

When Size Matters:
Effects of Context and Isolation Training on Orthographic Learning

Monyka L. Rodrigues

A Thesis in the Department of Education

Presented in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy (Education) at

Concordia University

Montreal, Quebec, Canada

May 2023

© Monyka L. Rodrigues, 2023

**CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Monyka Legault-Rodrigues

Entitled: When size matters: Effects of Context and Isolation Training on Orthographic Learning

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

<u>Dr. Saul Carliner</u>	Chair
<u>Dr. Nicole Conrad</u>	External Examiner
<u>Dr. Helena Osana</u>	Examiner
<u>Dr. Kim McDonough</u>	Examiner
<u>Dr. Julie Corrigan</u>	Examiner
<u>Dr. Sandra Martin-Chang</u>	Thesis Supervisor (s)

Approved by

Dr. Saul Carliner

Chair of Department or Graduate Program Director

July 6th 2023

Date of Defence

Dean,

Abstract

When Size Matters:

Effects of Context and Isolation Training on Orthographic Learning

Monyka L. Rodrigues, PhD
Concordia University, 2023

When learning to read, children are regularly exposed to words with novel symbols and spellings in stories and lists. I investigated the differential effect of self-teaching in and out of context, as well as the moderating effect of stimulus set size, word length, and word regularity, on the formation of orthographic representations. In Study 1, adult participants practiced reading regularised spellings of real words that were written in an artificial orthography across four experiments ($N = 114$). Participants trained with either a small set of words or with a large set of words. The first dependent variable was word reading accuracy measured during training (over 6 trials) and 3-days later at post-test. The second and third dependent variables were spelling recognition and production also measured 3-days later at post-test. Generalization of learning was also explored. Results indicate that training in context bolstered reading accuracy during and after training. Yet the highest spelling scores were noted following training in isolation. Further, self-teaching a large set of diverse words in isolation increased reading accuracy, closing the gap with context. Overall, self-teaching in context helped establish orthographic representations that supported reading accuracy. Whereas reading in isolation more highly specified orthographic representations, enabling spelling accuracy—especially when training with a large set of words. Study 2 explored whether these effects could be obtained with children. Children read regular and irregular sets of words of varying length written in standard English orthography across two experiments ($N = 64$). Reading practice took place in both context and isolation and stimulus set size was experimentally manipulated. Children participants were then asked to complete a variety of reading and spelling retention tasks at

post-test. Reading in context supported reading accuracy during training and 1 week later. Like adults, children spelled more words correctly from the isolation training condition. The greatest isolation advantage was for exception words. We also noted that when children had fewer opportunities to map orthographic representations onto pronunciations, training in context was more advantageous than training in isolation. This signals the scope and limitations of reading in context and highlights the facilitatory effect of reading amount.

Acknowledgements

Where to begin. My dissertation work was sandwiched between moving back to Montréal, losing my grandfather to cancer, making new friends in the program, developing a close relationship with my supervisor, the cool projects I worked on, surviving a pandemic, a huge last push to collect data and write it all up, and a severely broken foot. So, I sit here at my laptop thinking about all the people who got me to the finish line. My grandmother who asked me every week *pis, ça avance-tu ton travail dans tes papiers?* My father who was always ready to question the face validity of my research; he does not know it but, it pushed me to practice my elevator pitch. This dissertation is for them. My mother who—I do not even know where to start—thank you for everything. Serge, who cannot wait to call me Dr. Sunshine.

My number one fans.

And Sandi, my PhD supervisor, you are the epitome of a mentor. From you I learned too many things to enumerate. But two things stand out the most. Caring about other people is a superpower too often overlooked. How to receive and provide critical feedback that is actionable. Every conversation with you made me a better student and a better teacher. To my committee members, Kim and Helen. Thank you for your support and feedback throughout the process. To Holly, thank you for all the stats chats! Each of you enriched experiences.

Nasim, I am struggling to pick the novel that best represents the friendship we built but, I know that Frodo's and Samwise's journey to Mordor best represents our path during graduate studies. You provided solace when I needed it, and you encouraged me every mile of the way. Anne, whenever I needed a little bit of sunshine, you were at the ready like the break of dawn. Pierre, my best friend. Everyone deserves a cheerleader like you. At every moment throughout my studies, you were right there with pompoms, glitter, and the strength and stability of the base. Thank you for helping me fly, my friends!

It was challenging, and I loved it.

Contribution of Authors

The first author of the two manuscripts is Monyka L. Rodrigues, who created the materials, recruited the participants, collected and coded the data, ran the analyses, and wrote the contents within. The data for Experiment 2 was collected by Shiqiang Eden Jing and Experiment 4 was collected by Stephanie MacKinnon and Zoe Stern. The second author of each manuscript is Sandra Martin-Chang, who provided guidance at all stages of the research program. This included refining research questions and hypotheses, lending expertise in interpreting the findings, and the revision of the manuscripts.

Table of Contents

LIST OF TABLES	X
LIST OF FIGURES.....	XI
GENERAL INTRODUCTION	1
<i>Educational Problem.....</i>	<i>1</i>
<i>Self-teaching in and out of Context.....</i>	<i>5</i>
<i>Outcome Measures.....</i>	<i>8</i>
<i>Predictor Variables.....</i>	<i>14</i>
GENERAL RESEARCH AIMS.....	19
SUMMARY OF METHODS.....	21
WHEN SIZE MATTERS: EFFECT OF CONTEXT AND SET SIZE ON ORTHOGRAPHIC LEARNING ..	24
ABSTRACT	25
INTRODUCTION	26
EXPERIMENT 1	31
<i>Method.....</i>	<i>32</i>
<i>Results and Discussion.....</i>	<i>36</i>
EXPERIMENT 2	41
<i>Method.....</i>	<i>43</i>
<i>Results and Discussion.....</i>	<i>45</i>
EXPERIMENT 3	53
<i>Method.....</i>	<i>54</i>
EXPERIMENT 4	59
<i>Method.....</i>	<i>60</i>
<i>Results and Discussion.....</i>	<i>61</i>
GENERAL DISCUSSION.....	63
<i>Regular Words and Contextual Facilitation during Training and Retention</i>	<i>64</i>

