHK-AC&oi=fnd&pg=PA11&dq=Venezky+(1999)&ots=mEtbRTYW7y&sig=viqeHyrbUyGxcCJ9dFeG2q-o2Kw&redir_esc=y#v=onepage&q=Venezky%20(1999)&f=false

Tunmer, W. E., & Chapman, J. W. (2012). The Simple View of Reading Redux: Vocabulary Knowledge and the Independent Components Hypothesis. *Journal of Learning Disabilities*, *45*(5), 453–466. https://doi.org/10.1177/0022219411432685

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). Test Review: Comprehensive Test of Phonological Processing (CTOPP). *Canadian Journal of School Psychology*, *30*(2), 155-162.

Wang, E., Zhou, G., Wang, Y., Huang, H., Hou, Y., & Liu, J. (2017). Word recognition semantic processing defects in Chinese children with developmental dyslexia: An event-related potential study. *NeuroQuantology*, *15*(2). DOI:10.14704/nq.2017.15.2.1066

Woolley, G., & Woolley, G. (2011). *Reading comprehension* (pp. 15-34). Springer Netherlands

Table 2.1

Demographic table of the study population.

	TD (n = 22)	RD (n = 32)
Age (years)	9	9.2
Sex (male/female)	15/7	14/18
Dominant language at home	English	English
Mothers Education	Bachelor and above	Bachelor and above

Table 2.2

Mean and SD of TD and RD based on standard scores of reading measures.

	TD (n = 22)		RD (n = 32)	
	Mean	SD	Mean	SD
WIAT – RC**	100.95	21.53	72.19	20.83
WIAT – WR**	113.64	15.42	72.28	12.02
WIAT - EX. VOCA*	95.82	19.04	81.47	15.02
CTOPP - EL**	11.41	2.36	6.63	2.02
CC2 - WR**	32.64	10.74	14.56	11.53
CC2 - IR**	24.36	9.27	9.47	8.29
CC2 - NW**	26.27	13.16	9.72	9.85
SfV**	13.68	3.84	10.63	3.56

Key: WIAT RC = reading comprehension; WIAT WR = word reading; WIAT EX. Voca = expressive vocabulary; CTOPP - EL = Elision; CTOPP - PA = phonological awareness; CC2-WR = Castle and Coltheart word reading; CC2- IR, Castle and Coltheart irregular word reading; CC2- NW= Castle and Coltheart non-word reading.
p values * < 0.05, ** < 0.001.

Table 2.3

Mean and SD (in parenthesis) of Reaction time and accuracy of the N400 ERP task in the congruent and incongruent conditions.

	Congruent		Incongruent	
	RD (*n* =30)	TD (*n* = 21)	RD (*n* =30)	TD (*n* = 21)
Accuracy	52.36 (21.56)	88.35 (17.09)	54.18 (25.94)	87.84 (18.74)
Reaction Time	1173 (289.41)	920.18 (307.70)	1214 (332.78)	949 (339.53)

Table 2.4

Mean and SD (in parenthesis) of N400 Amplitude and peak Latency of both groups.

	Congruent		Incongruent	
	RD (*n* =30)	TD (*n* = 21)	RD (*n* =30)	TD (*n* = 21)
N400 Amplitude	-4.73 (1.44)	-7.31 (1.74)	-4.25 (1.07)	-9.57 (2.26)
N400 Latency	402.92 (11.16)	393.69(11.91)	411.45 (10.49)	393.97 (11.52)

Table 2.5 *Correlation of the reading measures in both groups.*

	1	2	3	4	5	6	7	8
				TD (*n* =22)				
1.RC	1							
2.WR	0.1	1						
3.Ex. Voca.	**.60****	0.29	1					
4.PA	**.49***	0.32	**.55****	1				
5.CC2 WR	0.11	**.80****	-0	-0.03	1			
6.CC2 IR	0.16	**.69****	0.13	-0.02	**.91****	1		
7.CC2 NW	0.16	**.85****	0.17	0.24	**.91****	**.82****	1	
8.SfV	**.51****	**.61****	**.47***	**.66****	**.54****	**.49***	**.66****	1
				RD (*n* = 32)				
1.RC	1							
2.WR	**.59****	1						
3.Ex. Voca.	0.1	0.17	1					
4.PA	0.3	0.31	.49	1				
5.CC2 WR	**.67****	**.75****	0.3	0.25	1			
6.CC2 IR	**.55****	**.78****	0.18	0.12	**.85****	1		
7.CC2 NW	**.68****	**.71****	0.12	0.27	**.82****	**.63****	1	
8.SfV	**.49****	.36	0.21	-0.03	**.69****	**.63****	**.59****	1

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant with Bonferroni correction for multiple correction (p = .006).

Table 2.6

Correlation between the N400 and SfV in the TD group

	SfV
N400 Effect	**.46***
N400 Latency	.27

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

Table 2.7

Correlation between the N400 latency and SfV in the RD group

	SfV
N400 Effect	.02
N400 Latency	**-.43***

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

Figure 2.1

N400 effect in the TD and RD group.

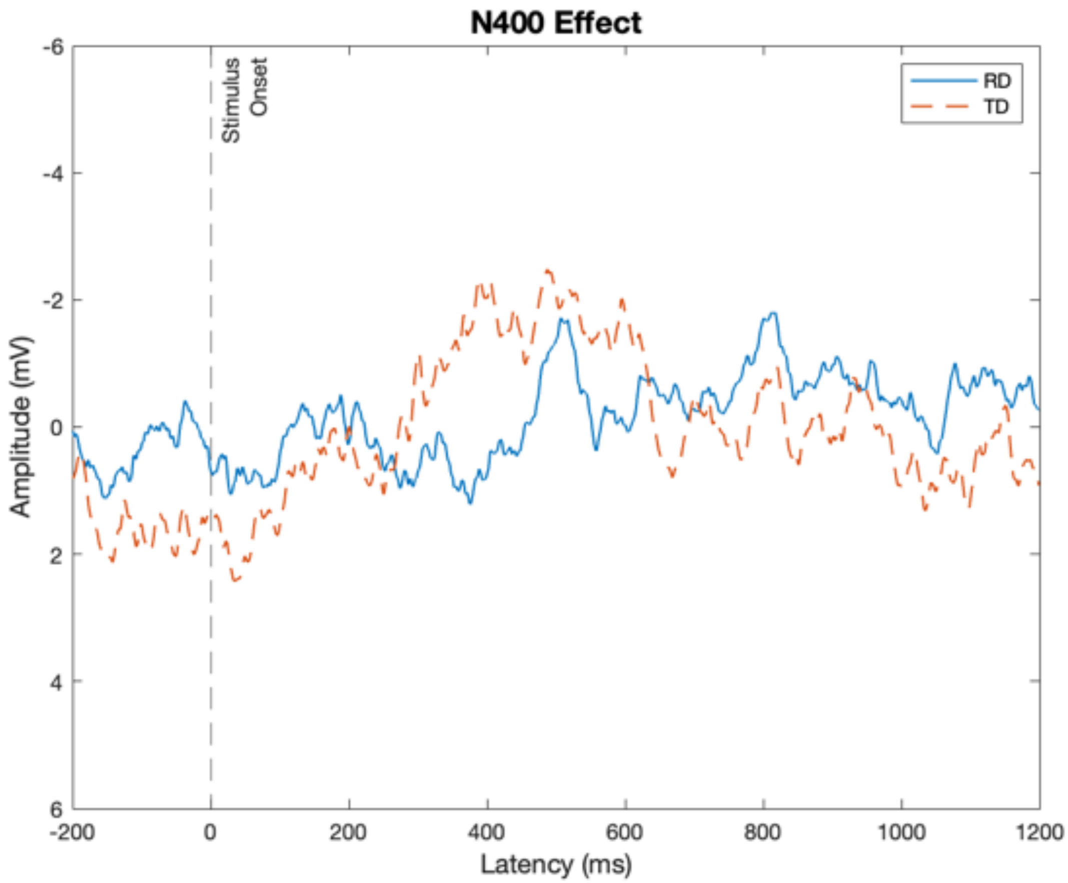
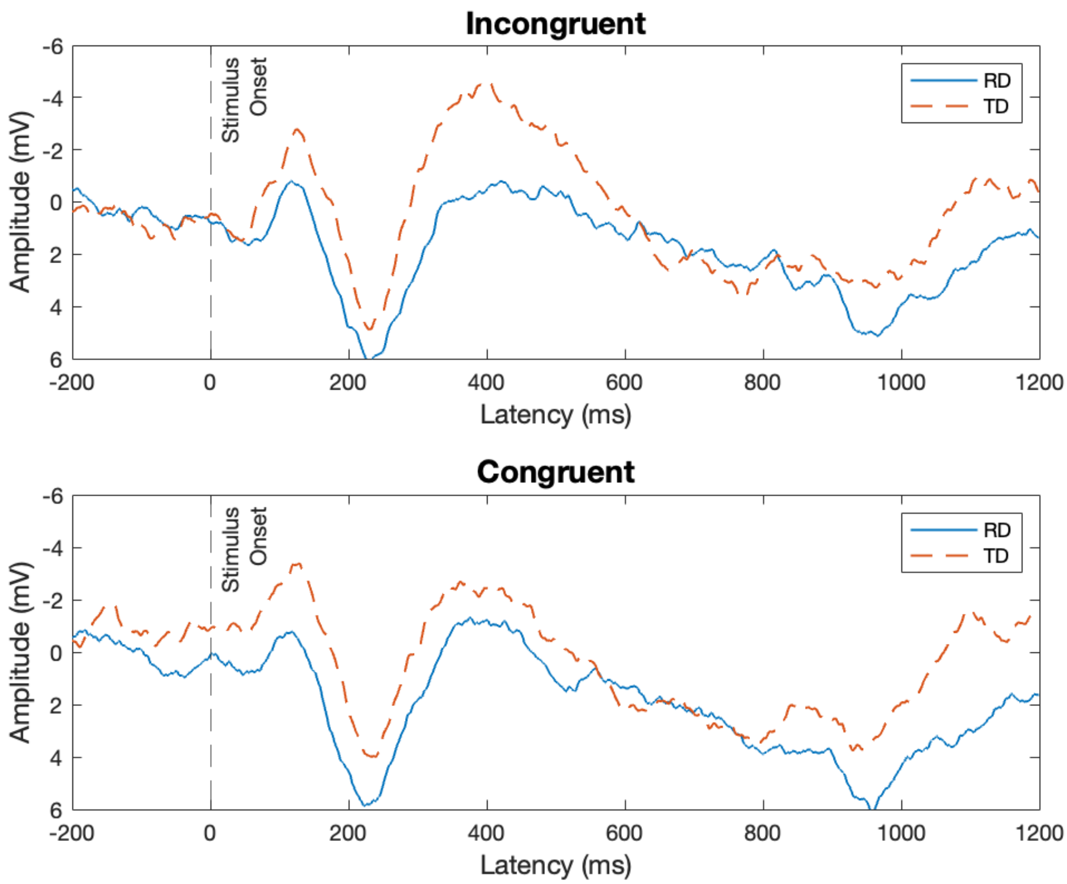


Figure 2.2

N400 mean amplitude in the congruent and incongruent conditions in the two groups



Acknowledgements

We would like to express our sincere gratitude to Dr. Dave St Amour for generously providing us with all the EEG laboratory equipment that was necessary for the study to take place. We extend our gratitude to the school boards, the families, and the participants for their collaboration, time, and effort in completing this study. We would also like to thank Dr. Jesse Michaels for generating figure 1.

Funding

This research was partially funded by a William Dawson Research Scholar Award to Dr. Armando Bertone and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC; RGPIN-2020-05052) to Dr. Gigi Luk.

Chapter 7. Bridging Manuscript 2 and Manuscript 3

In Manuscript 2, the results revealed that (i) SfV delays exist in RD over matched TD readers in group contrasts, (ii) the N400 was aberrant in children with RD, and (iii) there was a correlation between reading measures and N400 amplitude and latency. The hypothesized correlation path differed between the two groups, with a correlation found in N400 latency in the RD group. While previous studies have enhanced our comprehension of how SfV affects both typical and struggling readers, there is a lack of evidence concerning children who face difficulties with word reading. At present, there is limited evidence indicating whether interventions targeting SfV lead to improvements in cognitive flexibility beyond the specific words taught. It is crucial to establish substantial evidence of the transfer of skills to untaught exception words to propose the argument that SfV plays a vital role in the developmental process of acquiring irregular word. The next step involved using the N400 as a neural measure in a targeted reading intervention using SfV. The intervention teaches children with RD a strategy to pronounce printed words using variant Grapheme-Phoneme Correspondences (GPCs) and their lexical knowledge (vocabulary knowledge of words) to correct pronunciation in a randomized control study (Chapter 8) to examine potential reading interventions such as SfV and its link to N400 in English. The N400 would be measured pre-post to determine if it predicts improvements in reading made in the SfV condition over an active control intervention (CBP).

Chapter 8. Manuscript 3

The behavioral and neurophysiological effects of set-for-variability reading intervention for children with sustained word reading difficulties.

Badriah Basma (Corresponding author): Department of Education and Counselling Psychology
McGill University, Montreal, Quebec, Canada
Email: badriah.basma@mail.mcgill.ca
Phone: +1 (514) 802-2675

Robert Savage: Faculty of Education
York University, Toronto, Ontario, Canada
Email: robert.savage@edu.yorku.ca

Gigi Luk: Department of Education and Counselling Psychology
McGill University, Montreal, Quebec, Canada
Email: gigi.luk@mcgill.ca

Armando Bertone: Department of Education and Counselling Psychology
McGill University, Montreal, Quebec, Canada
Email: armando.bertone@mcgill.ca

Correspondence to:

Badriah Basma, Department of Education & Counselling Psychology, McGill University, 3700 McTavish St, Montréal, QC, Canada, H3A 1Y2.

Tel: +1 514 398 6655; e-mail: badriah.basma@mail.mcgill.ca

Abstract

Purpose: Set-for-variability (SfV) reading intervention assists individuals with a reading disability (RD) in producing alternative pronunciations for unfamiliar words. Whether such interventions result in neural change, as examined through eth N400 event-related potentials (ERPs), is rarely assessed. The first aim of this study is to assess if SfV intervention improved reading skills in children with RD. The second aim is to assess if SfV intervention improves the N400 characterized by earlier latency and an increase in N400 amplitude in the incongruent condition of the N400 task.

Methods: A randomized control trial (RCT) of a set-for-variability reading intervention was implemented in 30 children, mean age- 9 years with sustained word reading difficulties. Reading and N400 ERP measures were obtained pre and post a ten-week reading intervention. Two reading intervention groups participated in a 10–12-hour online program over ten weeks. The Set for Variability group focused on managing phoneme variation in decoding regular and exception words using synthetic phonics, while the Current Best Practices group received synthetic phonics and sight word reading.

Results: Analyses of covariance yielded significant interaction effects whereby the Set-for-Variability group demonstrated (1) post-test improvement in reading exception words and oral set-for-variability and (2) increased negative N400 amplitude and improved reading accuracy in the incongruent condition of the N400 task.

Conclusion: (1) set-for-variability may be an important reading intervention that helps children manage print-sound inconsistency in reading acquisition in English, and (2) the N400 pattern (amplitude) is a neurological correlate of behavioural changes evident with this efficacious literacy intervention. Keywords: *reading intervention, poor readers, N400, set-for-variability*

Introduction

Reading comprehension is a complex process that depends on language abilities, including word recognition and oral vocabulary (Dyson et al., 2017; Perfetti, et al., 1992; Landi & Perfetti, 2007; Zipke, 2016). Skilled word reading in English involves negotiating the marked inconsistency of English orthography (Share, 1999, 2011). Substantial empirical evidence strongly suggests that poor word reading is one of the primary causes of reading comprehension deficits (e.g., Hulme & Snowling, 2011). Efficient word reading requires precisely specified orthographic and phonological representations of words (Perfetti et al., 1992). Strong vocabulary representations are closely associated with high-quality orthographic and phonological representations of words (Perfetti & Hart, 2001). Effective interventions for struggling readers must address likely weaknesses in these lexical domains.

Substrates of these neurobiological processes correspond to the acquisition of these exact lexical representations. For example, Landi and Perfetti (2007) suggest that there are Event-Related Potential (ERP) differences, particularly within the N400 range, are observable in semantic tasks in adults with reading difficulties. Importantly, analogous patterns of N400 deficit are also evident in children with poor reading skills (Author et. al., submitted). This present study seeks to assess the effects of SfV remedial reading intervention explicitly theorized and constructed to help children with reading disabilities manage printed word-to-sound (orthography-to-phonology) inconsistency in English (Seymour et al., 2003), characterized by improvement in reading skills, specifically exception word reading. and to assess if the SfV intervention would induce a change in the N400 amplitude and onset latency, characterized by an increase in the N400 amplitude in the incongruent condition (semantic anomaly) and earlier onset latency.

Teaching RD young children

Teaching phonics to individuals with reading difficulties (RD) is effective (Ehri,2020). In synthetic phonics, young children learn grapheme-phoneme correspondence (GPC) rules to read words (Share, 2008). Beginner readers can then independently derive the pronunciation of a new word by applying their knowledge of GPC relationships to synthesize word pronunciations, and subsequently storing word knowledge. This concept is known as the self-teaching hypothesis (Share, 1995, 1998, 2011) and describes how children progress from spelling out unfamiliar words to rapid reading of the same word. According to Pritchard et al. (2018), children "create an opportunity to self-teach new orthographic knowledge" (p 722). However, Elbro et al. (2012) argue that not all words, even regular ones, can be easily read through GPCs because blending individual phonemes is complicated by GPC inconsistency, the presence of schwas in phoneme strings produced by the phonic assembly, and other fundamental featural differences between linear phonemes strings produced by synthetic phonics and word pronunciations.

Empirically there is evidence indicating that flexibility in applying a range of candidate GPCs to derive word pronunciations is developmentally important. Tunmer and Chapman (2012) conducted a study where children were asked to identify orally presented regularized pronunciation of irregular words, such as yacht spelled as *ya- ch-t* or stomach spelled as *stow-match,* or words such as *kind* spelled as if to rhyme with *pinned*. According to Tunmer and Chapman, to read these words accurately, children must alter one of the variable phonemes of a given grapheme in a word until this produces a word pronunciation they recognize. Once achieved, children can be taught to check the meaning of the word in sentence context, adjust their pronunciation of the word accordingly. or revisit the variable GPC to produce a meaningful sentence. For example, a reader may decode the word "wasp" to rhyme with "clasp." After first noticing that /wæsp/ is not a real word a child would then need to flexibly apply different pronunciations from their listening vocabulary until an acceptable word is found. This strategic mediation between vocabulary and reading irregular words is “Set-for-Variability” (SfV) (Venezky, 1999). SfV is a “process that cleans up the mismatch between orthography to phonology conversion and word pronunciation.” (Steacy et al., 2019, p. 523). Tunmer and Chapman (2012) present data from a three-year longitudinal study involving 140 typical students. Children were assessed on phonemic awareness, vocabulary, decoding skills, syntactic knowledge, SfV, word reading and reading comprehension. Results demonstrated that SfV predicted irregular word reading and indirectly affected reading comprehension via decoding and word recognition. Other studies also report that SfV predicts both regular and pseudoword reading (Elbro et al, 2012; Steacy et al., 2019).

Intervention research also supports the role for SfV and related processes. Dyson and colleagues (2017) conducted a brief four-week intervention study focusing on teaching children the ‘mispronunciation correction’ technique to explore potential improvements in irregular word reading post-intervention. In their study, 81 children from grades one and two participated, randomly assigned to either the intervention or control group. In the mispronunciation correction group, children were asked to sound out a set of irregular words and determine if they knew the word. If unfamiliar, they were trained to substitute it with a word sounding like the target word. Post-training, children in the mispronunciation correction group exhibited better reading skills for irregular words, demonstrating the ability to apply the mispronunciation correction strategy to an additional set of untaught experimental words. This suggests a potential generalizing effect of mispronunciation correction in reading irregular words that are unfamiliar to the child. However, this pattern was not evident in another set of standardized exception words, the Castle and Coltheart 2-word list (CC2, Castles et al., 2009). Similarly, Colenbrander et al. (2022)

provided brief (10-minute) mispronunciation correction instruction to typical kindergarten children and found improved orthographic representation of trained words but no evidence of improved representation for transfer words.

Savage et al. (2018) conducted a longer 10-week intervention, teaching both SfV and ‘direct mapping’ (reading texts rich in the specific grapheme-phoneme correspondence lessons recently taught) with grade one children at risk of developing reading disabilities. The SfV strategy was also taught through an approach focused mainly on managing alternate vowel and consonant pronunciations to derive conventional word pronunciations. Results indicated that children in the SfV group had significantly higher scores in standardized word reading and spelling at the post-test, along with sentence comprehension at a delayed post-test five months after the intervention concluded.

While existing intervention studies have contributed to our understanding of the impact of SfV on typical and struggling readers in Grades 1 and 2, evidence is lacking regarding older children with more sustained word reading difficulties. Currently, only modest evidence exists regarding whether SfV intervention generalizes to improve mental flexibility in SfV and the generalization of exception word reading beyond taught items. Establishing clear evidence of ‘transfer’ generalization to untaught exception words is essential for supporting claims that SfV is a developmental process for irregular word acquisition, and that SfV instruction is similarly important. These key claims are assessed in this study.

SfV Instruction and Event-Related Potentials

From a neurophysiological perspective, there is evidence to suggest that struggling readers also show distinct Event-Related Potential (ERP) profiles. A recent systematic review and meta-analysis by Basma et al., (2023) aimed to determine whether significant differences in

N400 response to various lexico-semantic tasks exist between neurotypical readers and readers with dyslexia. Twenty studies met stringent selection criteria and were included in the meta-analysis, revealing a relatively large overall effect size (Hedge's $g = 0.65$, $p < 0.001$) for typical versus atypical readers. Of particular interest are the larger effects associated with visual (object and print) modality ($g = 0.692$, $p < 0.001$) and for semantically incongruent sentence tasks ($g = 0.948$, $p < .001$) (over non-sentence tasks ($g = 0.6$, $p < .01$) and Chinese ($g = 0.968$, $p < .001$) versus alphabets ($g = 0.77$-0.79, $p < .001$). Notably, no analyses demonstrated strong evidence of publication bias. These results suggest that the N400 reliably indexes deficient processes during reading, highlighting differences between typical and atypical readers. The stronger associations of N400 with semantically incongruent sentences and printed words suggest implications for semantic problem-solving in reading, resembling processes related to SfV. Basma et al., (2023) provided preliminary evidence suggesting a negative association between SfV and N400 latency incongruent conditions, implying that proficient readers engage in earlier word information processing than poor readers. Minimal direct data exists on this specific question.

To date, only one study (Hasko et al., 2014) has approached ERPs using a behavioural reading intervention approach. Hasko et al. demonstrated changes in the N400 amplitude post-reading intervention delivered in German. Two interventions were conducted. The first focused on grapheme-phoneme correspondence, while the second focused on letter pronunciation. Information on intervention-specific improvement was not reported by Hasko et al., (2014), nor was it for the intervention involved semantic processes or the teaching of SfV. German orthography, unlike English, has few exceptions to GPC rules (Seymour et al., 2003). Importantly, Hasko et al. (2014) categorized children receiving intervention at post-intervention into treatment responders and non-responders rather than exploring the overall treatment group effects of a randomized control trial (RCT). RCTs typically offer stronger methodological control for extraneous variables and better error estimates, making them a preferred approach here. Distinct exploration of candidate SFV reading interventions and ERP-SfV links in English is thus warranted.

Our research study aims to address two questions. The first is whether SfV intervention that focus on teaching children with reading disabilities to manage grapheme-phoneme inconsistency in written English improves reading skills. We hypothesize that children who receive SfV intervention will exhibit improvements in reading skills, particularly in irregular word reading and the Tunmer and Chapman oral SfV measure (Tunmer &Chapman, 2012). The second question explores whether an SfV reading intervention, having improved reading skills, changes the N400 in incongruent conditions, marked by earlier latency and an increased negative amplitude. We hypothesize observing both an increased amplitude on the N400 in the incongruent condition and an earlier latency post-intervention after the SfV intervention.

Method

Participants. General methods and recruitment procedures was similar to that used by Basma et al., submitted) with participants included in their comparison-control study. Thirty (N=30) children, aged between 6.5 and 11.5 years (Mean age = 9.2) participated in this study (Male = 12; Female: 18). English was the primary language spoken at home for all participants. The inclusion criteria to be included in the intervention study encompassed either a confirmed diagnosis of RD or performance two grades below the current grade level on the WIAT Word Reading and WIAT Reading Comprehension standard scores, along with a below-average score on the Phonological Awareness composite score of the Comprehensive Test of Phonological Process (CTOPP) standardized test. Sixty-three percent (63%) of the total sample had a confirmed diagnosis of a specific learning disability in reading based on DSM-5 (APA, 2013) criteria, while 17% had reports from parents and teachers indicating reading performance below age-expected grade level, 10% were referred from speech and language clinics due to reading difficulties, and 10% were transferred from speech and language clinics as reading one or two grades below grade level. Ethics approval was obtained from the university board of ethics before the start of data collection.

General Procedure Design. The pre-intervention procedures were like that of Author et al., (submitted). Following a brief laboratory tour, an overview of the procedures, and signing assent/consent forms, all participants were then asked to complete a series of reading measures (refer to the Reading Measures section below for details) in a designated quiet room. After completing the reading tests, participants were seated in a separate, sound-attenuated room where they were placed 70 cm away from a visual display (ViewPixx©), and fitted with an electrode cap. As a form of distraction, participants were given the option to watch either Netflix or Disney channel while the electrodes were being fitted. Once the cap was in place, participants performed eye movements and jaw clenching, with the RA capturing images to demonstrate how these actions are reflected in brain waves. Participants were then instructed to sit, relax, and minimize eye blinking unless they saw a "+" on the display. The stimuli task was thoroughly explained, followed by a brief practice session involving four tasks. The task only proceeded once the participant was ready. Every 25 trials, the participant was given the option to have a longer break to prevent fatigue. The task resumed only when the participant indicated readiness by pressing the keyboard bar. The entire process, including reading and EEG measures along with planned and unplanned breaks, was completed within 2.5 hours.

A randomized, two-arm, pre-test-post-test trial study of the effects of a SfV reading intervention on (1) behavioral reading and SfV outcomes, and (2) the N400 amplitude and latency was then implemented. First, pairs of children were matched as closely as possible on the Wechsler Individual Achievement Test (WIAT) Word Reading was measured and then randomly placed in either the treatment or control arms of the trial. The interventions were delivered via Zoom during the Covid-19 pandemic. The treatment arm then received a Set-for-Variability intervention with phonics and item vocabulary, and the control arm received Current Best Practices (CBP) teaching, which exposed children to the same process that SfV was exposed to but in the absence of the SfV component (interventions are described in detail below). Children in both groups received 10 weeks of teaching (13.43 of contact) from research assistants (RAs). The interventions were delivered via Zoom during the Covid-19 pandemic.

Reading measures

<u>Wechsler Individual Achievement Test-III</u> (WIAT-III; Wechsler, 2009) is a standardized measure of academic achievement in the subjects of reading, writing, math, and spoken language. The study employed the standard scores of the following three WIAT subtests (1) *WIAT Reading Comprehension*. A subtest to assess the reader's inferential and literal reading comprehension abilities through the reading of various passages and the completion of various question types. Starting at the appropriate grade level, the reader has the option of reading silently or aloud. The examiner asks comprehension questions following each passage. The examiner moves back to one grade level if the reader receives two points or less on the passage comprehension questions. If the reader goes back three grades, the test is stopped. The published

internal reliability coefficient is *r* = 0.90. (2) *Word Reading* is a subtest to assess a reader's capacity for accurate word identification. A list of words is provided for the reader to read out loud. As the person reads on, the difficulty level of the words rises. The reader receives 1 point for each accurate response. If the reader commits four consecutive errors, the measure is discontinued. The published internal reliability coefficient is *r* = 0.98. (3) *Expressive Vocabulary* measure assesses a reader's vocabulary and word retrieval skills. After presenting them, the examiner asks the reader to say the term that goes with the given picture and definition. Testing is discontinued after four consecutive errors. The published internal reliability coefficient is *r* = .75.

<u>Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999)</u>. This test measures the phonological processes required for successful reading abilities. This study used three subtests. *Elision:* This test assesses the person's capacity to eliminate syllables and, later, phoneme sounds and syllables from a spoken word. This subtest contains 34 items. Participants were asked to say compound terms, and then the examiner asked them to say the word after dropping a word. As an illustration, the examiner asks the participant to state "cowboy." After the participant responds, the examiner requests that they say "cowboy" without "boy." The test increases in difficulty as the participants continue. After three consecutive errors, the test is terminated. The standard score was used in the analysis. The published internal reliability coefficient is *r = .82. Blending words.* This subtest is designed to assess a participant's capacity for blending words to produce the intended word pronunciation. This subtest contains 33 items. For example, the participant hears an audio that asks them to put the letters "t" and "oi" to form the word "toy." After three consecutive measures, the subtest is ended. The published internal reliability coefficient is *r* =.75. *Phoneme isolation*. This subtest is designed to assess the

participant's proficiency in identifying individual sounds in words. This subtest contains 32 items. In vowel consonant vowel (CVC) words, the participant is usually asked to identify the first and last sounds. The participant is next instructed to pick out the middle vowel words with four and five sounds. The published internal reliability coefficient is *r* =.83

Castles and Coltheart reading test (CC2; Castles et al., 2009). consists of three sets of words: 40 irregular words, 40 regular words and 40 nonwords. The examiner hands the participant one card at a time to read. As the reader continues, the items become more difficult. When five consecutive errors are made within a subscale, testing is discontinued. According to McArthur et al. (2015), this test has a reliability of 0.92 for regular words, 0.94 for irregular words, and 0.85 for nonwords.

Set-for-Variability (SfV; Tunmer & Chapman, 2012). This test is based on standard approaches in the field from Tunmer & Chapman's (2012) study. The test measures children's ability to pronounce regularized exception words correctly. For this study, twenty items from Tunmer and Chapman's (1998) mispronunciation task were selected and recorded to ensure consistency. The examiner read the script: "I have a friend who will read some sentences, but they will say the wrong word at the end of each sentence. Could you help my friend by saying the right word?" After the script, children listened to recorded sentences with mispronounced words at the end. Examples include:

He couldn't find his **money** (pronounce: moaney)

In France, they have great **weather** (pronounce: weet-her)

The man repaired the broken **watch** (pronounce: rhyme with catch)

Responses were scored as 1 (correct) or 0 (incorrect). The published inter-rater reliability is r = .86 (Tunmer & Chapman, 2012). Raw scores were used in the analysis since this test does not have standard scores.

Electrophysiological measures. Participants completed a sentence-judgement task adapted from Brandeis et al. (1995) and Sabisch et al. (2006) studies. The sentence-judgment task was implemented using PsychoPy© software. The task involved presenting visually displayed sentences that were either semantically congruent or incongruent, composed of words commonly found in children's storybooks. The sentence structure was either subject-verb-object or subject-verb-prepositional object, with the critical word being the last word determining semantic congruency or incongruency. To familiarize the participants with the task, a research assistant, trained for this study, provided instruction and example trials were demonstrated. The task ad 100 sentences; 50 congruent and 50 incongruent semantic endings. The task was designed into four blocks. Each block included 25 sentences with randomized congruent and incongruent sentences. The sentence length ranged from five to seven words, and the final word varied from four to eight letters. Examples of the sentences in the congruent condition included:

1. My father is eating a pizza.
2. Leaves turn color in the fall.

The incongruent sentences were created from the congruent ones by replacing the critical word with a word (the final word of the sentence) incongruent with the meaning of the preceding words. The following are examples of incongruent condition sentences:

1. My father is eating a sock.
2. Leaves turn color in the lamp.

Each trial started with a yellow fixation cross + displayed for 600 ms, followed by a sequential presentation of individual words in lowercase letters, each lasting 700 ms at the fixation point. . The final (critical) word or non-word in each sentence string was succeeded by a period. Participants were instructed to read each sentence silently and then prompted with a blue fixation + for 2500 ms, signaling the need for a judgment via a button press; green for meaningful and red for not meaningful. Between each stimulus block, participants were given breaks to breathe, stretch, and relax before progressing to the next block.

The complete ERP measurement took around 45 minutes to complete. Participants received a gift card upon participation in the pre-test and post-test. The participants also received a monetary gift for participating in the intervention. The complete test (reading and EEG measures) lasted 2.5 hours.

Event Related Potential (ERP) data acquisition and processing. ERPs were recorded during task completion using a Brain Vision Quickamp system with 64 channels. Electroencephalography (EEG) data were collected from 32 active Ag/AgCl electrodes positioned in an elastic cap following the 10/20 system (Morely, 2016), with references at the left and right mastoid and AFz as the ground. Impedance was maintained below 5 kΩ to minimize artifact. Data were filtered with a low cut-off of 0.1 Hz and a high cut-off of 30 Hz. Independent Component Analysis (ICA) was used to remove artifacts like eyeblinks before averaging. Additional artifacts, including those with a gradient exceeding 15 µV in a 200 ms window, low activity below 0.5 µV in a 200 ms window, and amplitude over 200 µV in a 50 ms window, were also removed. Vertical VEOG was recorded with electrodes placed above and below the right eye, while horizontal EOG (HEOG) was recorded with electrodes placed 1.5 cm lateral to the left and right external canthi. EEG data were epoched offline from -200 to 1200 ms, with the onset of the final word in each sentence set at 0 ms. Trials with electrooculogram (EOG) artifacts, such as high-amplitude signals due to blinking or low-frequency signals from eye movements exceeding ±75 µV, were excluded. A 64-channel waveform plot was initially inspected for all participants. Based on previous N400 research (Duncan et al., 2009), the 300-500 ms time window and specific electrodes were selected for further analysis. The final analysis included 19 channels: F3, FZ, F4, C3, CZ, C4, CP5, CP1, CP2, PZ, P3, P4, FC1, FC2, OZ, O1, O2, T7, and T8, following the guidelines by Duncan et al. (2009) and the review by Šoškic et al. (2022). Each participant required at least 35 clean trials for inclusion in the analysis, and thus three participants with fewer than 35 clean trials were excluded.

Intervention

Set-for-Variability Intervention. The main purpose of SfV is to develop children's capacity to manage GPC inconsistency in producing approximate phonological representations for unfamiliar written forms of words, and children are aided in matching phoneme strings derived from phonic assembly to conventional word forms. To build this repertoire of spelling-sound-variation, a researcher-designed Set-for-Variability intervention was adapted from the Savage et al. (2018) study by having training sessions that focus on teaching students how to alternate consonant and vowel digraph pronunciations and reading regular, as well as irregular words. Foundational synthetic phonics is assumed in the SfV intervention. In addition to the Savage et al. (2018) adaptation, we taught children attention to exception word item oral vocabulary. The Common and Best-Practice (CBP) control intervention content was also adapted from Savage et al., (2018).