<i>Stimulus Set Size</i>	67
<i>Transfer to Spelling</i>	68
<i>Limitations and Future Directions</i>	69
<i>Conclusion and Implications for Practice</i>	70
CONCEPTUAL REPLICATION WITH CHILDREN PARTICIPANTS	72
DIFFERENTIAL EFFECTS OF CONTEXT AND ISOLATION TRAINING ON ORTHOGRAPHIC LEARNING	74
ABSTRACT	75
DIFFERENTIAL EFFECTS OF CONTEXT AND ISOLATION TRAINING ON ORTHOGRAPHIC LEARNING	76
<i>Orthographic Learning in and out of Context</i>	77
<i>The Present Study</i>	85
EXPERIMENT 5	86
<i>Method</i>	86
<i>Results and Discussion</i>	90
EXPERIMENT 6	98
<i>Method</i>	98
<i>Results and Discussion</i>	101
GENERAL DISCUSSION.....	106
<i>Limitations and Future Directions</i>	111
<i>Conclusion and Implications for Practice</i>	112
DISSERTATION DISCUSSION	114
<i>Theoretical Implications</i>	115
<i>Practical Implications</i>	125
<i>Limitations and Future Directions</i>	130

<i>Conclusion</i>	131
REFERENCES	133
APPENDIX A	154
APPENDIX B	157
APPENDIX C	159
APPENDIX D	160
APPENDIX E	161
APPENDIX F	162
APPENDIX G	165
APPENDIX H	169
APPENDIX I	172
APPENDIX J	176
APPENDIX K	179
APPENDIX L	185
APPENDIX M	186
APPENDIX N	190
APPENDIX O	192
APPENDIX P	194
APPENDIX Q	197

List of Tables

Table 1 33

Table 2 39

Table 3 44

Table 4 49

Table 5 87

Table 6 92

Table 7 104

List of Figures

Figure 1.....	37
Figure 2.....	47
Figure 3.....	51
Figure 4.....	56
Figure 5.....	57
Figure 6.....	62
Figure 7.....	91
Figure 8.....	95
Figure 9.....	97
Figure 10.....	103
Figure 11.....	106

General Introduction

Educational Problem

The Right to Read inquiry report, released February 2022 by the Ontario Human Rights Commission, demanded an end to guessing words from context. Instead, they insisted that students receive phonics instruction to develop decoding skills. Decoding skills have been recognized for at least 20 years as foundational for the development of strong reading skills (e.g., Gough & Juel 1991; Ehri, 1991, 1994; Ehri et al., 2001; Stanovich, 1986; Rayner et al., 2001). Nonetheless, the report has renewed interest in the *reading wars*, which refers to a long-standing and contentious debate in the field of education regarding the best method for teaching reading (Castles et al., 2018; Kim, 2008; Tunmer, & Chapman, 2002). The dispute centers around two main approaches: whole-language and phonics-based instruction.

The whole language movement suggests that children can learn how to read through prediction by actively generating guesses about words based on context cues and initial letters of words (Goodman, 1967). They argue that reading instruction should be modeled on language acquisition, focusing on meaning making, and not on the printed letter and letter combinations. Thus, whole-language instruction promotes focusing on comprehension and engaging students with authentic literature. Supporters argue that children naturally acquire reading skills through exposure to meaningful texts and that explicit phonics instruction is unnecessary.

An offshoot of the whole-language movement is the three-cueing approach (also known as the searchlight model) which is pervasive in the Canadian school system (Ontario Human Rights Commission, 2022). Young readers are taught to use semantic, syntactic, and letter-sound cues simultaneously to guess unfamiliar words. Thus, children today are still taught to use cues from context to monitor for meaningfulness and guess unfamiliar words, and to only make corrections when comprehension breaks down. However, relying on guessing to identify unfamiliar words is unlikely to result in reading development because in connected text, the

predictability of content words is roughly 10% and 40% for function words (e.g., a, an, to; Gough, 1983; Pollatsek et al., 2006; Rayner et al., 2013). Tunmer & Chapman (2002) expertly stated: “The meaning of text therefore depends disproportionately on the meanings of its least familiar and predictable words.” (p. 344). Thus, it is not a feasible or scalable strategy.

In comparison, phonics-based instruction emphasizes systematic teaching of letter-sound correspondences and decoding skills (Ehri, 2020; Castles et al., 2018). For example, students learn to analyze letter-sound mappings for short common words. Phonics is shown to support reading fluency and comprehension because it helps children decode unfamiliar words (Ehri et al., 2001; Castles et al., 2008; Perfetti & Hart, 2002). In fact, through a computation analysis, Solity and Vousden (2009) demonstrated that if children learn 64 of the most common letter-sound mappings and the first 100 most frequent words (often irregular), they can read 90% of words they will encounter in a school setting. Once children are familiar with some of the letter-sound mappings, they use their knowledge to build their written lexicons through decoding (reviewed in detail below; Ehri, 1992, 1995, 2005, 2020; Share, 1995, 1999, 2011). Indeed, decades of research demonstrates that the practice of mapping sounds onto letters result in positive learning trials (i.e., correct word identifications), which in turn helps connect the orthographic and phonological representations with the word’s meaning in memory (e.g., Ehri, 1991, 1992, 1995, 2005, 2020; Perfetti 1992; Perfetti & Hart, 2002; Share, 1995, 1999, 2011). Evidence also suggests that phonics is beneficial for struggling readers (e.g., Ehri, 2020). With practice, then, these connected representations support rapid and efficient word reading, which in turn frees up cognitive resources for comprehension mechanisms. Castles and colleagues (2018) rightly suggest that:

“...the relationship between letters and sounds is necessary and nonnegotiable when learning to read in alphabetic writing systems and that this is most successfully achieved through phonics instruction. Yet reading scientists,

teachers, and the public know that reading involves more than alphabetic skills.

To become confident, successful readers, children need to learn to recognize words and compute their meanings rapidly..." (p. 6)

The science of reading community (for review, see Goodwin & Jiménez, 2020) promotes teaching decoding skills to reach the end game of students independently reading in context. Facilitating reading acquisition early on not only promotes skill development but also reading for pleasure. Indeed, seminal work by Stanovich (1986), and since corroborated (van Bergen et al., 2021), suggests that strong early reading skills trigger early positive literacy experiences, which creates a beneficial feedback loop wherein reading enjoyment fosters reading skills in later grades. Moreover, reading meaningful texts have far reaching positive impact on vocabulary, verbal fluency, comprehension, spelling skills as well as cultural and declarative knowledge (Allington, & McGill-Franzen, 2021; Grolig et al., 2020; Mar & Rain, 2015; Martin-Chang et al., 2021; Rodrigues et al., 2022; Tremblay et al., 2020; van Bergen et al., 2021).

In addition to the accruing benefits of reading for pleasure, reading in context has in-the-moment benefits. Several researchers have demonstrated that reading in context helps to disambiguate words that are difficult to read (Nation & Snowling, 1998; Perfetti et al., 1979; Stanovich, et al., 1985; Stanovich, West, & Feeman, 1981; Tunmer & Chapman, 1998). Seeing a word in the context of a sentence facilitates its identification by reducing the number of possible words that fit into the semantic field provided by the text. This is not guessing; under natural circumstances, words in literary texts are not easily predictable (Pollatsek et al., 2006; Rayner et al., 2013). What is more, reading researchers do not endorse the use of guessing from context for word identification or as a strategy for developing reading skills (Share & Stanovich, 1995; Stanovich, 1984). Instead, it is understood that when decoding is slow and effortful, the information extracted from context is combined with decoding attempts (Castle et

al., 2018; Kim, 2008; Share & Stanovich, 1995; Tunmer & Chapman, 1998, 2002). This cognitive process is called contextual facilitation (Stanovich et al., 1981).