Intervention Procedure. Research assistants (RAs) received 2.5 hours of training in SfV and CBP intervention. The intervention training was delivered in person by the first author, with the second author assisting. The training consisted of instruction to teach both methods, a lesson plan and materials for each session. All the lessons were on a PowerPoint slide accessible via OneDrive. All RAs were randomly assigned to the participants. In addition, each RA taught one participant in the SfV and one in the CBP to reduce RA-based bias. All RAs could contact the lead researcher with questions at any point during the intervention. The children received an intervention online for 30 minutes three times a week or 90 minutes on a weekend administered online via Zoom because of the pandemic. Children received an average of 12-14 hours of intervention. All lessons included a review of the previous lesson (2-5 minutes), teaching (5 minutes), practice activity game (10 minutes) and vocabulary (5 minutes) – SfV (5 minutes). In the CBP group, the participants received sight word instruction instead of SfV (5 minutes).

Set-for-Variability. Children were taught using a variation of the 5-step model described in Boldrini et al., 2022 (see Appendix 1 for precise prose used here). In this approach, the children who used phonic blending were encouraged to evaluate the lexicality of the result and to modify their response by replacing a variant GPC with another pronunciation and re-synthesizing the phonemes. To get to this stage of using SfV to read words, students received training in *grapheme-phoneme correspondence and phonic blending*, *item vocabulary*, and *oral set-for-variability*. These are prerequisites for accurate mispronunciation correction using SfV (Boldrini et al., 2022). The first step is to teach children that a grapheme can have multiple pronunciations. To do that, we teach students the common variability in vowels and consonants starting with six common consonant variants in the English language: *s* in *s*it and *s* in bin*s*, *c* in face and *c*ool, *y* in *y*ou and happ*y*, *g* in *g*orilla and ca*g*e, *th* in *th*ings and *th*e, and finally ed in finish*ed* (t) and pour*ed* (d). Children then learned the frequent and variant pronunciation of vowels: ee, ea, oo, ou, oa, ai, ay. They received training in blending common consonant-consonant structures (CC) (sp- st, bl, dr, cr) and final -CC in words such -st, -nd, -ld, -sk, -mp within CCVC and CVCC structures delivered through games. Once students master these steps, the second step is to teach students oral SfV for words and exception words. In doing so, SfV teaching used both phoneme strings and regularised word pronunciations. This aims to teach students to link spelling pronunciations to regular and exception words. The third step is to teach the vocabulary of exception words. The final step is to teach SfV for mispronunciation correction, at this stage, using printed words.

Current Best Practices (CBP). The control group received Current-Best Practices (CBP). The participants received all of the same content as the SfV group except for the SfV content, which was replaced by sight word instruction using flashcards.

Treatment Integrity. To ensure the validity of the intervention, Treatment Integrity was assessed using a substantially modified form of the Treatment Integrity rubric used by Savage et al. (2018). The rubric consisted of four criteria: (i) the contents of the intervention being applied. (ii) Time management (iii) Teaching quality (iv) Intervention components. The lead researcher and an assigned RA observed all RAs at different times and evaluated their practice according to a given rubric to ensure that the RA was delivering the intervention according to the training (Borrelli, 2011). The inter-rater reliability for the treatment integrity assessed by the Pearson correlation coefficient was .92.

Results

Data analyses Participants were matched based on age and gender. T-tests were conducted to compare the SfV group and the CBP group on these variables. Results indicated no statistically significant differences in age, t (28) = 1.20, p = .240 or gender t (28) .357, p = .724. Demographic information is detailed in Table 1. The descriptive statistics of the pre-and post-intervention reading measures for the SfV and CBP groups are presented in Table 2, and reading measures correlation at pre-and post-test are presented in Tables 3 and 4. The N400 behavioral measures (accuracy and reaction time) are in Tables 5 and 6. The accuracy rate for the N400 task is the percentage of correct responses out of the 100 given trials. The reaction time for the N400 task measures how quickly the participant responded following the onset of the stimulus (final word of the sentence). The N400 amplitude and latency are presented in Tables 7 and 8. Inspection of all tables means standard deviations and effect sizes suggest a change in most measures between pre-and post-tests in both intervention groups. Differential improvement is suggested in the mean differences and effect sizes for oral SfV and exception word reading in the SfV condition in Table 2. Before conducting inferential analyses, the Shapiro-Wilk test of normality showed that all measures reported (Table 1-8) scores were normally distributed, $p > .05$. Skewness scores ranged between -3 to +3 and Kurtosis between -10 to +10, which are acceptable values (Kallner, 2017). Test of Box's $M = 1.161$, $p = .179$, indicates that homogeneity of variance matrices, linearity, and multicollinearity were all satisfactory. The groups received an average of 13.43 hours of intervention. The CBP received 13.6 hours and the SfV group received 13.26 hours. A T-test for intervention hours was conducted and showed no significant difference between the number of hours received in each group, t (28) = -1.170, p =.252

Reading Inferential analyses

To inquire about the change in scores between the pre-test and post-test within each group (CBP; SfV), we ran a paired t-test for each group with reading measures as the dependent variable. For the SfV group, the paired t-test results showed a significant change in reading comprehension post-test scores (M = 13.93, SD = 12.4, CI [7.05 – 20.08]), t (14) = 4.34, $p < .001$; WIAT expressive vocabulary (M = 8.30, SD = X, CI [-1.20 – 7.86]), t(14) = 3.81, p = .002;

phonological awareness (*M* = 9.8, *SD* = 14.75, CI [1.63 – 17.96]), *t*(14) = 2.57, *p* = .022; Castles and Coltheart word reading (*M* = 9.33, *SD* = 10.63, CI [1.89 – 8.77]), *t*(14) = 3.39, *p* = .004); Castles and Coltheart irregular word reading (*M* = 5.33, *SD* = 6.20, CI [1.89 – 8.77]), *t*(14) = 5.33, *p* = .005; and finally SfV (*M* = 4.86, *SD* = 3.54, CI [2.9 – 6.8]), *t*(14) = 5.32, *p* < .001.

For the CBP group, the paired t-test results showed a significant change in reading comprehension post-test scores (*M* = 8.80, *SD* = 11.08, CI [2.66-14.93]), *t* (14) = 3.07, *p* = .008 and phonological awareness (*M* = 8.6, *SD* = 12.62, CI [1.60- 15.59]), *t* (14) = 2.63, *p* = .019.

To explore pre-test differences between the two groups, an ANOVA was conducted. The results showed significant pre-test differences between the two groups on CC2 word reading *F* (1, 28) = 4.69, *p* = .039. ANOVA showed no significant differences at the pre-test for reading comprehension *F* (1, 28) = .082, *p* = .777., WIAT word reading *F* (1, 28) = ..668, *p* = ..421, WIAT expressive vocabulary *F* (1, 28) = .1.243, *p* = ..274, CTOPP phonological awareness *F* (1, 28) = ..070, *p* = ..793, CC2 irregular word reading *F* (1, 28) = .1.010, *p* = .324*,* CC2 nonword reading, *F* (1, 28) = ..764, *p* =. 390, and SfV *F* (1, 28) = .367, *p* = .549.

To explore the post-test differences between the groups, a reading Measures x Group (SfV versus CBP) between-subject analysis of WIAT Reading Comprehension, Word Reading, & Expressive Vocabulary, CTOPP subtest (Blending words, phoneme isolation, elision, and phonological awareness), as well as Castle and Coltheart Word reading, irregular word reading, and nonword reading subtests, and SfV was conducted. The independent variable was Group (SfV, CBP). The main effect of the group was significant on both oral SfV *F* (1,28) = 6.259*, p* = 0.018, ES = 1.54, CI [0.7-2.35] and CC2 irregular word reading measure, *F* (1,28) = 6.476, *p* = 0.017, ES = 0.64, CI [-0.09-1.37] confirming post-test advantages for the SfV trial arm.

Correlation analysis was conducted to provide supplementary insight into the data to better understand the relationship between the reading measures and pre-and post-test, as well as the relationship between the N400 effect and both SfV and irregular word reading, using Bonferroni correction for multiple comparisons. At pre-test, RC was correlated with WIAT-WR (r = .693, p <.001), CC2 WR was correlated with CC2 IR (r = .817, p <.001) and NW (r = .800, p <.001), IR was correlated with NW (r = .796, p < .001), SfV was correlated with CC2-WR (r =.684, p <.001), IR (r = .622, p < .00), and NW (r = .598, p <.001). At post-test most of the reading measures were correlated with p < .001 (Table 5). The correlation between the N400 effect with SfV and irregular word reading at pre-test showed a correlation between N400 and SfV (r = 528, p = .004), and irregular word reading (r =.604, p <.001). At post-test, the N400 effect was correlated with SfV post (r = .534, p = .003).

Inferential analysis: N400

Behavioral measures

We ran repeated measures ANOVA with Group (CBP, SfV) as between-subject factor and time (pre-post), condition (congruent; incongruent) as within-subject factor. The results showed that there was a main effect of time F (1,26) = 8.47, p = .007, η_p^2 = .189, where both groups (CBP; SfV) improved in accuracy at post-test in the congruent condition. There was an interaction effect between time (pre; post) and Group (CBP; SfV) F (1, 26) = 6.07, p = .021, η_p^2 = .246, where the SfV group had better improvement scores in the incongruent condition and post-test (see Table 6).

EEG Analysis

N400 mean average amplitude A repeated measures ANOVA was conducted with Condition (congruent; incongruent), time (pre; post) as within-subject factors and Group (CBP; SfV) as between-subject factors. Results showed a significant interaction effect of time and condition, F

(1, 26) = 14.47, $p < .001$, $\eta_p^2 = .167$ where there was an increase in the N400 amplitude post the reading intervention. The results also showed that there was an interaction between Group x Condition, F (126) = 5.24, $p = .025$, $\eta_p^2 = .068$, and an interaction between Group x time x Condition, F (1, 26) = 63.59, $p < .001$, $\eta_p^2 = .469$, where the SfV group had an increase in the N400 amplitude in the incongruent condition at post-test.

N400 effect. The N400 effect (incongruent minus congruent) in Repeated Measures ANOVA, showed no main effect of time F (1, 26) = 1.97, p = .172, nor a main effect of groups F (1, 26) = 1.034, $p = .318$. There was, however, an interaction effect between Time X Group, F (1, 26) = 8.26, $p = .008$, $\eta_p^2 = .241$ where the N400 effect was more negative in amplitude in the SfV group.

N400 Fractional Peak Latency. Fractional peak latency is determining latency by identifying the peak amplitude and then tracing back in the waveform to the point where 50% of that peak voltage is reached. This method is more effective in identifying the onset latency, leading to more precise results (Luck, 2015; Neal et al., 2019). ANOVA was conducted to test hypotheses about significant differences between the two groups (SfV, CBP) in effects on the N400 post-test fractional peak latency in both conditions (congruent; incongruent). The main effect of the Group was significant in the incongruent condition F (1,32) = 31.50, $p < .001$, $\eta_p^2 = .496$ where the SfV group had an earlier latency in the incongruent condition at the post-test. The group had no main effect in the congruent condition F (1, 32) = 2.35, $p = .135$ (See Table 8).

Discussion

The stated aims of this study were twofold: 1) to explore whether an SfV- intervention that focused on teaching children with reading disabilities to manage grapheme-phoneme inconsistency in written English improves reading performance, and 2) to explore whether the

SfV reading intervention, having improved literacy outcomes, also changes the N400 amplitude and latency in children with reading disabilities. To answer these two questions, we randomized 30 participants with sustained reading difficulties in word reading into two groups: SfV and CBP. Both groups received ten weeks (12-14 hours) of online reading intervention, differing only in the presence of SfV strategy instruction (SfV arm) versus sight word reading instruction (CBP arm). Both interventions shared a focus on phonics, GPCs, and oral vocabulary. Theorized reading and related oral language behavioral outcomes were measured before and after the reading intervention using standardized and curriculum-based measures. We also measured ERP before and after both interventions and asked children to read semantically congruent and incongruent sentences at both times.

As expected, behavioral measures showed changes in the reading measures pre- and post-reading intervention. Both groups showed similar improvement in WIAT reading comprehension, word reading, vocabulary, and CTOPP phonological awareness tests. This may suggest that both interventions positively impacted reading, which is consistent with broader literature that show that phonics improves reading skills (e.g., Hulme et al. 2022). However, our study did not have a control group that did not receive phonics instruction to rule out the effects of maturation and classroom and other instruction. Hence, caution is required in the interpretation of these wider results.

Two distinct outcomes were theorized before intervention: that the SfV group should show significant changes in the Tunmer and Chapman (1998) oral SfV measure and the CC2 irregular word reading measure. We report significant changes in CC2 irregular word reading and in the oral SfV measure in the SfV intervention condition. This is consistent with previous studies that have suggested that SfV improves reading and SfV measures (Dyson et al., 2017;

Savage et al., 2018) in younger typical and at-risk poor readers in grades one and two and extends the pattern to somewhat older children aged around the grade four level, with sustained reading difficulties.

Our results also address a key unresolved question of regrading generalization in reading novel, untaught irregular words. Colenbrander et al. (2022) taught mispronunciation correction and found that children then showed an improved orthographic representation of trained words but not transfer words. In their study, Dyson et al. (2017) also did not report any changes in untaught CC2 exception word reading, suggesting inconsistent generalization effects. In the present study, the effect size analysis and inferential analyses showed the theorized differential impacts on oral SfV and exception word reading in the SfV intervention group. In both cases, there was strict separation between taught items and the items featured in transfer tests. The change in CC2 word reading in the SfV group but not the inferential analyses suggests that SfV may also generalize to CC2 word reading.

In Author et al., 2024, the RD group demonstrated a smaller N400 amplitude, and a delayed latency compared to a typically developing age-matched group. In this study, the same RD group, randomized into the SfV and CBP groups, showed improvement in the N400 amplitude with SfV. data, the results showed that the SfV group had an increased negative amplitude post-reading intervention in the incongruent condition, sentences with incongruent endings. These changes with intervention are consistent with Hasko et al. (2014), who also found that improvements to their reading intervention in German showed increased amplitude in the N400. Although we predicted a significant N400 latency effect, it was non-significant in our study. Nonetheless, the SfV group had an earlier latency (17*ms* earlier than the pre-test) in the incongruent condition, suggesting that teaching exception words can be generalized to earlier semantic processes of information. The effect size here was -0.95. for incongruent and .3 for congruent sentences in the SfV group, specifically, suggesting a practically important effect in the predicted direction. The absence of a statistically significant effect here may reflect the challenges children with sustained reading difficulties have in automating reading fluently (Shaywitz & Shaywitz, 2003). It may also be that more sustained SfV intervention is needed and/or a significant effect is evident at a delayed post-test.

An untheorized finding was the significant N400 behavioral improvement in accuracy for the SfV group. While not formally theorized as an outcome, if the SfV intervention promotes a generalizable strategy for reading irregular words, it should also affect post-test reading accuracy in the congruent and incongruent sentence tasks, especially if sentences contain untaught words with variant (irregular GPCs). Word-level features were not formally controlled in the N400 sentence tasks, but sentences did include many irregular words such as a *knob*, *pond*, *volcano*, and *dawn*.

We cautiously argue that our results suggest that teaching SfV can improve sustained word reading difficulties in a predominantly late elementary-aged poor reader sample, and training effects can be observed in brain activity as seen with the N400. These are above and beyond the known effects of phonics and vocabulary instruction shared in both intervention groups. This is the first study in the literature to undertake an SfV reading intervention and track neural changes. The improvement in the N400 amplitude concurs with the numerous MRI studies combined with reading intervention proposed, based on findings, an increase in brain activation in the left hemisphere reading regions post a reading intervention (Karipidis et al., 2018; Rezaie et al., 2011; Richards & Berninger, 2008; Shaywitz et al., 2004; Simos et al., 2007). These wider findings also suggest that a degree of recovery of brain activation through intervention is possible for poor readers.

Limitations

Like any study, this study has several limitations. Given our study's the small sample size, and restricted age group, the degree to which we can generalize these effects must be taken cautiously. However, our study was powerful enough to find medium-size effects that predicted behavioural and neurophysiological outcomes. Inspection of 95% confidence intervals, particularly for the oral SfV task and the N400 amplitude result, shows they did not include zero. The lower tail estimate here was d = .7 and d =.6, respectively, which adds to plausible replicability. The sample may also be somewhat untypical as, 76% of sample parents had a bachelor's degree. This study was primarily proposed to be run in school, but because of the COVID-19 pandemic, the study was changed to an online intervention to limit exposure. Hence, we tried to increase the number of participants but were limited for logistical reasons because of the pandemic. Further studies should replicate this study with a larger sample size in representative schools.

While children were matched on WIAT word reading on some measures after randomization, differences existed in some outcome measures at the pre-test. This was partly controlled by using MANCOVA to control for pre-test scores. Future studies should aim for better control in matching participants at pre-test prior to randomization to avoid pre-existing differences in outcome measures. Another potential limitation is that our sample included those with poor language skills and reading below grade level. Hence, we did not have a 'pure' research dyslexic population with poor word reading but typical verbal abilities. Nevertheless, we argue that the reading skills of our participants are consistent with the ‘real world’ context.