The effect of context is pervasive. Compared to a no-context condition, word reading accuracy is better when the target word is embedded in meaningful text. It has been suggested that “difficulties in one domain (e.g., phonology) might be compensated for by adequate or precocious skills in another (e.g., semantics)” (Nation & Snowling, 1998, p. 996). The supportive mechanism of context occurs for adults encountering low frequency or challenging words and is a typical situation for children that are honing their reading skills (Perfetti & Roth, 1981; Share & Stanovich, 1995; Stanovich, 1984; Stanovich & West, 1979, 1981; Rayner, 1998; 2009; Rayner et al., 2006; West et al., 1983). That is to say, when bottom-up decoding processes are insufficient for word identification, top-down processes from context are incidentally recruited to compensate. Several works have demonstrated the shortcomings of the whole language approach (for review, see Castles et al., 2018; Kim, 2008).

In fact, contextual facilitation was in part coined because one of the main predictions made by the whole language theorists was falsified. Goodman (1967) said that skilled reading chiefly relies on generating guesses from the syntactic and semantic redundancies of language. As such they argued that novice and struggling readers would be less able to make use of contextual information. However, researchers who compared good and poor decoders consistently demonstrate that those who were less skilled showed more facilitation than those with stronger reading skills (e.g., Landi et al., 2006; Nation & Snowling, 1998; Tunmer & Chapman, 1998). What is more, both Perfetti (Perfetti et al., 1979; Perfetti & Roth, 1981) and Stanovich (Stanovich, 1984; Stanovich & West, 1981, 1983; Stanovich et al., 1981) reported data demonstrating that “words and individuals are interchangeable” (Stanovich, 1986, p. 370); suggesting that when reading is experimentally made difficult, skilled readers also manifest a reliance on context for word identification.

What is open for debate is whether the in-the-moment facilitation effect generated from context promotes the formation of lexical representations in memory to support subsequent reading and spelling. Here, we present six training studies that evaluated orthographic learning. We exposed adults (Study 1) and children (Study 2) to words in and out of context, and measured word learning by assessing concurrent reading and delayed reading and spelling accuracy.

Self-teaching in and out of Context

In alphabetic orthographies, spoken words can be segmented into individual phonemes. They are the smallest unit of spoken language. Graphemes are the written equivalent of phonemes. Graphemes are the smallest unit of the written language. In the English language, a grapheme is sometimes comprised of a single letter (e.g., *a* book) and sometimes as a string of letters (e.g., *ray*, *vein*, *wait*, *great*, and *eight*). Decoding thus, is the act of translating graphemes into phonemes (i.e., letters into speech sounds). Decades of evidence demonstrates that children build their written lexicons through decoding (Ehri, 1992, 1995, 2005, 2020; Share, 1995, 1999, 2011). Indeed, the self-teaching hypothesis (Share, 1995) proposes that translating graphemes into phonemes serves as a reliable mechanism for developing lexical representations in memory. Lexical representations can be graphemes, syllables, or whole words that include information about the pronunciation (phonology), meaning (semantics), and spelling pattern (orthography) stored in long-term memory (Perfetti & Hart, 2002).

When translating print to sound, the reader's attention is focused on the letter-order and the relation between graphemes and phonemes within words. Each time word identification succeeds through decoding, the reader experiences a positive training trial; creating a feedback situation in which the reader associates the visual representation of a word (i.e., orthography) with its sound (i.e., phonology) and meaning(s) (i.e., semantics; Share, 1995). In this view, decoding operates as a self-teaching mechanism that results in orthographic learning

(Cunningham et al., 2002; Schwartz et al., 2014; Shahar-Yames & Share, 2008; Share, 1999, 2004). In turn, orthographic learning develops lexical representations in memory that enable fluent, automatic reading (Castles & Nation, 2006; Perfetti, 1992). Once children establish this self-teaching mechanism, they are well positioned to garner a host of positive training trials through the combination of oral vocabulary, decoding skills, and exposure to printed text (Share & Stanovich, 1995; J. N. Taylor & Perfetti, 2016).

To be clear, when reading in or out of context, the primary strategy for identifying words is decoding; however, in context, a secondary, compensatory strategy is available (Martin-Chang, 2017; Nation & Snowling, 1998; Tunmer & Chapman, 1998, 2002). For example, when a reader comes across the word chameleon for the first time in print, they might decode it as /čæ-mɛ-li-on/, then crosscheck that pronunciation against the contextual information (perhaps a text about lizards or a story about a boy who often changes his mind to make friends), which triggers access to the familiar spoken word that sounds like /kə-mi-liən/. Conversely, when being faced with chameleon out of context (e.g., part of a spelling list to be studied, the name of a store, or street name), word reading strategies are restricted to decoding. Without facilitation from context, the probability of deriving the correct pronunciation can be limited.

Outside of explicit reading and spelling instruction, the majority of reading practice happens in context (e.g., argumentative, expository, narrative, descriptive texts). As young children become literate, they receive basal readers from their teachers to read in the classroom and at home (Foorman et al., 2004; Juel, 2006). Older children are mainly exposed to print through longer texts found in non-fiction books and fiction novels (Guthrie et al., 2012; Kim et al., 2017; Locher et al., 2019). Similarly, adults consume the majority of print through reading for pleasure with a smaller portion through reading for school or work (Locher & Pfof, 2019). Because so much time spent reading is spent in context, it is worthwhile to investigate the impact of the effect of context on orthographic learning.

As such, contextual facilitation studies evolved into training studies, moving beyond measuring the effect of context on word identification to assessing the effects of context training on word acquisition. Still, facilitation is defined as the difference in word reading speed or accuracy between a context and an isolated word reading condition. For example, in the context condition, participants practice reading a set of target words embedded in sentences or passages. In the isolation condition, participants read another set of target words in a list. They are asked to read the targets as quickly and accurately as possible. Then, various outcome measures, such as word reading, spelling recognition, and spelling accuracy are compared across training conditions. The advantage of training studies is the possibility of testing causal links between training in and out of context and the formation of lexical representations in memory.