Our primary goal was to explore the effect of SfV on both reading and brain behavior. Hence, we controlled the reading intervention content to ensure that both groups received the same intervention in the absence of the SfV element for the CBP group to ensure that the hypothesized changes were due to SfV. However, there is arguably a need for additional untreated control or non-reading-related active control to understand further if the changes we saw in reading and neural profiles were due to the intervention and not due to the maturation effects, as we saw in phonological awareness improvement. Lastly, the duration of the intervention was no longer than 10-12 hours, limited again by resources. A longer intervention is recommended to see the effects on standardized outcomes like the WIAT reading achievement test. Longer intervention for such poor readers may be needed to show effects on all standardized outcomes and some N400 outcomes. such as latency.

Conclusion

In this study, we aimed to explore the effects of the SfV intervention, which focused on teaching children with RD to manage grapheme-phoneme inconsistency in written English. We sought to assess its impact on improving reading outcomes and on changes in the N400 related to reading improvement in 9-year-old children. This improvement was marked by an improved (increased) negative N400 amplitude in the incongruent condition, further suggesting that SfV intervention can improve neural measures of lexical-semantic processing. This is the first study to have used SfV intervention with poor reading participants and to study the relationship between SfV and

ERPs. Larger replication studies are needed to understand the relationship between SfV and ERP components.

References

Author et al., (2023). The N400 in readers with dyslexia: A systematic review and meta-analysis. *International Journal of Psychophysiology*, 112283. https://doi.org/10.1016/j.ijpsycho.2023.112283

Boldrini, G., Fox, A., & Savage, R. S. (2022). Flexible Phonics: A complementary 'next generation' approach for teaching early reading. *Literacy, 57, 1*, 72-86. https://doi.org/10.1111/lit.12308

Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improve children's ability to learn to read words. *Scientific Studies of Reading*, *21*(5), 392-407. https://doi.org/10.1080/10888438.2017.1315424

Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly*, *55*, S45-S60. https://doi.org/10.1002/rrq.334

Górecka, J., & Makiewicz, P. (2019). The dependence of electrode impedance on the number of performed EEG examinations. *Sensors*, *19*(11), 2608. https://doi.org/10.3390/s19112608

Gupta, R., Mitra, M., & Bera, J. (2014). *ECG acquisition and automated remote processing* (Vol. 20). New Delhi: Springer India. Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2014). What does the brain of children with developmental dyslexia tell us about reading improvement? ERP evidence from an intervention study. *Frontiers in Human Neuroscience*, *8*, 441, 1-20 https://doi.org/10.3389/fnhum.2014.00441

Hulme, C., & Snowling, M. J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, *20*(3), 139-142. https://doi.org/10.1177/0963721411408673

Kallner, A. (2017). *Laboratory statistics: methods in chemistry and health sciences*. Elsevier

Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., Walitza, S., & Brem, S. (2018). Simulating reading acquisition: The link between reading outcome and multimodal brain signatures of letter–speech sound learning in prereaders. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-24909-8

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and language*, *102*(1), 30-45. https://doi.org/10.1016/j.bandl.2006.11.001

Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *The Lancet*, *357*(9263), 1191-1194. https://doi.org/10.1016/S0140-6736(00)04337-3

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). CONSORT 2010 Explanation and elaboration: Updated guidelines for reporting parallel group randomized trials. *British Medical Journal, 340*, c869. https://doi.org/10.1016/j.ijsu.2011.10.001

Morley, A., Hill, L., & Kaditis, A. G. (2016). 10-20 system EEG Placement. *European Respiratory Society, European Respiratory Society*.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skills. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). American Psychological Association. https://doi.org/10.1037/10459-004

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In *Advances in Psychology* (94, pp. 227-248). North-Holland. https://doi.org/10.1016/S0166-4115(08)62798-3

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A computational model of the self-teaching hypothesis based on the dual-route cascaded model of reading. *Cognitive Science*, *42*(3), 722-770. https://doi.org/10.1111/cogs.12571

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

Richards, T. L., & Berninger, V. W. (2008). Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. *Journal of Neurolinguistics*, *21*(4), 294–304. https://doi.org/10.1016/j.jneuroling.2007.07.002Savage, R., Georgiou, G., Parrila, R., & Maiorino, K. (2018). Preventative reading interventions teaching direct mapping of graphemes in texts and set-for-variability aid at-risk learners. *Scientific Studies of Reading*, *22*(3), 225-247. https://doi.org/10.1080/10888438.2018.1427753

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology 94*, 2, 143–174. https://doi.org/10.1348/000712603321661859

Share D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–226. https://doi.org/10.1016/0010-0277(94)00645-2

Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. In *Advances in child development and behavior* (Vol. 36, pp. 31-82). *Journal of Advances in Child Development and Behavior*. https://doi.org/10.1016/S0065-2407(08)00002-5

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95-129. https://doi.org/10.1006/jecp.1998.2481

Share, D. L. (2011). On the role of phonology in reading acquisition: The self-teaching hypothesis. In S. A. Brady, D. Braze, & C. A. Fowler (Eds.), *Explaining individual differences in reading: Theory and evidence* (pp. 45–68). Psychology Press.

Shaywitz, S. E., & Shaywitz, B. A. (2003). Dyslexia (specific reading disability). *Pediatrics in Review*, *24*(5), 147-153.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., Holahan, J. M., Marchione, K. E., Fletcher, J. M., Lyon, G. R., & Gore, J. C. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention. *Biological Psychiatry*, *55*(9), 926–933. https://doi.org/10.1016/j.biopsych.2003.12.019

Steacy, L. M., Wade-Woolley, L., Rueckl, J. G., Pugh, K. R., Elliott, J. D., & Compton, D. L. (2019). The role of set for variability in irregular word reading: Word and child predictors in typically developing readers and students at-risk for reading disabilities. *Scientific Studies of Reading*, *23*(6), 523-532. https://doi.org/10.1080/10888438.2019.1620749

Tunmer, W. E., & Chapman, J. W. (1998). Language prediction skills, phonological recoding ability, and beginning reading. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling: development and disorders* (pp. 33–68). Hillsdale, NJ: Erlbaum.

Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, *16*(2), 122-140. https://doi.org/10.1080/10888438.2010.542527

Venezky, R. (1999). *The American way of spelling*. New York: Guilford Press

Zipke, M. (2016). The importance of flexibility of pronunciation in learning to decode: A training study in set for variability. *First Language*, *36*(1), 71-86.

Acknowledgements

We would like to express our sincere gratitude to Dr. Dave Saint-Amour for generously providing us with EEG equipment used in the study. We extend our gratitude to the school boards, the families, and the participants for their collaboration, time, and effort in completing this study. We would like to express our gratitude to every single parent who continuously encouraged their children to complete the 10-week online reading intervention. Special thanks to all our undergraduate and graduate research trainees who delivered the intervention with much professionalism.

Funding

This research was partially funded by a William Dawson Research Scholar Award to Dr. Armando Bertone and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC; RGPIN-2020-05052) to Dr. Gigi Luk.

Table 3.1

Demographic table of the study population

	SfV (*n* = 15)	CBP (*n* = 15)
Age (years) pre	9.5	8.8
Sex (male/female)	7/8	6/9
Dominant language at home	English	English
Mothers Education	Bachelor and above	Bachelor and above
Intervention hours	13.26	13.6

Table 3.2

Mean and SD of standard reading scores for both groups pre- and post-intervention.

	SfV (*n* = 15)			CBP (*n* = 15)		
	Pre	Post	Effect size	Pre	Post	Effect size
WIAT RC	72.82(16.548)	85.47 (20.329)	0.6 CI [-0.05 – 1.41]	70.80 (19.91)	81.73 (18.66)	0.5 CI [-0.16-1.29]
WIATWR	70.33(12.653)	73.73 (12.931)	0.26 CI [-0.45-0.98]	73.53(13.469)	76.33 (17.795)	0.19 CI [-0.53-0.89]
WIAT Voca	79.60(18.558)	88.07(20.614)	0.41 CI-0.29-1.15]	85.73(11.634)	90.40(11.915)	0.39 CI [-0.32-1.11]
Elision	6.6 (2.197)	7.6(2.028)	0.5 CI [-0.251-1.2]	6.93(2.052)	7.27 (2.492)	0.13 CI [-0.56-0.89]
BW	4.33(2.127)	7.73(3.575)	1.1 CI [0.38- 1.9]	5.00 (1.732)	6.87 (2.386)	0.9 CI [0.14-1.64]
PI	7.93(3.77)	9.13(3.067)	0.34 CI [-0.3-1.07]	7.29 (2.785)	9.07(3.54)	0.55 CI-0.17-1.28]
PA	79.07(14.839)	88.87 (13.989)	0.67 CI [-0.056-1.41]	77.93(7.411)	86.53 (13.789)	0.62 CI [-0.33-1.51]
CC2WR	10.53(10.295)	19.87(10.875)	0.88 CI [0.13-1.63]	19.27(11.732)	17.2(12.963)	-0.15 CI-0.88-0.54]
CC2 IW	7.93 (8.972)	13.27 (7.63)	0.64 CI [-0.09-1.37]	11.07(8.084)	9.87(8.288)	-0.2 CI-0.92-0.51]
CC2 NW	8.6 (10.548)	9.87 (8.999)	0.12 CI [-0.58-0.84]	11.8(9.481)	8.13(8.114)	-.04 CI [1.13—1.10]
SfV	10.4 (3.397)	15.275 (2.915)	1.54 CI [0.72-2.35]	11.2(3.821)	12.53(3.067)	0.3 CI [-0.33-1.10]

Note. WIAT RC= reading comprehension, WIAT WR = WIAT word reading, WIAT Voca = WIAT expressive vocabulary, BW= blending words, PI = phoneme isolation, PA = phonological awareness, CC2WR= Castle and Coltheart word reading, CC2 IW = Castle and Coltheart irregular word reading, CC2 NW = Castle and Coltheart nonword reading, SfV= Set for Variability.

Table 3.3

Correlation of reading measures at pre- and post-test

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

1.RC Pre	1															
2.WR Pre	.693**	1														
3.Voca Pre	0.2	0.3	1													
4.PA pre	0.2	0.3	.494**	1												
5.WR-2 Pre	.451*	.453*	0.4	0.2	1											
6.IR Pre	.443*	.371*	0.3	0.1	.817**	1										
7.NW pre	.434*	.443*	0.3	0	.800**	.796**	1									
8. SfV pre	.434*	0.4	0.3	-0	.684**	.622**	.598**	1								
9.RC Post	.787**	.694**	.380*	0.3	0.3	0.3	.387*	0.3	1							
10.WR post	.597**	.844**	0.1	0.4	.400*	0.3	.424*	0.4	.594**	1						
11.Voca Post	0.1	0.1	.765**	.465**	0.2	0.1	0.2	0.2	0.3	0	1					
12.PA post	.413*	.429*	0.2	.438*	0.2	0	0.2	0.2	.480**	.490**	.386*	1				
13.WR-2- Post	.610**	.753**	0.1	0.1	0.3	0.2	0.3	0.3	.724**	.731**	0.1	.518**	1			
14.IR Post	.748**	.706**	0.2	0.1	.410*	.482**	.468**	.458*	.789**	.644**	0.1	.415*	.857**	1		
15.NW post	.512**	.661**	0.1	0.2	0.2	0.2	.395*	0.3	.697**	.713**	0.1	.511**	.853**	.677**	1	
16.SfV post	.541**	.436*	0.2	0.1	0.1	0.1	0.2	0.3	.732**	.416*	.367*	.554**	.737**	.760**	.653**	1

**. Correlation is significant at the 0.001 level (2-tailed) with Bonferroni correction.

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

Table 3.4

Correlation of SfV, irregular word reading and N400 effect.

	1	2	3	4	5	6
1.SfV pre	1					
2.IR pre	.622***	1				
3.N400 pre	.528**	.604**	1			
4.SfVpost	.336	.147	.147	1		
5.IR post	.458*	.482**	.482**	.760***	1	
6. N400 post	.326	-.029	-.132	.534**	.325	1

p = .008 (Bonferroni correction for multiple comparison)
*** Correlation is significant at .001 level
**Correlation is significant at the < .01 level
*Correlation is only significant at < .05 level

Table 3.5

Mean, SD, effect Size, and Confidence interval of the reaction time pre- and post-

	SfV(*n*=15)				CBP(*n*=15)			
	pre	post	ES	CI	pre	post	ES	CI
Congruent	1199 (315)	1052 (246)	0.5	-0.20-1.24	1165 (264)	1205 (214)	-0.16	-0.88-0.55
Incongruent	1213 (319)	1127 (271)	0.3	-0.42-1.01	1236(346)	1186 (216)	0.17	-0.5-0.8

Table 3.6

Mean, SD, effect Size and Confidence interval of the accuracy rate % pre- and post-

	SfV (*n*=15)				CBP (*n*=15)			
	pre	post	ES	CI	pre	post	ES	CI
Congruent	40 (23.84)	68 (21.16)	1.24	0.46-2.02	42.43 (18.55)	51 (17.62)	0.4	-0.22-1.22
Incongruent	36 (22.55)	63 (29.580)	1.02	0.26-1.78	48.14 (16.87	50 (26.53)	0.09	0.62-0.8

Table 3.7

Mean, SD, effect size, and confidence interval of the N400 amplitude pre-post-

	SfV (*n* = 15)				CBP (*n* = 15)			
	pre	post	ES	CI	pre	post	ES	CI
Congruent	-5.667 (2.001)	-4.228 (1.413)	0.8	0.08-1.57	-3.632 (1.173)	-4.422 (1.193)	-0.66	1.42-0.05
Incongruent	-3.609 (.9301)	-5.590 (1.723)	1.14	0.62-2.23	-5.339 (1.49)	-4.934 (1.368)	0.28	-0.436-1002

Table 3.8

Mean, SD, effect size, and confidence interval of N400 fractional peak latency pre-post-.

	SfV (*n* = 15)				CBP (*n* = 15)			
	pre	post	ES	CI	pre	post	ES	CI
Congruent	327(31.84)	296 (151.8)	-0.28	-0.92-0.35	315 (133.35)	323 (73.52)	0.07	-0.56 - 0.71
Incongruent	353 (101.80)	330 (62.29)	-0.59	-1.24- -0.05	355 (53.64)	376 (40.18)	0.44	-0.2-1.08

Figure 3.1

N400 amplitude and latency of both groups in the incongruent condition post intervention.

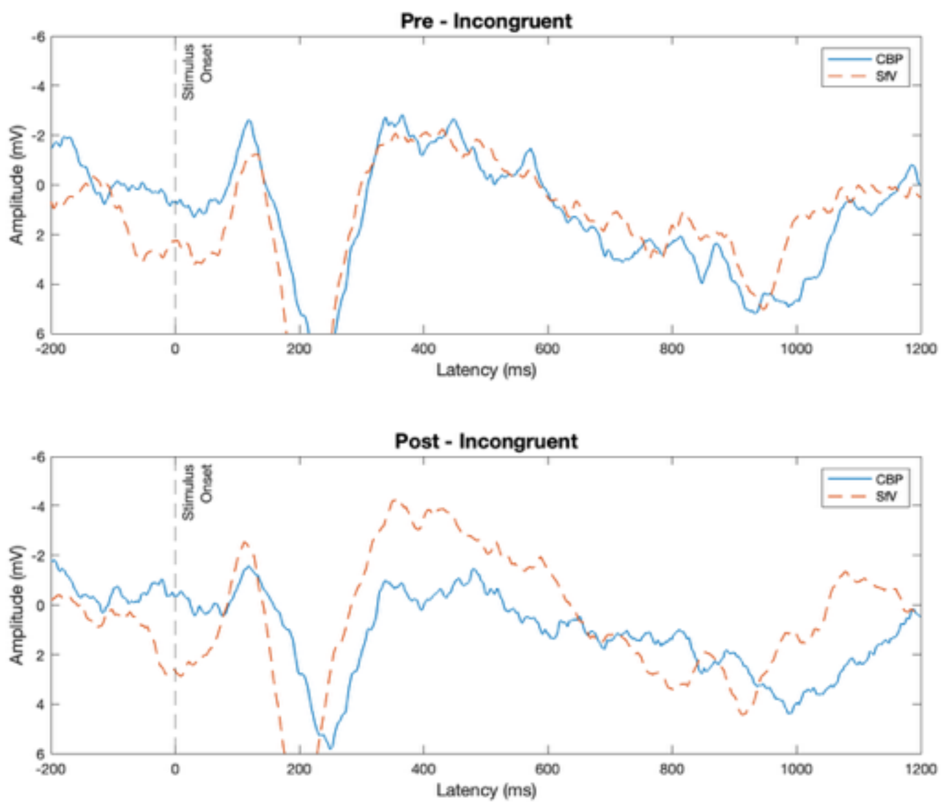
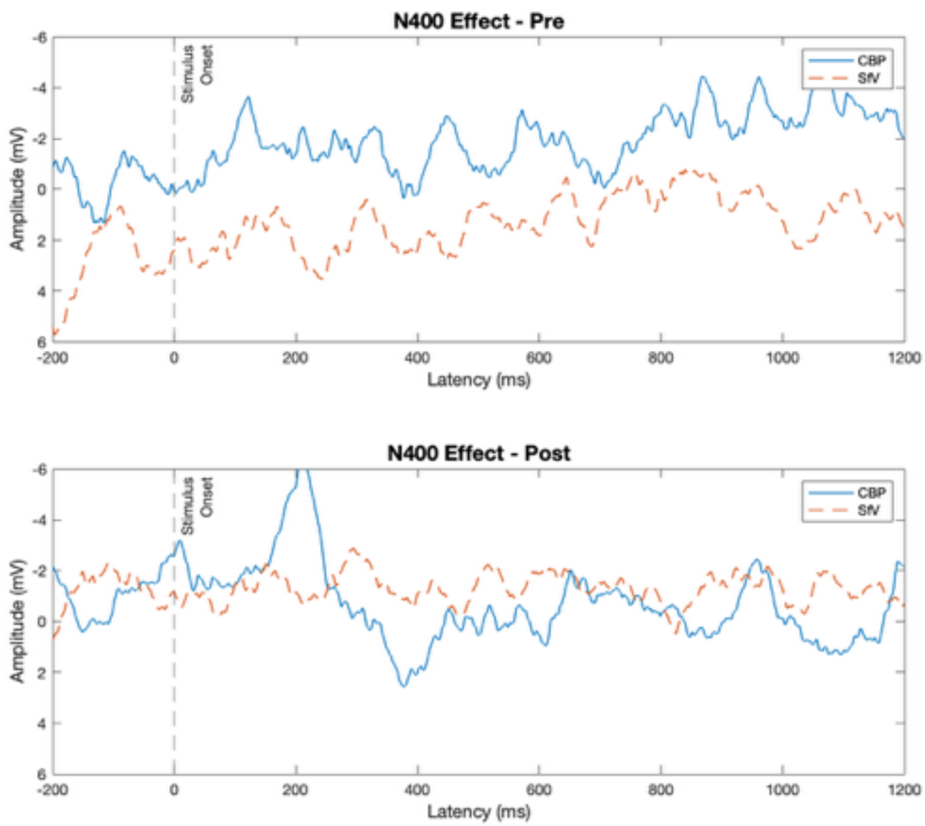


Figure 3.2

N400 effect pre and post in both groups (SfV; CBP).



Appendix 1

The 5-step model used to teach SfV.

A. Children blend a letter string looking for and applying phonics rules (grapheme-phoneme correspondence).

B. Ask the participant the following*: What do we first do when we read a word?* Answer: *We use our phonic rules and sound it out*. Example. *The vowel is short in bat, sit, and rug.*

C. With new words, and exception words, the participant evaluates his/her attempt to synthesize pronunciation by first applying the phonics rules, then ask the participant: *What do we do next children?* answer: *I ask myself – is this a word I know?* (Vocabulary and meaning)

D. If the child's answer is "no" Then ask the child to try and replace the consonant or a vowel with another consonant or vowel pronunciation that they might know (comment variants of letter names and "uh" sounds. (I would supply some of these rules as "words banks" at the beginning".

What do we do then? Answer: *we flip/swap in another sound we know for the letters*.

E. The participant then synthesizes the revised phoneme string. Lead the children by asking "*What do we do next*" Answer: *Then, we use our other phonic rules and sound them out*

Chapter 9. Discussion

The three studies presented in this thesis each contributed to a central thesis argument that (i) the N400 measures a lexical semantic aspect of language tasks, (ii) the SfV measures similarly lexical -semantic aspects of language, (iii) the N400 indexes SfV and (iv) reading intervention involving instruction in SfV improve both reading and N400 amplitude. These were evidenced by a reliable N400 profile difference between TD and RD children in the literature (Manuscript 1) and in an empirical study (Manuscript 2). Furthermore, reader group differences in SfV and a correlation of aspects of N400 with SfV reported in Manuscript 2 underscore the rationale for an SfV intervention along with an N400 ERP task in a randomized control study (Manuscript 3). The intervention was created by teaching children the ability to flexibly apply a variety of candidate GPCs in deriving word pronunciations. This is derived from theories of SfV and closely related reading theories (Tunmer & Chapman, 2012; Venezky, 1999). The resultant intervention targeted behavioral components of reading. The efficacy of the intervention was then examined by evaluating its impact on both reading measures (especially exception word reading), taught language abilities (SfV) and neurophysiological changes (N400 amplitude and latency) of students with reading difficulties. To do so, the SfV intervention was compared to a standard CBP intervention. Results of significant post-test advantages for the SfV group on irregular word reading, SfV, and N400 amplitude directly support claims they are linked. The research presented here aimed to address several gaps in the literature on reading instruction, including the need for integration between theoretical and empirical perspectives on neurophysiological correlates of reading difficulties and the need for a well-designed intervention that assesses the potential of combined behavioral and neurophysiological tracking of reading interventions.

In Chapter 4, results of the systematic review and meta-analysis sought to identify the difference in mean N400 amplitude and latency in TD and RD and to explore the impacts of moderators. The results revealed a substantial effect size of 0.66, indicating a significant difference in the mean of the N400 amplitude between TD and RD. This observed substantial difference between semantically congruent and incongruent task (g = 0.948) among other moderator factors discussed in Chapter 2, suggests that the N400 serves, to a considerable extent, as an indicator of semantic processes. The consistency of the N400 amplitude mean difference provided robust neurophysiological evidence that RD individuals have lexical semantic processing difficulties. This neurophysiological evidence not only enhanced our understanding of the neurocognitive processes involved in reading difficulties but also has potentially important implications for pedagogy. Recognizing the challenges in semantic processing highlighted by the N400 amplitude differences underscores the need for principled and effective reading interventions tailored to address these specific needs in individuals with RD. Additionally, the need for more data on the N400 latency was identified. Existing research (Helenius et al., 2009; Jednoróg et al., 2010; Rüsseler et al., 2007) suggests a delayed N400 latency among individuals with RD. While the primary objective in this thesis was to examine the specific effect of SfV on reading improvement and neurophysiological change in young children with RD. We first sought to address a number of preliminary questions: Firstly, do SfV delays exist in age- and demographic-matched RD versus TD groups? Secondly, do samples of younger children below the age of 10 show N400 sentence incongruency effects? Finally, we asked - what is the correlation between the N400 amplitude and latency and reading-related measures, and most specifically, the SfV measure. Additionally, there was a need to ascertain whether the correlation differs between the congruent and incongruent conditions, a question for which we did not find an answer in the systematic review and meta-analysis.

Chapter 6, an empirical study designed to investigate SfV delays in age- and demographic-matched RD versus TD groups, confirming for the first time in such samples, a delay in SfV use in the RD sample. The study also explored differences in N400 amplitude and latency between the TD and RD. The results showed that young TD children had a more enhanced negative N400 amplitude and a faster latency in the incongruent condition over RD children. An N400 effect remained absent in young children with RD, which indicates a deficiency in lexical-semantic processing. As anticipated, distinct processes of words and sentence reading and N400 emerged between the TD and RD. N400 latency was correlated with SfV in the RD group. Therefore, we established the imperative to extend these observed patterns and employ SfV intervention, in conjunction with the anticipated improvement in reading among young children with RD, to investigate potential intervention-driven plasticity in the underlying neurophysiological N400 components.

In Chapter 8, we investigated whether SfV intervention could enhance reading performance and secondly, we explored whether the SfV reading intervention, upon improving literacy outcomes, would also induce changes in the N400 amplitude and latency in children with RD. Thirty participants with sustained RDs in word reading were randomly assigned to two groups: SfV and CBP. Both groups underwent 10 weeks (12-14 hours) of online reading intervention, differing only in the inclusion of SfV strategy instruction (SfV arm) or sight word reading instruction (CBP arm). We observed improvement in both groups post the reading intervention. The SfV group improved in both in SfV measure and irregular word reading in contrast to the CBP control group. This suggests that the SfV intervention improves theorized aspects of reading and reading-related SfV measures that likely reflect the management of GPC inconsistency in young children with RD. The SfV group also exhibited an increased negative N400 amplitude in the incongruent condition after the intervention—indicating that it underpins the specific observed behavioural change in reading. It is posited that our findings indicate that instruction in SfV has the potential to ameliorate sustained word reading difficulties in a primarily late elementary-aged poor word reader sample. Additionally, apparent training effects are observable in brain activity, as evidenced by changes in the N400.

Implications for Theory

The Role of the N400 in Lexical-Semantic Processing

In Chapter 4, the meta-analysis and systematic review study presented a reliable N400 difference between TD and RD. The results in this study show various moderators were identified that played a role in the N400 effect difference between the two populations. This suggests that the N400 is a reliable index of lexical-semantic processing. In Chapter 6, Study 2 revealed an N400 effect for incongruent sentences in the TD group. In addition, it showed a correlation between the N400 amplitude and word reading. A trend in the correlation tables was observed, even when not statistically significant, with word reading measures and not with reading comprehension measures. This suggests that the N400 is a reliable index of word reading. Therefore, the question that arises is whether the N400 just a facilitator for meaning processing? An index of purely lexical (orthographic/phonological) processes? (and perhaps that is why we saw a correlation with word reading), or is it an index of an interactive process of the cascaded effect of word reading and thus a weaker connection with semantic processing?

Challenges in performing lexical decision tasks may indicate difficulties in the fundamental process of decoding words (Coch & Holcomb, 2003). From this perspective, a fundamental difficulty in phonological awareness contributes to obstacles in GPC processing, which, in turn, may be connected to delayed semantic processing. Thus, difficulty in phonological stimulus tasks is sometimes associated with a reduced N400 amplitude (McPherson et al., 1998,1999). Another potential explanation for the mechanism underlying the N400 involves a relatively late, post-word-level integrative process. In many psycholinguistic models, integrative processes are employed by language comprehension systems to integrate information from various lower-level processes (such as lexical/semantic, syntactic, and pragmatic) into an ongoing discourse representation (Coch & Benoit, 2015; Landi & Perfetti, 2007). This interpretation aligns with the findings from our two current studies. According to this characterization, N400 amplitudes for words at the end of sentences indicate the ease with which integrative processes can assimilate information. Additionally, Coch and Holcomb (2003) propose that the N400 does not exclusively represent phonological or semantic processing. Instead, it reflects a dynamic and interactive system supporting the processing of meaning, involving sequential and interactive information of word-level processes (e.g., decoding, sight word recognition), followed by the activation and selection of meaning in the mental lexicon, ultimately leading to comprehension. On the other hand, sentence endings that are more congruent and expected, result in smaller N400 responses than incongruent ones (Kutas & Hillyard, 1980a, 1980b, 1984). This sentence-processing data can in principle be explained through a more contextual lexical interpretation. For instance, when the final word of a sentence is highly anticipated, its meaning is likely activated in advance (Deacon et al., 2012). While presenting the words constituting a sentence, the anticipation of the final word likely initiates top-down activation of semantic information. These top-down semantic influences may facilitate the orthographic or phonological processing essential for extracting meaning from the word in a connected word reading system such as a Triangle model (Plaut et al., 1994), though it should be noted that such triangle models do not implement text-level processing influences on words.

From a purely neurophysiological perspective, Kutas and Federmeier (2011) argue in the review of 30 years of research on the N400 that it is challenging to suggest a precise neurophysiological characterization of the N400 and indicate it as an index of a specific mental process because it depends on factors like response behavior. experiment. This could explain the modulation of the N400 by various moderators that we saw in the Chapter 3; it probably arises from the activity of multiple neural generators (Lau et al., 2008), The N400 could rather be an interactive system that supports meaning processes based on what was evident in all three studies. However, further replication studies with larger sample sizes with sentence incongruency tasks are needed to explore the relationships between reading measures and various aspects of the N400 to gain a better understanding of the N400 mechanism.

The N400 and SfV

This thesis explored the potential connections between SfV and irregular word reading on the one hand and N400 measures on the other to test the central thesis about SfV and N400 both indexing lexical-semantic processes at behavioural and neurophysiological levels respectively. As for the behavioural difference between children with different reading abilities, the results of Study 2 in Chapter 4 have indicated a difference in the SfV between TD and RD. The RD group performed lower than the TD on the SfV measure despite multiple controls for candidate extraneous variables. The RD group was less accurate and slower in RT in the N400 behavioral task compared to the TD group. The correlation analysis showed that the N400 latency was correlated with SfV in the RD group. In Study 3 (Chapter 8), we found that the SfV-taught group had an increased N400 amplitude and better reading performance in SfV and irregular word reading. A broad and potentially causal association between N400 and SfV is suggested if intervention effects are replicated in larger samples in the future.

It may be possible to explore which aspects of N400 are associated with the learning process as children acquire words. In the correlation study, a significant effect was observed between N400 latency and SfV in the RD group. Following a 12-hour intervention focused on SfV, the post-intervention assessment revealed a more negative N400 amplitude. Why did we observe a correlation with N400 latency in the correlation study, while witnessing enhancement in N400 amplitude following the intervention in the intervention study? Explanations as to why this pattern of results was observed are naturally speculative at this stage. It could however be theorised first, that the correlation with latency in Study 2 suggests that the RD readers had more challenges in the cognitive processing of the lexical-semantic aspects of words. From this view, N400 latency indexes fluency in accessing the limited number of already stored word pronunciations children with RD have, which is thus evident to a greater degree among the stronger readers of the RD group.

In Studies 2 and 3 (Chapters 6, 8), a model of direct instruction in SfV strategy is theorised to have a direct effect on improving irregular word reading and then an indirect effect on reading comprehension (Tunmer & Chapman, 2012). There was also evidence of generalization here to novel irregular words. Hence, one plausible explanation is the SfV intervention with its focus on explicit strategy instruction that improved their irregular word reading, might be that it helped many in the RD overcome their prior difficulty at the cognitive processing level by giving an explicit meta-cognitive tool that allows wider self-teaching for irregular words and thus of sentences, and which is then reflected in an enhanced amplitude post-intervention.

From a learning theory standpoint, there may exist an inherent and untaught proficiency in SfV among better readers who might simply explicitly 'work it out' for themselves and thus be indexed by pre-intervention N400 latency-SfV associations. In the context of a connectionist model, proficient readers might establish interconnected processes involving orthographic, phonological, and semantic aspects when encountering words that allow acquisition to occur naturally through experience (e.g., Plaut et al., 1994; Ziegler et al 2020). It is thus even possible that for typical and stronger readers among those with RDs, the use of SfV is an implicit learning process. Poor readers with less powerful connections among orthography, phonology and semantics in their neural networks cannot so benefit (Ziegler et al., 2020). By contrast, a purposive explicitly taught generative strategy for children with RD provides a powerful meta-cognitive tool for these children to overcome computational limitations they otherwise experience. The SfV meta-cognitive instruction produces greater neural unity and orchestration at N400 in taught children and is thus indexed by distinct amplitude effects at intervention post-test for the first time. From this view, meta-cognitive strategy instruction creates connected neural processing capacity that is then brought to bear to read more successfully. Children can then benefit from printed word exposure towards acquisition. It is also plausible that the consequent better comprehension abilities exhibited by more proficient readers within the RD group contributed to their capacity to process sentence meaning, in contrast to less skilled readers, thereby resulting in the observed amplitude effects post-intervention.

Several future studies would perhaps answer some of these questions. First and foremost, a study with a larger sample size is needed to see if similar results will take place and establish that these results are robust. Perhaps a replicated study with a delayed post-test to see if we have a consistent result both behaviorally and neurophysiologically, is the first step. A future study could also adopt a different study framework with an SfV intervention and the N400 latency and amplitude being measured at multiple time points as SfV use demonstrably moves from acquisition through to adoption to mastery and generalization phases. This might give a better and more dynamic understanding of the reasons for the shift between latency and amplitude over time and with learning over the pre-post-test design. An additional prospective study could perhaps involve conducting a cross-sectional study encompassing diverse levels of competence in SfV across distinct RD groups. This study would perhaps aim to examine disparities in N400 amplitude and latency before and after the implementation of an SfV intervention.

Implications for Research and Teaching Practice

It was observed in the thesis that young good readers frequently come across unfamiliar letter combinations. In such cases, neither guessing from context nor direct instruction on these words is likely to be effective in producing generative approaches. This assumption of centrality of generative power forms the basis of the concept of self-teaching, which suggests that the ability to independently derive the pronunciation of a new word, based on sub-lexical connections between spelling and sound, is the most effective way to acquire the orthographic knowledge necessary for skilled word recognition. The majority of our understanding of written language is learned independently (Share 1995, 1999). It's possible that children also teach themselves how to read words with irregular spelling patterns as well. They might do this by using their increasing knowledge of how letters and sounds correspond to figure out irregular spellings and by trying out different ways of saying the word until they find one that matches a word they know and makes sense in the sentence. This process of figuring out words with variable spellings, SfV, was discussed earlier in the introduction and multiple discussion points in studies 2 and 3. Dyson et al. (2017) ran a 4-week SfV with TD readers. After the 4-week intervention, the children in their study improved in reading irregular taught words and irregular untaught words. Dyson's study, however, did not find a significant effect on Castle and Coltheart's irregular word list. In Chapter 5, we found improvement in the SfV measure, and we were able to see a generalization effect to CC2 irregular word reading. One plausible explanation of why effects were evident here and not in Dyson et al. is the effect of the intervention hours. The children in Dyson's study only received 2.6 hours of intervention over 4 weeks while our participants received a minimum of 12 hours over 10 weeks.

In addition, we used CBP practices as an active control intervention. The CBP included the same procedure of SfV intervention, in the absence of the SfV component. The CBP group showed improvement in all reading measures between the pre- and post-test. However, the CBP appears insufficient for generalizing to irregular word reading, as we did not find significant results in the CC2 irregular word list. In this context, (SfV) assumes significance, serving as a complementary extension to phonics instruction. I believe this intervention study shows evidence that there is a direct connection between a child's ability to fix mispronounced irregular words (via SfV) and their ability to learn to read specifically in an RD group. These findings indicate that interventions incorporating "lexicalized" phonic strategies for teaching variable vowel rules, prove to be more effective compared to common best-practice interventions utilizing widely validated research practices. In this context, SfV assumes significance, serving as a complementary extension to phonics instruction.