Since the seminal research of Stanovich et al. (1981), there is wide consensus that context bootstraps decoding as it happens (Ehri, 2014; Landi, 2013; Rayner, 2012). However, the impact on word learning is less certain. Some researchers have hypothesized that because reading in context increases the number of opportunities for mapping word phonology and semantics onto orthography (Martin-Chang, 2017; Nation & Snowling, 1998; Tunmer & Chapman, 2002), that this success should then translate to stored lexical representations. Whereas other researchers have argued that when decoding is partially supplanted by context, it detracts attention from a word's orthographic form resulting in "a paradoxical relation between reading and learning" (Land et al., 2006, p. 1), where children are *reading* more words, but *learning* fewer of them. They further postulated that reading out of context is ideal for creating lasting lexical representations in memory because it is more likely to recruit effortful bottom-up processes (Landi, 2013; Landi et al., 2006).

At present, there is no consensus in the field. However, I would argue that both hypotheses are substantive and may not in fact be contradicting. The novel contribution of my

program of research is to examine whether these seemingly incongruent interpretations may be explained by methodological differences across studies. I evaluate the literature as well as and extend the research by contrasting key variables.

Researchers have implemented a range of training methods, exposing adults and children to orthographic forms with and without story context (or semantic information through definitions). The type of words trained have often differed from one study to another. Participants have trained with real words and nonwords. These items have been written using regular and exception spelling patterns, as well as with standard and artificial orthographies. Furthermore, across studies a variety of measures of learning have been used. These range from reading in context, reading in isolation, to spelling recognition and spelling production. In short, the lack of consistency makes it difficult to draw conclusions across studies. And yet, as was argued above, much of the reading done in and outside of the classroom occurs in context, therefore, it is of practical and theoretical interest to know the boundaries of contextual facilitation on the development of orthographic knowledge.

Outcome Measures

Transfer to Context

As outlined above, during training readers generally fare better in context. What remains unclear is whether training in context fosters word learning beyond training; said differently, will gains made during training *transfer* to novel reading situations. The choice of outcome measures seems to greatly impact the interpretation of the utility of contextual facilitation for orthographic learning. Over a series of training studies, Martin-Chang and colleagues contrasted different outcomes measures (Martin-Chang & Levy, 2005, 2006; Martin-Chang et al., 2007; Martin-Chang, 2017). Children in second and fourth grade practiced reading 170 words (Martin-Chang & Levy, 2005). Eighty-five words were trained in isolation and the other half were trained in short stories. Children were exposed to each target word 12 times over six

trials. If a word was misread, whole-word corrective feedback was provided, immediately pairing orthography with whole-word phonology. During training, performance for children in Grade 4 was nearly at ceiling however, children in Grade 2 read more words accurately in context relative to isolation. One day following the last training session, children read new passages that contained the trained words, but in novel short stories. Results revealed that words initially trained in context were read with better accuracy than words originally trained in isolation (Martin-Chang & Levy, 2005). Note that because all target words were presented in a novel context during retention, contextual facilitation was comparable for trained target words. Yet, participants did not read words trained in isolation as well as those trained in context. This indicates that the gains made during context training were relatively stable when transferring to a novel context. It also indicates that contextual facilitation did not sufficiently improve the performance for words trained in isolation relative to those trained in context (see also Martin-Chang et al., 2007; Martin-Chang, 2017).

Transfer to Isolation

Landi and colleagues' (2006) worked with younger children in Grades 1 and 2 and assessed retention in isolation. Children were asked to read a list of words and any that could not be read accurately were used to create individualized training sets ($M = 33$ words). One week later, half of the words were read once in isolation and the other half were read once in predictable two-sentence passages (adult mean cloze probability was 67%). During training, children accurately read almost three times the number of target words in context compared to isolation, with performance near ceiling in context. A week later, when all target words were presented in isolation during retention, word reading accuracy was similar across the training conditions. However, participants exhibited a noticeable drop in performance for the words originally trained in context. This was made obvious when scores were conditionalized. The number of correctly read words at retention were divided by the number of correctly read words

during training. The conditionalized scores were higher for words trained in isolation. The authors concluded that although children's in-the-moment performance was best when reading in context, training words in isolation was more likely to foster lexical representations that could be accessed at a later time to support fluent reading out of context (Landi et al., 2006). In a second experiment, each target word was read three times during training instead of once. Yet, the overall pattern of results remained the same.

These results echo Martin-Chang & Levy (2006) who also assessed word reading accuracy during retention in isolation. Children in Grade 3 practiced reading 96 target words (48 in isolation and another 48 in context) 12 times over six trials. As expected, children read more words accurately during context training than in isolation. By the last trial, children's word reading accuracy was at ceiling in the context condition, creating a significant advance over words trained in isolation. Here as well, words trained in isolation were read with similar accuracy as words trained in context during retention. At the same time, words previously trained in isolation suffered less loss than those trained in context. Thus, similar to Landi et al., 2006, this pattern of results suggests that words trained in isolation were better remembered (also see Martin-Chang et al., 2007).

Transferring to list rather than to a novel context holds substantial merit, both experimentally and practically. The premise behind engaging in reading practice is to graduate words from low to high quality representations. Doing so permits readers to access words under any circumstance (Perfetti & Hart, 2002; J. N. Taylor & Perfetti, 2016). With time and effort, words are fluently read without resorting to laborious decoding or requiring the support from context. As such, a post-test retention task involving reading in isolation (rather than in context) is a more sensitive measure of lexical quality. Reading without the support from context means the reader can only rely on their vocabulary and decoding skills. Success reading in isolation

attests to lexical representations being good enough to support fluent reading. A milestone in of itself.

Transfer to Spelling

Reading and spelling are not equally challenging skills to acquire (Conrad, 2008; Ouellette, 2010; Ouellette & Sénéchal, 2008). Often, lower quality lexical representations are good enough for reading, especially with the support from context. However, spelling relies on more precise representations in memory (Martin-Chang et al., 2014; Ouellette et al., 2017; Perfetti & Hart, 2002; Rossi et al., 2019; J. N. Taylor & Perfetti, 2016).

In essence, a word that is completely unfamiliar does not have a lexical representation in memory (Perfetti & Hart, 2002; J. N. Taylor & Perfetti, 2016). A word with low lexical quality reflects weaknesses in any of the lexical components (orthographic, phonological, or semantic) or between the connections of these components. Finally, a word with high lexical quality represents strong connections between precise orthographic, phonological, and semantic representations of the word in memory. High quality orthographic representations, then, reflect “specificity and completeness of the information stored in lexical memory about both the identity and order of the letters defining a particular word, which ensures that a particular written word directly activates its lexical representation with little interference from other similar words” (Veldre & Andrews, 2014, p. 705). For example, seeing the word *decent* and having immediate access to its phonology and semantics without any or very little interference from its orthographic neighbour *descent* reflects high quality lexical representation in memory. Orthographic knowledge is hypothesized to result from continually refining and retrieving word-specific representations from memory on a word-by-word basis (Luke et al., 2015; Martin-Chang et al., 2014; Perfetti & Hart, 2002; Perfetti, 1992, 2007; J. N. Taylor & Perfetti, 2016).