The correlations observed between neural and behavioral measures have the potential to offer valuable insights to educators. Specifically, these findings may guide instructional decisions related to word reading for students with RD, emphasizing tailored strategies to address their specific lexical-semantic challenges. Concurrently, for TD students, the focus may shift towards implementing more advanced reading comprehension strategies. This approach aligns with the principle of differentiation in the classroom, wherein instructional methods are tailored to meet the individual learning needs of students with distinct cognitive profiles (Tomlinson & Imbeau, 2023). In addition, the details provided on the N400 amplitude and latency offered us potential insights into the lexical-semantic processes areas of process deficits in reading (Osterhout & Holcomb, 1995). Consequently, it enabled us to design the SfV reading intervention that effectively targeted the lexical-semantic processing difficulties that RD children face. Thereby, the N400 component of our thesis aided in the potential interventions of reading difficulties. Improvement was observed within the SfV group, accompanied by enhanced generalization to untaught irregular words and a more negative amplitude of the N400. Nonetheless, it is crucial to note that not all participants exhibited improved reading scores after the intervention. The intervention showed that it can improve reading online. This suggests that it can be implemented online and reduce the time needed to be implemented in person during school hours. While the intervention demonstrates efficacy and is recommended for educators and parents, it should be emphasized that this intervention serves as a valuable strategy to aid readers with RD rather than a definitive cure for reading difficulties.

Limitations

The work presented in his thesis was subject to certain limitations. As discussed in the correlation and intervention study, the recruitment experienced impediments due to the impact of the COVID-19 pandemic. From March 2020 to September 2021, access to the university campus, including our laboratory, was restricted. These restrictions significantly impacted the timeline for completing and submitting my thesis. As a result, the intended sample size was reduced due to these circumstances. The intended intervention study was due to occur within school settings featuring a larger sample of clusters of small groups of students. However, due to the constraints imposed by government restrictions related to the COVID-19 pandemic, the implementation of the initial plan was impeded. Hence, this compelled me to reconsider and reframe the study as an online reading intervention. I initially wanted to look for clusters of students who have an official diagnosis of reading disability. This proved quite challenging to find. The RD and TD reader groups were matched on WIAT word reading. However, pre-existing literacy differences between the groups on other measures remained. In study 3, we only had a small sample size of 30 participants in the RD group. Despite this modest sample size theorized significant effects were still evident while using conservative analysis such as rigorous study selection for meta-analysis with mediation effects (Manuscript #1), ANCOVA (Manuscripts #2 and #3), and two-tail testing of closely theorized directional correlation effects (Manuscript #2). Particularly in Manuscript #2, the TD group had a WR mean average =103. This might have affected the results that we had in both the N400 and reading measures. Perhaps future studies would include more representative average readers to compare with RD groups. Additionally, we are cautious about overwhelming the participants (specifically the RD groups) with the reading and EEG measures, such that we did not include a measure of cognitive ability such as verbal and non-verbal reasoning. Perhaps future studies could focus on word reading, irregular word reading, and a measure of non-verbal intelligence to have a better understanding of the participants' cognitive abilities. In addition, within the context of Montreal, a multicultural city, it is noteworthy that none of our participants were monolingual. All individuals involved in both studies predominantly spoke both English and French and, in some instances, a third language. This limits our generalization as our sample may not be representative of a wider Canadian sample.

Conclusion

This dissertation enhances the existing body of literature concerning the correlation among neurophysiological processes, reading and SfV. Additionally, it contributes to our understanding of lexical-semantic processes from a neurophysiological perspective (N400) and a behavioral (SfV) perspective by exploring and understanding the neurophysiological changes that occur after a successful intervention (SfV) aimed at improving reading irregular word reading. To date, Manuscript #1, the meta-analysis and systematic review is the first attempt in the literature to investigate the distinctions of the N400 mean and amplitude effect between TD and RD individuals, while also scrutinizing the moderators influencing this effect. Manuscript #2 is the first research study directly examining the correlation between reading measures, specifically SfV and the N400 in two distinct groups (TD; RD). This thesis represented an initial investigation reporting SfV delay in RD over matched TD groups, and even where expressive vocabulary is controlled statistically. Employing a sentence incongruity ERP paradigm, the study delves into N400 patterns and SfV in children, both with and without RD. Furthermore, no studies have directly evaluated the efficacy of SfV as a reading intervention within samples experiencing reading difficulties (RD), and hence Manuscript #3. Notably, it is the first study to document N400-linked effects in congruent versus incongruent printed sentence tasks in English among children under the age of 10, who are still in the process of acquiring written English skills, and the first intervention study to use SfV intervention and track neurophysiological changes. Together these studies go some distance towards establishing the main argument advanced at the start of the thesis that the N400 and SfV both reliably measure lexical semantic aspects of language, with the N400 operating as a neurophysiological index of SfV and such that a reading intervention involving instruction in SfV improves both reading and N400.

Several potential avenues for future research exist beyond those already identified, and their pursuit could significantly contribute to a more comprehensive understanding of the N400 in individuals with RD. These prospective directions may encompass, but are not limited to, the following: Conducting longitudinal studies that track individuals with RD over an extended period including phases of sustained reading intervention could provide insights into the developmental trajectory of N400 patterns. Examining how N400 responses evolve may reveal critical information about the persistence or modification of neurocognitive processes associated with reading difficulties. Additionally, expanding the scope of research to include diverse subtypes of RD such as at-risk readers, and readers with poor word reading but good reading comprehension and considering variations in severity and etiology. It is interesting to see how these sub-types of RD could elucidate whether N400 patterns differ across these subgroups. This approach may contribute to a more nuanced understanding of the neural correlates of RD.

Another avenue for future work is integrating multiple neuroimaging modalities, such as functional magnetic resonance imaging (fMRI) or magnetoencephalography (MEG), alongside electroencephalography (EEG), which may offer a more comprehensive view of the neural mechanisms underlying the N400 in RD. Combining these techniques can enhance the spatial characteristics, complement the temporal resolution strengths of EEG, and provide additional hypotheses related to lexical-semantic processes, which can subsequently be theorized and empirically tested. In our study, we included appropriate assessment of reading measures but perhaps incorporating a comprehensive cognitive assessment alongside N400 investigations, including measures of phonological processing, working memory, and attention, can help explain the relationship between these cognitive domains and N400 patterns in RD. This holistic approach may contribute to a more integrated understanding of reading difficulties. By addressing these future directions, researchers can advance our understanding of the N400 and SfV in the context of RD, ultimately contributing to the development of more targeted and effective interventions for individuals with reading difficulties and build upon the novel work presented in this thesis.

References

Aboud, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental Science*, *19*(4), 632-656. Doi:10.1111/desc.12422

Araújo, S., Faísca, L., Reis, A., Marques, J. F., & Petersson, K. M. (2016). Visual naming deficits in dyslexia: An ERP investigation of different processing domains. *Neuropsychologia*, *91*, 61-76. https://doi.org/10.1016/j.neuropsychologia.2016.07.007

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

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking Up for Vocabulary: Reading Skill Differences in Young Adults. *Journal of Learning Disabilities*, *40*(3), 226–243. https://doi.org/10.1177/00222194070400030401

Barbiero, C., Lonciari, I., Montico, M., Monasta, L., Penge, R., Vio, C., Tressoldi, P.E., Ferluga, V., Bigoni, A., Tullio, A. and Carrozzi, M. (2012). The Submerged Dyslexia Iceberg: How Many School Children Are Not Diagnosed? Results from an Italian Study. *PloS one*, *7*(10), e48082. https://doi.org/10.1371/journal.pone.0048082

Bergmann, J., Hutzler, F., Klimesch, W., & Wimmer, H. (2005). How is dysfluent reading reflected in the ERP? *Journal of Neurolinguistics*, *18*(2), 153-165. https://doi.org/10.1016/j.jneuroling.2004.11.004

Bowey, J. A. (2005). Predicting individual differences in learning to read. *The science of reading: A handbook*, 155-172. https://doi.org/10.1002/9780470757642.ch9

Cardoso-Martins, C., Mesquita, T. C. L., & Ehri, L. (2011). Letter names and phonological awareness help children to learn letter–sound relations. *Journal of Experimental Child Psychology*, *109*(1), 25–38. https://doi.org/10.1016/j.jecp.2010.12.006

Caylak, E. (2009). Neurobiological approaches on brains of children with dyslexia. *Academic Radiology*, *16*(8), 1003-1024 https://doi.org/10.1016/j.acra.2009.02.012

Caylak, E. (2010). The studies about phonological deficit theory in children with developmental dyslexia. *American Journal of Neuroscience*, *1*(1), 1-12.

Coch, D., & Benoit, C. (2015). N400 event-related potential and standardized measures of reading in late elementary school children: Correlated or independent?. *Mind, Brain, and Education*, *9*(3), 145-153. https://doi.org/10.1111/mbe.12083

Coch, D., & Holcomb, P. J. (2003). The N400 in beginning readers. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *43*(2), 146-166. https://doi.org/10.1002/dev.10129

Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, *4*(1), 5-17.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological review*, *100*(4), 589. https://doi.org/10.1037/0033-295X.100.4.589

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1197–1211. https://doi.org/10.1037/0096-1523.20.6.1197

Coles, M. G. H., & Rugg, M. D. (1995). Event-related brain potentials: An introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Oxford psychology series, No. 25. Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 1-26). New York, NY, US: Oxford University Press.

Deacon, S. H., Benere, J., & Castles, A. (2012). Chicken or egg? Untangling the relationship between orthographic processing skill and reading accuracy. *Cognition*, *122*(1), 110-117. https://doi.org/10.1016/j.cognition.2011.09.003

Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. DSM-5 (2013). American Psychiatric Association. American Psychiatric Association Publishing.

Djalali, A., Pierrehumbert, J., & Clark, B. (2008). The effect of focus on bridging inferences. In *LSA Annual Meeting, Chicago, IL. Jan* (pp. 3-6)

Dijk, T. A. V., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., … & Van Petten, C. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*, 1883-1908. https://doi.org/10.1016/j.clinph.2009.07.045

Duta, M. D., Styles, S. J., & Plunkett, K. (2012). ERP correlates of unexpected word forms in a picture–word study of infants and adults. *Developmental Cognitive Neuroscience*, *2*(2), 223-234. Doi: 10.1016/j.dcn.2012.01.003

Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improves children's ability to learn to read words. *Scientific Studies of Reading*, *21*(5), 392-407. https://doi.org/10.1080/10888438.2017.1315424

Ecalle, J., Kleinsz, N., & Magnan, A. (2013). Computer-assisted learning in young poor readers: The effect of grapho-syllabic training on the development of word reading and reading comprehension. *Computers in Human Behavior*, *29*(4), 1368–1376. https://doi.org/10.1016/j.chb.2013.01.041

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*(3), 250-287. https://doi.org/10.1598/RRQ.36.3.2

Elbro, C., de Jong, P. F., Houter, D., & Nielsen, A. M. (2012). From spelling pronunciation to lexical access: The second step in word decoding? *Scientific Studies of Reading*, *16*(4), 341-359. https://doi.org/10.1080/10888438.2011.568556

Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*(1), 1-12. https://doi.org/10.1111/j.1469-8986.2006.00381.x

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757-1761. Doi:10.1097/wnr.0b013e328318f014

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A Meta-Analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74–111. https://doi.org/10.3102/0034654313499616

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, *100*(3), 566. doi:10.1037/0022-0663.100.3.56

Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, *136*(2), 323–345. https://doi.org/10.1037/0096-3445.136.2.323

Gough, P. B., Hoover, W. A., & Peterson, C. L. (2013). Some observations on a simple view of reacting. In *Reading comprehension difficulties* (pp. 25-38). Routledge.

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6-10 ._doi:10.1177/074193258600700104

Grainger, J., & Ziegler, J. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, *2*, 54. https://doi.org/10.3389/fpsyg.2011.00054

Griffiths, Y., & Stuart, M. (2013). Reviewing evidence-based practice for pupils with dyslexia and literacy difficulties. *Journal of Research in Reading*, *36*(1), 96-116. **https://doi.org/10.1111/j.1467-9817.2011.01495.x**

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review, 111*(3), 662–720. https://doi-org.proxy3.library.mcgill.ca/10.1037/0033-295X.111.3.662

Hasko, S., Groth, K., reading processes in children with and without dyslexia: an ERP study. *Frontiers in* Bruder, J., Bartling, J., & Schulte-Körne, G. (2013). The time course of *Human Neuroscience, 7*, 570. https://doi.org/10.3389/fnhum.2013.00570

Hasko, S., Groth, K., Bruder, J., Bartling, J., & Schulte-Körne, G. (2014). What does the brain of children with developmental dyslexia tell us about reading improvement? ERP evidence from an intervention study. *Frontiers in Human Neuroscience, 8*, 441. https://doi.org/10.3389/fnhum.2014.00441

Helenius, P., Parviainen, T., Paetau, R., & Salmelin, R. (2009). Neural processing of spoken words in specific language impairment and dyslexia. *Brain*, *132*(7), 1918-1927. https://doi.org/10.1093/brain/awp134

Hernandez, E. J. (2016). Reducing bullying and preventing dropout through student engagement: A prevention-focused lens for school-based family counselors. *International journal for school-based family counseling*, *7*, 1-13.

Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language*, *117*(2), 88-99. https://doi.org/10.1016/j.bandl.2010.12.003

Hudson, N., Scheff, J., Tarsha, M., & Cutting, L. E. (2016). Reading comprehension and executive function neurobiological findings. *Perspectives on Language and Literacy*, *42*(2), 23.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological science*, *23*(6), 572-577. https://doi.org/10.1177/0956797611435921

Holcomb, P. J., Coffey, S. A., & Neville, H. J. (1992). Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology*, *8*(2-3), 203-241. https://doi.org/10.1080/87565649209540525

Jednoróg, K., Marchewka, A., Tacikowski, P., & Grabowska, A. (2010). Implicit phonological and semantic processing in children with developmental dyslexia: Evidence from event-related potentials. *Neuropsychologia*, *48*(9), 2447-2457. https://doi.org/10.1016/j.neuropsychologia.2010.04.017

Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and Linguistic Compass 1*, 571-591 https://doi.org/10.1111/j.1749-818X.2007.00037.x

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 Component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621-647. Doi:10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological psychology*, *11*(2), 99-116. https://doi.org/10.1016/0301-0511(80)90046-0

Kutas, M., & Hillyard, S. A. (1980b). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205. https://doi.org/10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Event-related potentials in cognitive science. In *Handbook of cognitive neuroscience* (pp. 387-409). Boston, MA: Springer US.

Kutas, M., & Iragui, V. (1998). The N400 in a semantic categorization task across 6 decades. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *108*(5), 456–471. https://doi.org/10.1016/S0168-5597(98)00023-9

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighbourhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*, *61*(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: an 8-year follow-up. *Journal of educational psychology*, *100*(1), 150. https://psycnet.apa.org/doi/10.1037/0022-0663.100.1.150

Landi, N. (2012). Learning to read words: Understanding the relationship between reading ability, lexical quality, and reading context. In *Reading-From words to multiple texts* (pp. 17-33). Routledge.

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and language*, *102*(1), 30-45. https://doi.org/10.1016/j.bandl.2006.11.001

Landi, N., Frost, S. J., Mencl, W. E., Sandak, R., & Pugh, K. R. (2013). Neurobiological bases of reading comprehension: Insights from neuroimaging studies of word-level and text-level processing in skilled and impaired readers. *Reading & Writing Quarterly*, *29*(2), 145-167. https://doi.org/10.1080/10573569.2013.758566

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature reviews neuroscience*, *9*(12), 920-933. https://doi.org/10.1038/nrn2532

Lovett, M. W., Steinbach, K. A., & Frijters, J. C. (2000). Remediating the Core Deficits of Developmental Reading Disability: A Double-Deficit Perspective. *Journal of Learning Disabilities*, *33*(4), 334–358. https://doi.org/10.1177/002221940003300406

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

Luck (2014). *An introduction to event related potential technique*. MIT Press: London, England

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101*(1), 3–33. https://doi.org/10.1037/0033-295X.101.1.3

McArthur, G., Kohnen, S., Jones, K., Eve, P., Banales, E., Larsen, L., & Castles, A. (2015). Replicability of sight word training and phonics training in poor readers: a randomised controlled trial. *PeerJ*, *3*, e922. https://doi.org/10.7717/peerj.922

McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H.-C., Kohnen, S., Banales, E., Anandakumar, T., Marinus, E., & Castles, A. (2018). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*, *11*. https://doi.org/10.1002/14651858.CD009115.pub3

McPherson, W. B., Ackerman, P. T., Holcomb, P. J., & Dykman, R. A. (1998). Event-related brain potentials elicited during phonological processing differentiate subgroups of reading disabled adolescents. *Brain and language*, *62*(2), 163-185. https://doi.org/10.1006/brln.1997.1893

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(1), 53-65. https://doi.org/10.1017/S0048577299971196

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological bulletin*, *138*(2), 322. https://doi-org.proxy3.library.mcgill.ca/10.1037/a0026744

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol. 1. Theory and methods* (pp. 357–390). John Wiley & Sons.