Orthographic learning can be conceptualized as an extension of learning letter–sound mappings (Ehri, 2014). With practice, graphemes, strings of graphemes, and whole word

representations are bonded to their phonology and semantics allowing fluent, accurate reading to proceed (Ehri, 1992, 2005). In this way, decoding and orthographic knowledge are reciprocally influenced. The quality of an orthographic representation, however, is more accurately indexed by spelling production rather than reading accuracy, mainly because accurate word reading is possible with weaker orthographic representations stored in memory—people can commonly read more words than they can spell (Luke et al., 2015; Martin-Chang et al., 2014; Perfetti, 1992, 2007). Here, we asked whether reading practice in context fosters the formation of word-specific representations to the same extent as reading in isolation.

Within the literature, evidence suggests that at least some orthographic learning takes place during passage reading (Cunningham, 2006; Cunningham et al., 2002; Share, 1999, 2004; Ricketts et al., 2011; Wang et al., 2012, 2013). When considering scores on spelling recognition and production tasks, researchers who contrasted training in and out of context found that isolation was no more effective than context training (Cunningham, 2006; Nation et al., 2007; Ouellette & Fraser, 2009; Wang et al., 2011). Thus, learning did occur, but no differences were found between context and isolation training. One study did find, however, that spelling production was best following isolation training compared to context (Martin-Chang et al., 2017).

Two tasks are often implemented to measure orthographic learning, an orthographic choice task and a spelling production task (Share, 1999, 2004). The orthographic *choice* task assesses participants ability to recognize correct spellings of trained words and reject misspellings. For example, the trained target nonword *wote* is presented among three foils, one is usually a homophone (e.g., *woat*), while the other two are visually similar distractors (e.g., *wode*, *woet*; Ouellette & Fraser, 2009). A variation of this task, the orthographic *decision* task, however, is said to be more sensitive. For the decision task, targets and misspelled foils are randomly presented one at a time, and participants are asked whether they recognize the

presented item from training. An even more difficult task that also measures orthographic learning is spelling production; participants are asked to spell a list of targets dictated by an experimenter.

Within this framework, Cunningham (2006) trained real words that were likely to be contained in the oral vocabulary of children in the first grade, but not yet consolidated with their orthographic representations. Children practiced eight targets, four in short passages and another four in scrambled passages (thereby eliminating both semantic and syntactic cues) and, each target word was presented six times in one trial. As expected, children read more words accurately with contextual support than without (Cunningham, 2006). Later, children correctly identified and spelled a large proportion of trained target words, but performance was similar across training conditions.

Martin-Chang et al. (2017) proceeded with a similar methodology. Children in the second grade practiced 50 target words, 25 were presented in context and another 25 in isolation, with 12 exposures over six trials. During training, children read more words accurately with the support of context than without. When spelling production was considered however, children spelled more word correctly that were initially trained in isolation compared to context. Notably, Cunningham (2006) implemented a small stimulus set size, while Martin-Chang et al. (2017) used six times as many words and longer passages. These differences challenge whether their results can be directly compared.

Nation and colleagues (2007) trained nine nonwords with children in second and third grade. Half of the children practiced reading the nonwords in short stories whereas, the other half practiced in isolation by sorting flashcards into a word and nonword pile. The number of exposures was manipulated, so that targets items were seen either once, twice, or four times. Across all conditions, children correctly decoded three-quarters of the nonwords and, the authors did not find any differences between the training conditions on the orthographic choice

task. In another study, Ouellette and Fraser (2009) asked children in Grade 4 to practice 10 nonwords. In contrast to Nation et al. (2007), the correct pronunciation was provided during an initial exposure phase. Children repeated each nonword aloud after the experimenter. During this exposure phase, children heard a short description for five of the nonwords accompanied by drawings and the other five nonwords were presented in isolation (without semantic information). During training, each nonword was presented six times in print and participants were asked to read the nonwords aloud. The findings revealed a facilitatory effect of semantic information on orthographic choice task, but not for spelling production. Although spelling is a more challenging undertaking than recognizing a nonword, the findings of orthographic choice do not align with those of Nation et al. (2007). However, there is perhaps an important methodological difference between the two studies. Nation et al. (2007) did not teach the pronunciation of the nonwords to the children, therefore when children encountered the nonwords, in or out of context, they had to rely on their decoding skills without the support of previous knowledge of the word's phonology. Furthermore, even if children in the context condition deduced the meaning of the target nonwords, the semantic knowledge from the context could not supplement partial decoding attempt—decoding can only be supplemented by pre-existing knowledge (Nation et al., 2007).

Predictor Variables

Ostensibly, the goal of reading acquisition is to move from slow, methodical decoding to fluent, accurate reading. As more words are successfully decoded, the laborious translation of graphemes-to-phonemes is supplanted by high-quality lexical representations that are autonomous from context (Ehri, 1992, 2005, 2014; Perfetti, 1992, Share, 1995). These representations are well-specified in memory, and redundant with pronunciation and meaning. Redundancy here means that when exposed to the phonology of a word, access to its semantics and orthography is immediate. When exposed to the semantics of a word, access to

its phonology and orthographic is immediate; the same logic applies when exposed to the orthography. It provides immediate access to the word's phonology and semantics. However, to achieve this end goal, readers must train a fluent and flexible reading system that is sensitive to orthographic regularities and yet can also cope with the many exceptions that exist within the English language spelling system (Share, 2008; Zeigler & Goswami, 2005). Moving individual words from unfamiliar to familiar, and in parallel, progressing from novice to expert reader, necessarily results from direct experience with words—an experience that is largely heterogenous across words and readers. "...at any particular point in time, a child may be reading some words slowly, effortfully, and with heavy reliance on context and alphabetic decoding, while reading other words rapidly, automatically, and with minimal contextual influence" (Castles & Nation, 2006, p. 152). This lexical-level knowledge differs between readers (Andrews & Bond, 2008; Hersch & Andrews, 2012; Kuperman & Van Dyke, 2011; J. N. Taylor & Perfetti, 2016).

As such, another methodological consideration that impacts the interpretation of whether training in context fosters word learning is the choice of stimuli (Nation & Snowling, 1998; J. S. H. Taylor, et al., 2011). Lexical-level factors, such as word regularity and number of items to be learned, can potentially influence both reading and spelling performance. Accordingly, the experiments presented here include two factors that may moderate the impact of contextual facilitation on forming lexical representations. They are word regularity and number of words to be learned (i.e., stimulus set size).

Word Regularity

Regular words are pronounced according to accepted letter-to-sound correspondence (e.g., mint, tint, lint). These examples are deemed regular because the short vowel sound is expected in a closed syllable. By contrast, irregular words contain atypical or less frequent letter-to-sound mappings, and thus do not rhyme with words that share the same orthographic

rime (e.g., a *pint* of beer; Rastle & Coltheart, 1999). Nation and Snowling (1998) posited that reading accuracy of irregular words would benefit more from top-down contextual cues than regular words because decoding can be insufficient to arrive at the correct pronunciation. Contextual information can tip the scales towards correct word pronunciation by activating semantically related words in memory and pairing these with the partial decoding attempt. Confirming their hypothesis, the authors found that children in the third and fourth grade read more words accurately in context than in isolation, and that contextual facilitation was greatest for irregular words (Nation & Snowling, 1998). Once again, the question becomes, does this greater effect of context for irregular words translate to reading and spelling words at a later time, namely, word acquisition. On the one hand, more success at reading tricky, irregular words in context could lead to better long-term gains because of the increased probability of pairing word phonology with orthography. On the other hand, leaning on context could divert cognitive resources away from processing a word's orthography thus, leading to weaker consolidation in memory.

In a training study, Wang and colleagues (2011) exposed children in second grade to four regular and four irregular nonwords. For example, the nonword *praif* was taught as either having a regular pronunciation (rhyming with *safe*) or an irregular pronunciation (rhyming with *laugh*). Prior to training, children learned the meaning and either the regular or irregular pronunciation of eight nonwords. This provided access to both the phonetic and semantic information prior to training. Thus, the nonwords were part of the children's oral vocabulary—as would be the case for real words that children could use orally, but not yet read or spell. During context and isolation training, children were exposed to each nonword four times. The authors found that children read more regular nonwords in context on the first trial, but that differences with isolation tapered off at the second training trial. Conversely, children consistently read more irregular nonwords in context than in isolation. In fact, without the support of context children

tended to produce regularized pronunciations of the irregular nonwords. Despite the added phonological instruction, the authors corroborated findings by Nation et al. (2007). When children were asked to identify correctly spelled targets amongst foils or spell the words themselves, performance was similar after both context and isolation training (Wang et al., 2011).

However, Wang and her colleagues also implemented a decision task (Wang et al., 2011). Instead of a forced choice among one target and three foils, each nonword (including foils) were randomized and presented individually. This orthographic decision task detected an effect of context for irregular nonwords. Perhaps making discrete judgments for each item increased the tasks sensitivity to weaker lexical representations that formed during training. The authors argued that contextual information facilitated orthographic learning, but only when children could not rely on regular spelling-sound mappings. However, given that there were only four words per condition, and that by the second trial, children's word reading accuracy was at ceiling for regular nonwords, the effect of context might have been obscured.

In a similar design, but working with adults, Taylor and colleagues (2011) implemented an artificial orthography and detected an ancillary effect of semantic information for orthographic learning. Adults were exposed to 36 pseudowords written with a series of nonalphabetic symbols and either practiced their pronunciations in isolation or practiced pronunciations paired with definitions. Furthermore, pseudowords varied in terms of vowel regularity and frequency. As anticipated, by the end of training, both high-frequency and regularized items were read accurately more often than low-frequency and irregular items. Of interest, when pseudowords were accompanied by semantic information, reading accuracy was boosted for the challenging irregular and low frequency (relative to the phonology only condition; J. S. H. Taylor et al., 2011). In an old-new task that is reminiscent of Wang et al. (2011), pairing pronunciations with definitions improved adults' ability to discriminate target items against foils. However, the

authors collapsed across regularity, so it is unknown whether the added semantic information moderated orthographic learning to a similar extent for regular and irregular items. Unanswered is whether these same effects would be observed for real words read in a naturalistic passage-level context.

Stimulus Set Size

Another methodological difference between the studies reviewed here is the number of words participants trained in and out of context. The majority of studies ensured that the same number of target words were trained in both context and isolation conditions. However, stimulus set size may tangibly influence task difficulty during training, and as a result moderate the formation of lexical representations in memory. Indeed, memory research suggests that the number of items to be learned affects participants' chances of retaining and retrieving chunks of information (Cowan & Alloway, 2009). Therefore, it is plausible that training a large number of words would make the task more difficult thereby attenuating learning.

It stands to reason that there may be a difference between learning to read 10 or fewer words and 80 or more words. Previous work demonstrates that context is drawn upon when reading is hindered (Stanovich, 1984; Stanovich & West, 1981; Stanovich et al., 1981). Challenges may arise from a variety of sources, including long, low-frequency words (Perfetti & Roth, 1981; Stanovich & West, 1981), irregular words (Nation & Snowling, 1998), degraded text (Stanovich & West, 1979), and weak reading skills (Nicholson, 1991; Stanovich et al., 1981, 1984). If training a large set of words increases task difficulty and increasing difficulty augments reliance on context, it follows that the number of words to be learned might impact orthographic learning.

At the same time, print exposure—which is understood to be the amount read over the lifetime—is not only associated with reading achievement, overall academic success (Allington & McGill-Franzen, 2021) but also, with better spelling skills (Stanovich & West, 1989). For

example, a correlational study by Cunningham and Stanovich (1990) demonstrated that print exposure was related to the development of children's orthographic knowledge. They assessed non-verbal intelligence, word identification, phonological awareness, and spelling recognition, as well as print exposure of third- and fourth-grade children. Orthographic processing was positively related to word identification. What is more, print exposure was uniquely related to orthographic processing. The authors stipulated that the more volume a child reads, the more orthographic representations are strengthened and refined, and ultimately, the more resources are made available to support reading comprehension.

The big question then is, can the literature on print exposure be reconciled with the findings that suggest that reading in context is not ideal for creating lexical representations in memory. The mixed findings regarding the benefits of contextual facilitation for word learning likely extends beyond outcome measures and in part be attributed to the nature and number of words to be learned. Thus, a factor that may impact orthographic learning is stimulus set size. This can be tested by experimentally manipulating print exposure; operationalized here as the number of unique words to be learned by participants. Yet, this methodological issue has only once been discussed as a potential moderator and has never been directly tested (Martin-Chang, 2017).

General Research Aims

To summarize, some researchers argue that effortful bottom-up processes are ideal for creating word-specific orthographic representations (Landi, 2013; Landi et al., 2006). When top-down resources cannot be recruited from context, the readers' attention is more likely to be focused on the individual graphemes, evaluating how to blend the sounds to arrive at a reasonable pronunciation. Whereas others claim that a combination of bottom-up and top-down processes is a more likely candidate for orthographic learning because reading in context

increases the likelihood of positive learning trials (Martin-Chang, 2017; Nation & Snowling, 1998; Tunmer & Chapman, 2002). Both theoretical perspectives are plausible.