Mousikou, P., Coltheart, M., Finkbeiner, M., & Saunders, S. (2010). Can the dual route cascaded computational model of reading offer a valid account of the masked onset priming effect? *The Quarterly Journal of Experimental Psychology*, *63*(5), 984-1003. https://doi.org/10.1080/17470210903156586

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

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.

Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry*, *51*(9), 1031–1039. https://doi.org/10.1111/j.1469-7610.2010.02254.x

Nation, K., & Norbury, C. F. (2005). Why reading comprehension fails: Insights from developmental disorders. *Topics in language disorders*, *25*(1), 21-32.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*(1), 85-101. https://doi.org/10.1006/jmla.1998.2564

Nobre, A. D. P., & Salles, J. F. D. (2016). Lexical-semantic processing and reading: Relations between semantic priming, visual word recognition and reading comprehension. *Educational Psychology*, *36*(4), 753-770. https://doi.org/10.1080/01443410.2014.950948

Noltemeyer, A. L., Joseph, L. M., & Kunesh, C. E. (2019). Effects of Supplemental Small Group Phonics Instruction on Kindergartners' Word Recognition Performance. *Reading Improvement*, *56*(3), 149–160.

Nuss, B. (2018). *Base-1 method: A structural-functional approach to word, sentence and discourse readability*. Waxmann Verlag.

Osterhout, L., & Holcomb, P. J. (1995). Event-related potentials and language comprehension. *Electrophysiology of mind: Event-related brain potentials and cognition*, *25*, 171-215.

Ott, P. (2008). *Activities for Successful Spelling: The Essential Guide*. Routledge.

O'Connor, R. E., Bell, K. M., Harty, K. R., Larkin, L. K., Sackor, S. M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology*, *94*(3), 474–485. https://doi.org/10.1037/0022-0663.94.3.474

Partanen, M., Siegel, L. S., & Giaschi, D. E. (2019). Effect of reading intervention and task difficulty on orthographic and phonological reading systems in the brain. *Neuropsychologia*, *130*, 13-25. https://doi.org/10.1016/j.neuropsychologia.2018.07.018

Perfetti, C. A. (1985). *Reading ability*. Oxford University Press. http://dx.doi.org/10.1037/10459-004

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. Doi:10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy*, *11*, 67-86.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 227–247). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470757642.ch13

Perfetti, C., & Stafura, J. (2014). Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Plante, E., Van Petten, C., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia*, *38*(13), 1669-1684

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological review*, *103*(1), 56.

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A Computational Model of the Self-Teaching Hypothesis Based on the Dual-Route Cascaded Model of Reading. *Cognitive Science*, *42*(3), 722-770. https://doi.org/10.1111/cogs.12571

Rüsseler, J., Becker, P., Johannes, S., & Münte, T. F. (2007). Semantic, syntactic, and phonological processing of written words in adult developmental dyslexic readers: an event-related brain potential study. *BMC neuroscience*, *8*(1), 1-10. https://doi.org/10.1186/1471-2202-8-52

Sabisch, B., Hahne, A., Glass, E., von Suchodoletz, W., & Friederici, A. D. (2006). Auditory language comprehension in children with developmental dyslexia: evidence from event-related brain potentials. *Journal of cognitive neuroscience*, *18*(10), 1676-1695. https://doi.org/10.1162/jocn.2006.18.10.1676

Savage, R. (2019). Literacy, basic processes, and interventions for struggling readers. *Oxford Research Encyclopedia, Education* 10.1093/acrefore/9780190264093.013.925

Savage, R., Georgiou, G., Parrila, R., & Maiorino, K. (2018). Preventative reading interventions teaching direct mapping of graphemes in texts and set-for-variability aid at-risk learners. *Scientific Studies of Reading*, *22*, 225-247. https://doi.org/10.1080/10888438.2018.1427753

Savage, R., Maiorino, K., Gavin, K., Horne-Robinson, H., Georgiou, PhD, G., & Deacon, H. (2023). Contrasting Direct Instruction in Morphological Decoding and Morphological Inquiry-Analysis Interventions in Grade 3 Children with Poor Morphological Awareness. *Journal of Learning Disabilities*, https://doi.org/10.1177/00222194231161117

Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (2002). Relationship of rapid automatized naming and phonological awareness in early reading development: Implications for the double-deficit hypothesis. *Journal of learning disabilities*, *35*(3), 245-256. https://doi.org/10.1177/002221940203500306

Scanlon, D. (2013). Specific learning disability and its newest definition: Which is comprehensive? And which is insufficient? *Journal of learning disabilities*, *46*(1), 26-33. https://doi.org/10.1177/0022219412464342

Schuberth, R.E., Eleimas, P.D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Preparation and Performance, 3*, 27-36.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Annals of Internal Medicine*, *152*(11), 726–732. https://doi.org/10.7326/0003-4819-152-11-201006010-00232

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568. https://doi-org.proxy3.library.mcgill.ca/10.1037/0033-295X.96.4.523

Seymour, P. H. (1997). Foundations of orthographic development. *Learning to spell: Research, theory, and practice across languages*, 319-337.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua none of reading acquisition. *Cognition*, *55*(2), 151-218.

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry*, *57*(11), 1301-1309. Doi: 10.1016/j.biopsych.2005.01.043

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., … & Gore, J. C. (2003). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological psychiatry*, *52*(2), 101-110.

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

Snow, C. (2002). *Reading for Understanding. Towards an R&D Program in Reading Comprehension*. RAND CORP SANTA MONICA CA. https://apps.dtic.mil/sti/citations/ADA402712

Sun, Y. F., Lee, J. S., & Kirby, R. (2010). Brain imaging findings in dyslexia. *Pediatrics & Neonatology*, *51*(2), 89-96. Doi:10.1016/S1875-9572(10)60017-4

Statistics Canada (2012). Canadian survey on disability (CSD). Retrieved from: https://www150.statcan.gc.ca/n1/pub/89-654-x/89-654-x2014001-eng.ht

Steacy, L. M., Elleman, A. M., & Compton, D. L. (2017). Opening the "black box" of learning to read. *Theories of reading development*, *15*, 99. https://dx.doi.org/10.1075/swll.15.06ste

Steacy, L. M., Wade-Woolley, L., Rueckl, J. G., Pugh, K. R., Elliott, J. D., & Compton, D. L. (2019). The Role of Set for Variability in Irregular Word Reading: Word and Child Predictors in Typically Developing Readers and Students At-Risk for Reading Disabilities. *Scientific Studies of Reading*, *23*(6), 523–532. https://doi.org/10.1080/10888438.2019.1620749

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

Tomlinson, C. A., & Imbeau, M. B. (2023). *Leading and managing a differentiated classroom*. Ascd.

Torppa, M., Parrila, R., Niemi, P., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2013). The double deficit hypothesis in the transparent Finnish orthography: A longitudinal study from kindergarten to Grade 2. *Reading and Writing*, *26*(8), 1353–1380. https://doi.org/10.1007/s11145-012-9423-2

Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, *16*(2), 122-140.

Venezky, R. L. (1999). *The American Way of Spelling: The Structure and Origins of American English Orthography*. Guilford Press.

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current Evidence on the Effects of Intensive Early Reading Interventions. *Journal of Learning Disabilities*, *51*(6), 612–624. https://doi.org/10.1177/0022219418775110

Weber, R. M. (2018). Listening for schwa in academic vocabulary. *Reading Psychology*, *39*(5), 468-491. https://doi.org/10.1080/02702711.2018.1464531

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

Wolf, M., & Denckla, M. B. (2005). *RAN/RAS: Rapid automatized naming and rapid alternating stimulus tests*. Austin, TX: Pro-ed.

Wolff, U. (2011). Effects of a Randomised Reading Intervention Study: An Application of Structural Equation Modelling. *Dyslexia*, *17*(4), 295–311. https://doi.org/10.1002/dys.438

Yeung, S. S. S., & Savage, R. (2020). Teaching Grapheme–Phoneme Correspondences Using a Direct Mapping Approach for At-Risk Second Language Learners: A Randomized Controlled Trial. *Journal of learning disabilities*, *53*(2), 131-144.

Yuill, N., & Oakhill, J. V. (1991). *Children's problems in text comprehension.* Cambridge, UK: Cambridge Univ. Press.

Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models. *Current Directions in Psychological Science*, *29*(3), 293–300. https://doi.org/10.1177/0963721420915873

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F.-X., & Perry, C. (2008). Developmental dyslexia and the dual-route model of reading: Simulating individual differences and subtypes. *Cognition*, *107*(1), 151–178. https://doi.org/10.1016/j.cognition.2007.09.004

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), https://doi.org/10.1098/rstb.2012.0397

Zipke, M. (2016). The importance of flexibility of pronunciation in learning to decode: A training study in set for variability. *First Language*, *36*(1), 71-86.

Copyright

Appendix A

McGill University
Research Ethics Board Office
www.mcgill.ca/research/research/compliance/human

REB File Number:	22-02-077
Project Title:	Assessing the Behavioral and Physiological Effect of a Reading Intervention on School-age Children
Principal Investigator:	Badriah Basma
Department:	Educational & Counselling Psychology
Supervisor:	Professor Gigi Luk

Approval Expiry Date: **19-Apr-2024**

- The *REB-3* reviewed and approved the Continuing Review application for the above project on 20-Mar-2023.
- Approval is granted only for the research and purposes described.
- The PI must inform the REB if there is a termination or interruption of their affiliation with the University.
- An **Amendment** form must be used to submit any proposed modifications to the approved research. Modifications to the approved research must be reviewed and approved by the REB before they can be implemented.
- Changes to funding or adding new funding to a previously unfunded study must be submitted as an Amendment.
- A **Continuing Review** form must be submitted before the above expiry date. Research cannot be conducted without a current ethics approval. Submit 2-3 weeks ahead of the expiry date. A total of 5 renewals are permitted after which time a new application will need to be submitted.
- A **Termination** form must be submitted to inform the REB when a project has been completed or terminated.
- A **Reportable New Information** form must be submitted if any unanticipated issues that may increase the risk level to participants or that may have other ethical implications or to report any protocol deviations that did not receive prior REB approval.
- The REB must be promptly notified of any new information that may affect the welfare or consent of participants.
- The REB must be notified of any suspension or cancellation imposed by a funding agency or regulatory body that is related to this study.
- The REB must be notified of any findings that may have ethical implications or may affect the decision of the REB.

Appendix B

FW: Research Proposal: Assessing the Behavioral and Physiological Effet of a Reading Intervention in School-Aged Children with a Reading Disability

Mancini, Gina G. <gmancini@emsb.qc.ca>
Tue 1/24/2023 1:12 PM
To:Badriah Yasir Basma, Miss <badriah.basma@mcgill.ca>
Cc:Sanalitro, Anna <ASanalitro@emsb.qc.ca>;Villalta, Anna <AVillalta@emsb.qc.ca>;Galanogeorgos, Athina <AGalanogeorgos@emsb.qc.ca>;Argyrakopoulos, Elpis <EArgyrakopoulos@emsb.qc.ca>;Bixby, Jason Lloyd <jbixby@emsb.qc.ca>;Napoletano, Anthony <ANapoletano@emsb.qc.ca>;Sansone, Patrizia <PSansone@emsb.qc.ca>;Dockery, Karla <KDockery@emsb.qc.ca>;Kij, Anna <AKij@emsb.qc.ca>;Daskalakis, Jim <jdaskalakis@emsb.qc.ca>;Manstavich, Steven <SMANSTAVICH@emsb.qc.ca>

Good day Ms Basma,

On behalf of Anna Sanalitro, Director of Educational Services, and Anna Sanalitro, Chairman of the Research Committee,
we are happy to inform you that the Education Policy Committee has now also approved your research proposal
and this study can now be presented to our schools.

You may contact the EMSB School Principals directly by sending them this email as confirmation of our approval
for this research.

The Research Committee wishes you success with this project and would very much appreciate receiving, at
the end of this study, a brief summary of your findings as well as a copy of your final report.

Respectfully,

Gina Mancini, B.B.A., B.Ed.
Secrétaire exécutif du comité de recherche , Service de l'éducation
Executive Secretary for the Research Committee , Education Services

Tel.: 514-483-7200 poste / ext. 7359
F: 514-483-7246
gmancini@emsb.qc.ca

Commission scolaire English-Montréal
English Montreal School Board

From: Badriah Yasir Basma, Miss <badriah.basma@mcgill.ca>
Sent: Tuesday, December 20, 2022 10:56 AM
To: Mancini, Gina G. <gmancini@emsb.qc.ca>
Subject: Re: Research Proposal: Assessing the Behavioral and Physiological Effet of a Reading Intervention in School-Aged Children with a Reading Disability

⚠ **Caution:** This email originated from outside of the EMSB. Use caution when clicking links or opening attachments. | **Avertissement:** Ce courriel provient de l'extérieur de la CSEM. Soyez prudent lorsque vous cliquez sur des liens ou ouvrez des pièces jointes

Appendix C

COMMISSION SCOLAIRE SIR-WILFRID-LAURIER
SIR WILFRID LAURIER SCHOOL BOARD

January 19, 2023

Dear Ms. Basma,

Your research project request entitled *Assessing the Behavioral and Physiological Effect of Reading Intervention in School-aged Children with a Reading Disability* has been reviewed.

We are pleased to inform you that your request to conduct research has been accepted.

We would appreciate it if you could communicate with Geneviève Légaré (glegare@swlauriersb.qc.ca) once you have confirmed the name and location where your research will be conducted.

Please note that it is the responsibility of the researcher(s) to contact the school(s) or center(s) directly to proceed with the research process. The principal must approve the project and the school's participation in order for the researcher to proceed. The principal reserves the right to decline the school's participation.

A report must be submitted to the Director of Pedagogical Services if the project is to be ongoing for more than one year. Upon completion of the project, a one-two page executive summary containing a brief description of the project, findings and future implications of the research is to be submitted to the Director of Pedagogical Services.

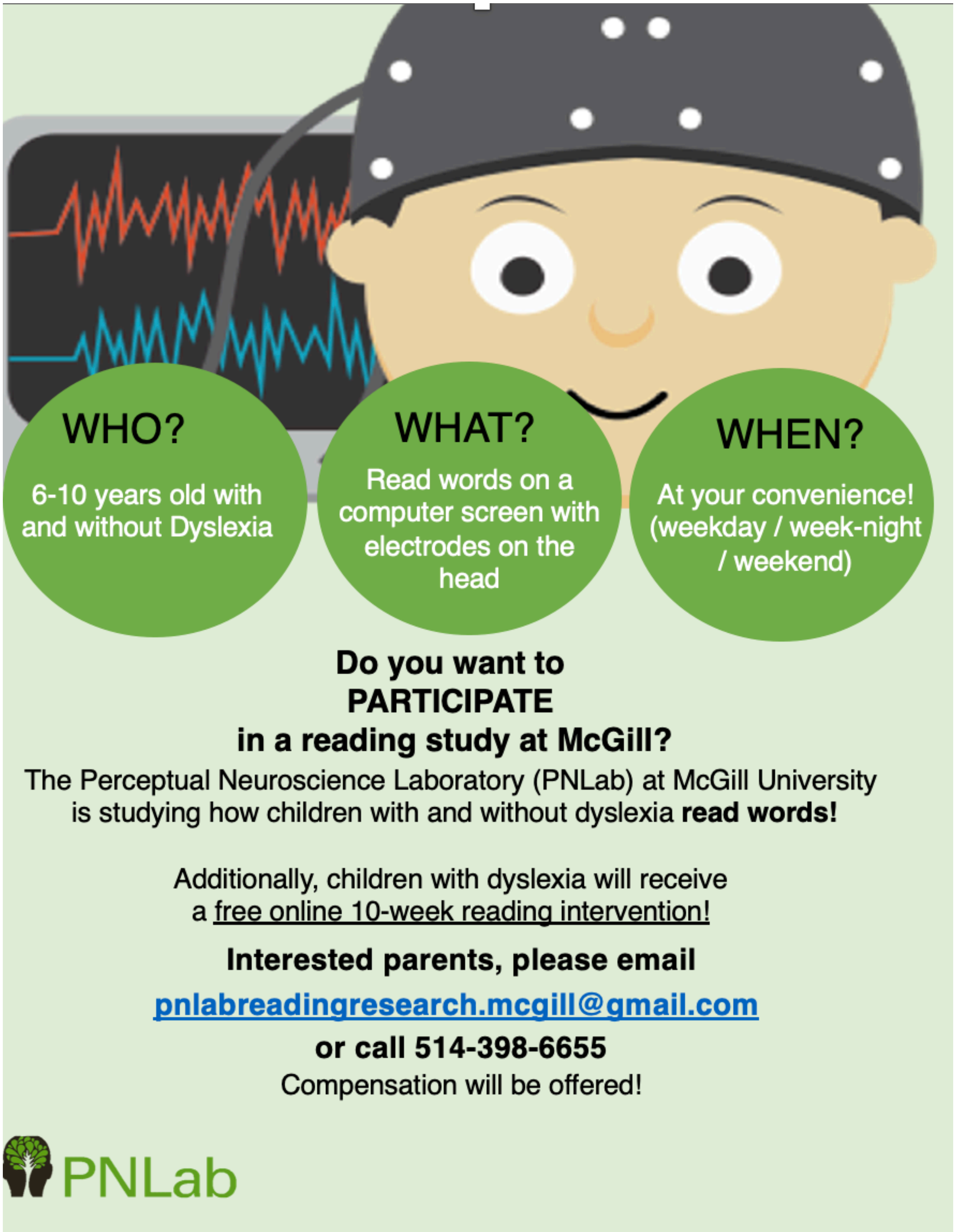
Regards,

Lynda DaSilveira
Director of Pedagogical Services
Sir Wilfrid Laurier School Board
(450) 621-5600 ext. 1377
ldasilveira@swlauriersb.qc.ca

cc: Geneviève Légaré

T 450-621-5600
1 866-621-5600
F 450-965-4208
Pedagogical Services
239, montée Lesage
Rosemère (Québec)
J7A 4C4

Appendix D


WHO?
6-10 years old with and without Dyslexia
WHAT?
Read words on a computer screen with electrodes on the head
WHEN?
At your convenience! (weekday / week-night / weekend)
Do you want to PARTICIPATE in a reading study at McGill?
The Perceptual Neuroscience Laboratory (PNLab) at McGill University is studying how children with and without dyslexia read words!
Additionally, children with dyslexia will receive a free online 10-week reading intervention!
Interested parents, please email
pnlabreadingresearch.mcgill@gmail.com
or call 514-398-6655
Compensation will be offered!
PNLab

Appendix E

Parental Consent form for the Intervention Group

Researcher: Badriah Basma, M.Ed., Ph.D. candidate
Department of Education and Counselling Psychology Human Development Program
Faculty of Education, McGill University

Supervisor: Dr. Gigi Luk
Department of Education and Counselling Psychology
Faculty of Education, McGill University

Co-Supervisor: Dr. Armando Bertone, Ph.D.
Department of Education and Counselling Psychology School Applied Psychology
Faculty of Education, McGill University

Co-Supervisor: Dr. Robert Savage, Ph.D.
Department of Education
Faculty of Education York University

Title of Project: Assessing the Behavioral and Physiological Effect of a Reading Intervention in School-age Children.