On the one hand, if top-down support from naturalistic texts provides the necessary cues to supplement decoding attempts, then readers should be exposed to accurate word pronunciations more often in context than in isolation (Martin-Chang, 2017; for review, see Share, 1995, 2011). Moreover, if successful decoding attempts bind phonology, semantics, and orthography, then reading in context may be more beneficial than reading in isolation. That is, contextual facilitation should increase instances of self-teaching (Martin-Chang, 2017). On the other hand, as Landi and colleagues (2006) postulated, if bottom-up decoding is supplanted by reading in context because it detracts attention from word orthography, resulting in weaker lexical representations in memory. Landi (2013) further hypothesizes that refinements in lexical representations are predicated upon “the association between readers’ generated phonology and the letters that are being attended to simultaneously” (p. 24). In their account, the self-teaching process likely operates when reading in either context or isolation, but they suggest that its impact may be attenuated in context (Landi, 2013; Landi et al., 2006).

Accordingly, Wang et al. (2011) expected improvements in orthographic learning as a result of training nonwords in isolation. They further speculated a negative effect of context for learning the trickier, irregular words, because contextual facilitation would divert even more cognitive resources away from processing word-level orthography. This was not the case. In fact, context boosted reading accuracy for only irregular words and orthographic learning was similar across both conditions with a small advantage for context when implementing an old-new spelling recognition task. Whereas others have reported the benefits of context for regular words as well as a mix of word types (Cunningham, 2006; Cunningham et al., 2002; Martin-Chang, 2017; Nation & Snowling, 1998; J. S. H. Taylor et al., 2011). A key methodological difference of the studies listed above was the number of words to be learned (i.e., stimulus set

size). Furthermore, in terms of outcome measures, when evaluating spelling production, Martin-Chang et al. (2017) reported that training in isolation enabled children to spell more words correctly compared to training in context. Therefore, it remains unclear whether contextual facilitation is moderated by word regularity, stimulus set size or both.

In Experiments 1 to 4 we asked three research questions. (1) Does contextual facilitation occur in a regularized orthography (Moats, 2003)? (2) If contextual facilitation does occur, what is the impact on word acquisition as measured by word reading accuracy during retention, spelling recognition, and spelling production? (3) Does stimulus set size impact orthographic learning in and out of context. In Experiments 5 and 6, we extended our research with adults to children. We investigated whether contextual facilitation would be observed for both regular and exception words and if contextual facilitation did occur, its impact on spelling recognition and spelling production. In Experiment 6, we also examined whether stimulus set size would moderate orthographic learning in and out of context.

Summary of Methods

Over six experiments, 178 participants were recruited. Study 1 included 114 English speaking adults. Study 2 aimed to replicate the findings with novice readers, and thus included 64 children in Grades 1 and 2. The adult data (Study 1) was collected online over three sessions (Experiment 2 included one in-person session). The child data (Study 2) was collected on school grounds over five sessions.

The experiments received certification of ethical acceptability for research involving human subjects by the Concordia University Human Research Ethics Committee (Certification Numbers: 30014272 & 30010794). On all occasions, consent was obtained from adult participants and parents of children (see Appendices A and B, respectively). Moreover, verbal assent was obtained from each participant before data collection (see Appendix C). Demographic information was also collected via survey (see Appendices D and E).

For Study 1, a series of word sets and stories were developed. We regularized a large set of real words by transcribing them into phonetic symbols (Moats, 2003). This permitted examining whether contextual facilitation would be observed for words presented in a regular orthography. Training phonetic symbols also exposed participants to a novel orthographic system, emulating the process of a novice reader mapping familiar phonemes onto unfamiliar graphemes. This created a situation akin to children in the partial alphabetic phase (Ehri, 1995). In this case, the target words were part of the adults' oral vocabulary, and they had partial knowledge of the orthographic system. In Experiment 1 and 2 participants trained both in context and isolation whereas, in Experiments 3 and 4 participants trained either in context or isolation. See appendices F to K for word sets and stories. Adopting this methodology, permitted addressing our first research question about contextual facilitation with regular words when reading is challenging enough. To answer the second research question, a series of outcome measures were adopted (see below). To answer the third research question, participants trained with either a small or large set size in Experiments 1 and 2. Whereas in Experiments 2 and 4, stimulus set size was a within-subject variable.

To delineate the boundaries in which orthographic learning occurred both in and out of context, measures tapping various levels of knowledge were implemented: word reading accuracy during training and retention in list, spelling recognition, and spelling production. By implementing multiple outcome measures, we were able to evaluate the continuum between orthographic knowledge that is *good enough* for reading to *precise enough* for spelling.

To assess contextual facilitation during training, participants were presented with a story and/or a list of words. For the context condition, the story included target words which appeared in phonemic notation (Moats, 2003) while, the remaining text was in the standard English alphabet. Both context stories are available in Appendices G, I and, K for Experiments 1, 2, and 4, respectively. For the isolation condition, the list of words was presented one word at a time in

phonetic notation. The word lists are available in Appendices F, H, and J for Experiments 1, 2, and 3 respectively. In each case, participants were asked to read as quickly and accurately as possible. To measure learning, the story and list were read three times during the first session and another three times during the second session (total: 6 trials). To assess whether contextual facilitation moderated word learning, reading accuracy was measured in a retention task. After a 3-day delay, participants read all trained target words randomized in a *new* list. Again, participants were asked to read as quickly and accurately as possible.

Participants also completed a spelling recognition and spelling production task. For the spelling recognition task (also known as an orthographic decision task), target words and foils were presented in a random order one at a time, and participants made old-new distinctions. Participants were instructed to respond yes, if they recognized the word and no, if they did not recognize the word. For the spelling production task, participants were asked to spell a series of trained target words in phonetic notation that had been exposed to during training.

For Study 2, a similar approach was undertaken however, because children in Grade 2 are novice readers, target words were presented in the standard English alphabet. To address the research questions, target words were manipulated through word length and regularity (Experiment 5) and set size (Experiment 6). Outcome measures were word reading accuracy during training and retention, and spelling production. The methodology was similar to that followed by adults with minor exceptions. Testing was carried out on school grounds during regular hours. Training took place 10 trials rather than 6—this was to ensure sufficient exposure for learning. See appendices L to O for word sets and stories implemented in Experiments 5 and 6 with children. Detailed descriptions of the relevant tasks and procedures are reported in the method section of each manuscript.

When Size Matters: Effect of Context and Set Size on Orthographic Learning

Monyka L. Rodrigues, Eden S. Jing, & Sandra Martin-Chang

Department of Education

Concordia University, Quebec, Canada

Part of this manuscript was submitted for publication. Within the dissertation, Experiment 4 is presented in whole rather than as a footnote. That is the only difference.