Introduction: Reading skills play an essential role in developing the ability to recognize individual words, read text fluently and accurately, and understand what we read. We are interested in seeing if a reading intervention based on letter-sound variation will influence reading ability in elementary school students with reading disabilities (e.g., Dyslexia). We ultimately want to determine if improving students' reading abilities with reading disabilities can improve their reading performance.

This study will assess your child's reading level using behavioural and physiological measures. For the behavioural measures, we will use reading assessment measures by asking them to complete different reading tasks: read a list of words, answer questions about short stories, identify and blend the sound of words, read non-real words, read, and correct mispronounced words, remember the meaning of words. For the physiological measures, we will use an electroencephalogram (EEG). The EEG will measure if your child can process whether the sentence has a whole meaning or not.

Procedure. All EEG and reading measures will occur at McGill University's PNLAB, located at 3744 McTavish St. (corner avenue des Pins), at a time convenient to you. Parking will be available to you in front of the building. The EEG is a harmless tool that records brain wave patterns. It will accurately measure your child's brain responses to a word in real-time. A cap will be placed on your child's head with sensors disks. To record the brain's activity, these disks will be filled will a water-soluble gel that allows the sensors to record brain activity at the scalp. These sensors will send signals to a computer to record the brain activity while your child reads words and answers questions about what they read. Once your child finishes the training, the sensors will be gently removed, and any gel remaining on your child's hair will be cleaned with a moist towel provided by our lab; this procedure is painless. Children

Appendix F

Parental Consent Form for Neurotypicals

Researcher: Badriah Basma, M.Ed., Ph.D. candidate
Department of Education and Counselling Psychology Human Development Program
Faculty of Education, McGill University

Supervisor: Dr. Gigi Luk
Department of Education and Counselling Psychology
Faculty of Education, McGill University

Co-Supervisor: Dr. Armando Bertone, Ph.D.
Department of Education and Counselling Psychology School Applied Psychology
Faculty of Education, McGill University

Co-Supervisor: Dr. Robert Savage, Ph.D.
Department of Education
Faculty of Education, York University

Title of Project: Assessing the Behavioral and Physiological Effect of a Reading Intervention in School-Children.

Introduction: Reading skills play an essential role in developing the ability to recognize individual words, read text fluently and accurately, and understand what we read. We are interested in seeing if a reading intervention based on letter-sound variation will influence reading ability in elementary school students with reading disabilities (e.g., Dyslexia). We ultimately want to determine if improving students' reading abilities with reading disabilities can improve their reading performance.

This study will assess your child's reading level using behavioural and physiological measures. For the behavioural measures, we will use reading assessment measures by asking them to complete different reading tasks: read a list of words, answer questions about short stories, identify and blend the sound of words, read non-real words, read and correct mispronounced words, remember the meaning of words. For the physiological measures, we will use an electroencephalogram (EEG). The EEG will measure if your child can process whether the sentence has a whole meaning or not.

Procedure. All EEG and reading measures will occur at McGill University's PNLAB, located at 3744 McTavish St. (corner avenue des Pins), at a time convenient to you. Parking will be available to you in front of the building. The EEG is a harmless tool that records brain wave patterns. It will accurately measure your child's brain responses to a word in real-time. A cap will be placed on your child's head with sensor disks. To record the brain's activity, these disks will be filled will a water-soluble gel that allows the sensors to record brain activity at the scalp. These sensors will send signals to a computer to record the brain activity while your child reads words and answers questions about what they read. Once your child finishes the training, the sensors will be gently removed, and any gel remaining on your child's hair will be cleaned with a moist towel provided by our lab; this procedure is painless. Children do not feel any electrical impulses. The total time spent in our lab - including already scheduled and requested breaks – will be 2.5 hours.

Appendix G

Mispronunciation correction task stimuli (Tunmer & Chapman, 1998)

The taught words are in bold.

Her granny is very **kind** (pronounce: rhyme with wind).

He got mud on his **shoe** (pronounce: show)

The dog had to have a **wash** (pronounce: rhyme with ash) He put suntan lotion on his body (pronounce: boady)

He couldn't find his **money** (pronounce: moaney)

In France, they have great **weather** (pronounce: weet-her)

The man repaired the broken **watch** (pronounce: rhyme with catch)

He spilt spaghetti all down his **front** (pronounce: froant)

The children's granny baked some **bread** (pronounce: breed)

We got very cold swimming in the **river** (pronounce: rive-er rhyme with fiver)

They searched for the **treasure** (pronounce tree-sore)

The friends shared a **biscuit** (pronounce: bis-coo-it)

The child used the blocks to build a **castle** (pronounce: cast-el)

The cake was shaped like a **heart** (pronounce: rhyme with hear-t)

He washed the plastic **bowl** (pronounce bowel)

For a snack he ate a **banana** (pronounce: ban-ay-nar)

Last year there was a big **flood** (pronounce: fl-oo-d)

The dog chased the **lamb** (pronounce lam-b)

He lost his **glove** (pronounce: rhyme with clove)

Appendix H

Set-for-Variability Intervention (SfV)

Example of a structured lesson.

<u>GPC & Direct Mapping</u>

Review (1min): Recall the word from the previous lesson, the sound, and how it is written.

Introduce a new word (This) + definition (1min)

Explain: "The word of the day" is "this" do you know what "this" means? Define it if participants do not learn.

Children at this stage do not yet see the word.

Spell the word (1min):

Have letter cards. Ask children how they think the word was spelled. Once they figure it out, they write the word on mini whiteboards or pencil/paper. Verify if the spelling is correct.

Introduce CPC of the day (2mins)

Explain which GPC in the word made the sound of the day by linking the spelling and the pronunciation of the GPC and articulated phonemes themselves. Explain how the letter sound is written. Explain that the S (grapheme) makes the sound sssssss (phoneme). The students can form the grapheme S in the air and make the sound sssss. The teacher writes down the word "this" and circles the grapheme s.

Find words with special GPC in the text (2mins): Read the text and ask the participants to spot the words that contain the GPC. Once the participant spot it, ask the participants to say the GPC and the word.

Storyline: This is a sunny day. A boy takes his dog for a walk. The boy throws a stick, and the dog runs after it. The dog sniffs around in the grass and starts barking. A large, spotty snake is sitting up and hissing, /sssssss/. The boy grabs the dog, and the snake slowly slithers away. Children create their sentences (this, as you mentioned, is for higher ability students) with younger or more poorly performing students.

Children create their sentences (extension for higher ability) (2 mins)

Ask participants to create their sentences using the GPC and word of the day (higher ability children) on a mini whiteboard. Teacher/TAs check for the spelling of the GPC / word of the day and the other words in the sentence. Lower-ability children try to read and write the grapheme.

The variant of the GPC *s*

Explicitly Explain (2 min) the grapheme S makes another sound: zzzzzzz in the word “is.”

s: Often (but sadly not always) ‘soft’ unvoiced ‘s’ occurs at the *beginning* of a word (e.g., ‘sun’ ‘sand’ ‘sea’) while the voiced ‘s’ often denotes a plural at the end of a word (‘suns,’ ‘seas’). However, beware of exceptions such as ‘presents! Thus teach the pattern as a *tendency* for this to happen.

Blending: Show students index cards with the GPC blends written on them. Practice saying each combination with the student. You have an example of using S Blends in index cards: *so, say, this, is, bus, as.*

To ensure that the children grasped the concept and can generalize it, Use the script below to teach the variable GPC ***S.*** Use *untaught* items and see if they got the variant. These are untaught items: **his, has, was**.

Teaching vocabulary: exceptional words

Today's word is **his**

A requirement for the work on **exception words** is that children know the meaning of these words (have the word in their vocabulary), so as a preliminary, they have been checked against published norms for the typical age of acquisition. The item ages of the acquisition were checked against norms from Brysbaert and Biemiller (2017).

Evidence suggests teaching vocabulary works well when taught across clustered items – words that are in some way linked around shared/related concepts.

When teaching vocabulary - saying, using, and seeing the print and writing words repeatedly is important.

Spacing out the teaching of vocabulary over several sessions and building in time to review words that have been taught is also key to helping children to embed words into their vocabulary.

Duration: 5 minutes.

Materials: whiteboard – markers

Write on the board, “fill in the blank in this sentence.”

This is _________ book (his).

I put _______ book on the desk. (his)

Make sure each child can do at least two sentences orally using the word “his” in their own words.

Oral Set-for-Variability (5 minutes)

Tell: We are going to play a Simon says game on body parts. Simon might say words that do not make sense. Remember, we use the flipping sound strategy to know the word's correct spelling.

For regular words: follow these instructions:

‘Simon says’ task (Simon says touch you...e.g. ‘ar’-‘m’). The participant should be able to touch their arm.

Simon says: touch your

Ea-r

Ar-m

L-i-p

L-e-g

H-i-p

H-a-n-d

B-a-ck

For **exception words**, follow these instructions.

The exception words should be sounded out using regularised pronunciation, so children must attempt to swap in an alternative sound for the word to make sense.

Teacher: *Simon says!*" touch your

N-o-se (with unvoiced /s/)

K-n-ees (with voiced /k/ /s/)

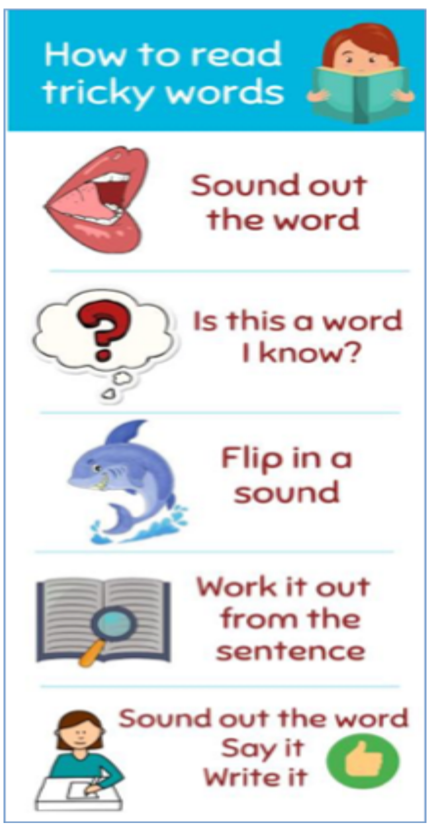
M-u-s-c-u-l (s) (with unvoiced /s/)

Sh-ou-l-d-er (with 'ou' as in 'out')

S-t-o-m-a-ch ('sto-mach,' 'stow'-'match')

F-oo-t (so 'oo' rhymes with boot)

Mispronunciation Correction (5 minutes)



Materials: magnetic letters

Tell today; we are going to learn a new common exception word. Put the magnetic letters "h," "I," and "s" together and blend the word (with unvoiced s)

Ask: does that sound like a word? Is that a word you know? It is not a word that I know either. This means that it might be one of those words where one of the sounds we need to flip so that it's a word that we know.

So, let's see if we can find this word in our book. I will start reading, and you tell me if you spot this word.

How to catch a star by Oliver Jeffers

Once there was a boy, and the boy loved stars very much. Every night the boy watched the stars from ***his*** (unvoiced s) window and wished he had one of ***his*** (unvoiced s) very own.

Once they point at the word, read the sentence and ask (does that make sense – his (unvoiced s). What word would you think might make sense more than his (unvoiced s)?

Once they get the correct answer, tell: correct, we had to change one of those sounds to make a word that we know, so which letter did we have to change the sound of? [the answer should be we changed the unvoiced *s* to a voiced *s*].

Continue reading the story. Stop at the target word and ask the participant to read it.

The remaining of the story goes as follows:

He dreamed how this star might be ***his*** friend. They would play hide and go seek and take long walks together. The boy decided he would try to capture the star. He thought that getting up early in the morning would be best because the star would be tired from being up in the sky all night. So, the next day, he set out at sunrise but could not see a star anywhere. He sat down and waited for one to appear. He waited, and he waited and ate lunch and waited, and after dinner, he waited for some. Finally, just before the sun disappeared, he saw a star. The boy tried to jump up and grab it, but he could not jump high enough. So very carefully, he climbed to the top of the tallest tree. He could find it, but the star was still far from reach. He thought he might get the star with a lifebelt from his father's boat, but it was too heavy for him to carry. He thought he could fly up in his spaceship and grab the star, but ***his*** spaceship had run out of petrol last Tuesday when he flew to the moon. Perhaps he could get a seagull to help him fly up in the sky to reach ***his*** star, but the only Siegel he could find didn't want to help. The boy thought he would never catch a star just then he noticed something floating in the water it was the prettiest star he had ever seen, just a baby star it must have fallen from the sky he tried to fish the star out with ***his*** hands. Still, he couldn't reach it then he had an idea the star might wash up on the shore. He ran back along the jetty to the beach, then he waited and walked and watched and waited, and sure enough, the star washed up on the bright golden sands. The boy had caught ***his*** star, a star of ***his*** very own.

Current Best Practices

The current best practices group will receive exactly the same phonics instruction without the set-for-variability component. It will also follow the same GPC table without the sound variation.

It will go as follows:

GPC & Direct Mapping

Review (1min): Recall the word from the previous lesson, what the sound was, and how it is written.

Introduce a new word (This) + definition (1min)

Explain: "The word of the day" is "this" do you know what "this" means? Define it if participants do not learn.

Children at this stage do not yet see the word.

Spell the word (1min):

Have letter cards. Ask children how they think the word was spelled. Once they figure it out, they write the word on mini whiteboards or pencil/paper. Verify if the spelling is correct.

Introduce CPC of the day (2mins)

Explain which GPC in the word made the sound of the day by linking the spelling and the pronunciation of the GPC and articulated phonemes themselves. Explain how the letter sound is written. Explain that the S (grapheme) makes the sound sssssss (phoneme). The students can form the grapheme S in the air and make the sound sssss. The teacher writes down the word "this" and circles the grapheme s.

Find words with special GPC in the text (2mins): Read the text and ask the participants to spot the words that contain the GPC. Once the participant spot it, ask the participants to say the GPC and the word.

Storyline: This is a sunny day. A boy takes his dog for a walk. The boy throws a stick, and the dog runs after it. The dog sniffs around in the grass and starts barking. A large, spotty snake is sitting up and hissing, /sssssss/. The boy grabs the dog, and the snake slowly slithers away.

Children create their sentences (this, as you mentioned, is for higher ability students) with younger or more poorly performing students

Children create their sentences (extension for higher ability) (2 mins)

Ask participants to create their sentences using the GPC and word of the day (higher ability children) on a mini whiteboard. Teacher/TAs check for the spelling of the GPC / word of the day and the other words in the sentence. Lower-ability children try to read and write the grapheme.

Vocabulary word of the day: 5 minutes

Materials: whiteboard – markers

Write on the board, “fill in the blank in this sentence.”

This shining ________ shows up in the sky at night.

I watch the shining ___________ at night with my brothers.

Make sure each child can do at least two sentences. Let them try even if they find it difficult.

Reading the letter “S”-

We will read this together. Once you hear the sound ssssss, let me know and we reread the word again.

Zoom!

Robert Munsch

When her mother came to pick her up at school, Lauretta said, "Look at this ratty old wheelchair! I've had it since forever. I need a new wheelchair!"

"Guess what?" said her mother. "We are getting one today! I wanted it to be a surprise!"

- -

So they went to the wheelchair store to get a nice new wheelchair.

- -

Lauretta's mother said, "How about this? Look at this! A nice new five-speed wheelchair"

Lauretta rode the wheelchair around the store:

ZOOOOOM

ZOOOOOM

ZOOOOOM

and said, "Too slow."

- -

Then Lauretta's mother said, "Well, how about this? Look at this! A nice new ten-speed wheelchair."

Lauretta rode the wheelchair around the store:

ZOOOOOOOOOOM

ZOOOOOOOOOOM ZOOOOOOOOOOM and said, "Too slow."