Abstract

Adults' self-taught different sets of words written in a false orthography across four experiments. Participants read under two experimental training conditions (context and isolation); half trained with a small set of 24 words while the others trained with a large set of 86 words. Training took place over six trials, followed by delayed reading and spelling post-tests. Reading in context consistently bolstered reading accuracy during and after training. In contrast, the highest spelling scores were noted following reading in isolation. The number of words to be learned (stimulus set size) moderated the effects of isolation training. Self-teaching a large set of diverse words in isolation increased reading accuracy, closing the gap with context. Overall, training in context helps establish orthographic representations that are *good enough* to support reading. Whereas, reading in isolation refines orthographic representations to be precise enough to support spelling, especially when training with a large set of words.

Keywords: artificial orthography, contextual facilitation, orthographic learning, reading, spelling

Introduction

Reading is a cognitive process that relies on rapid, accurate word identification (Nation, 2019; Perfetti, 2007). When word reading is hindered, both adults and children draw upon top-down support from context to arrive at the word's correct pronunciation (Stanovich et al., 1985). Though guessing from context is not recommended by reading researchers as a strategy for word identification, it is still the case that readers use context to supplement decoding when reading is difficult (Share & Stanovich, 1995; Tunmer & Chapman, 1998, 2002). What remains contentious is whether the support garnered from context results in orthographic learning that is "good enough" for subsequent reading and precise enough for spelling (Martin-Chang et al., 2017, p. 17). Here, orthographic learning was measured over three experiments after adults read words written with unfamiliar symbols, in and out of context.

Readers build orthographic representations on a word-by-word basis (Ehri, 1992, 2014). Share (1995) proposed that translating graphemes (letters and letter patterns) into phonemes (speech sounds) serves as a reliable self-teaching mechanism because the reader's attention is on the letter order and the relation between the letters within words. In turn, successful decoding provides the reader with positive learning trials at the word level. This creates a feedback loop in which the reader can associate a word's orthography (spelling) with its phonology (sound) and semantics (meaning). As children continue to practice, their reading moves from decoding that is plodding to word recognition that appears effortless. From a cognitive standpoint, this budding fluency is thought to result from lexical representations that are becoming fully specified in memory—a lexicalization process (Castles & Nation, 2006; J. N. Taylor & Perfetti, 2016).

There is wide consensus that context scaffolds reading when decoding is difficult, yet the impact on long-term word learning is equivocal. Some researchers hypothesize that reading in context creates more learning opportunities for amalgamating word phonology, orthography, and semantics in memory because word reading accuracy increases in context (Martin-Chang,

2017; Nation & Snowling, 1998; Tunmer & Chapman, 2002). Other researchers argue that reading in isolation relies on effortful bottom-up processes which are thought to be ideal for creating lasting orthographic representations in memory (Landi, 2013). Indeed, Landi and colleagues (2006) postulated that when decoding is partially supplanted by context, it detracts attention from a word's orthographic form resulting in "a paradoxical relation between reading and learning" (p. 1). Thus, although word reading is advantaged in context, the quick fix may come at the cost of forming well-specified lexical representations. Thus, at present, the field seems to be divided into two camps, however, these seemingly incongruent findings may be explained by methodological differences across studies.

For instance, the choice of outcome measures, such as whether participants are asked to read words in context or in isolation, or to spell words, seems to greatly impact how researchers view the utility of context. Over a series of training studies, Martin-Chang and colleagues explored different outcomes measures (Martin-Chang & Levy, 2005, 2006; Martin-Chang et al., 2007; Martin-Chang, 2017). One day following the last training session, children in Grades 2 and 4 were asked to read novel short stories containing previously trained target words (Martin-Chang & Levy, 2005). The authors found that words previously trained in context were read more accurately in a new context than words trained in isolation. Note that all words were presented in a new context during retention, thus the contextual facilitation was equated for all previously train target words. Yet, words trained in isolation were not read as well as to those trained in context. This indicates two things. First, the gains made during context training were relatively stable when transferring to a novel context. Second, the contextual information provided during retention was not enough to improve the performance for words trained in isolation.

In a follow-up study, Martin-Chang & Levy (2006) implemented the same methodology, again with children in Grade 3, except they varied the outcome measure so that it was word

reading accuracy in a new list. Once again, words trained in isolation had a lower overall success rate during training. But this time, words previously trained in isolation suffered less loss and thus were better remembered compared to words trained in context.

These results echo Landi and colleagues' (2006) work with younger children in Grades 1 and 2. They found that one week after training, when all target words were presented in isolation, children's word reading accuracy was similar across training conditions. Words trained in isolation were read well at test whereas, there was a pronounced drop in word reading accuracy for the words originally trained in context. This was made more obvious when the authors divided the number of correctly read words at retention by the number of correctly read words during training. These conditionalizing reading scores led the authors conclude that children's in-the-moment performance was better in context, but that training in isolation was more effective for developing lexical representations that support subsequent fluent reading out of context (Landi et al., 2006). This interpretation may warrant some tempering. When more words were read during the single training trial, as was the case in context, the retention score was necessarily divided by a larger number; thus, lowering the final score and creating a mathematical advantage for training isolation.

Beyond fluent, accurate word identification, forming orthographic representations that are precise enough for spelling is of value. Reading and spelling are not equally challenging skills to acquire (Conrad, 2008; Ouellette, 2010). Often, lower quality lexical representations (Perfetti & Hart, 2002; J. N. Taylor & Perfetti, 2016) are sufficient for reading—especially when cross-referenced with context—but more precise orthographic representations are required for spelling (Hersch & Andrews, 2012; Rossi et al., 2019; Veldre & Andrews, 2014). Therefore, studies that train children to read in and out context and later transfer to spelling, allow word learning to be viewed from yet another vantage point.

Interestingly, the contextual facilitation observed during training often fails to result in improved spelling performance. Thus, even though children read more words correctly with the top-down support of semantic information, they correctly spell a similar number of words after context and isolation training (Cunningham, 2006; Ouellette & Fraser, 2009; Wang et al., 2011), or spell even more words correctly from isolation training compared to context (Martin-Chang et al., 2017). Taken together, the evidence shows that training in isolation is just as good, if not better, than training in context for the lexicalization process.

Another methodological consideration that impacts the interpretation of whether training in context fosters word learning is the choice of the stimuli (Nation & Snowling, 1998; J. S. H. Taylor, et al., 2011). Lexical-level factors, such as word regularity and number of items to be learned, can potentially influence both reading and spelling performance. Words are deemed *regular* when they follow expected letter-sound correspondences (e.g., cat is a regular word because the short vowel sound is expected in closed syllables like, hat, mat, flat, splat). Conversely, *irregular words* violate expected letter-sound correspondences. For example, pint is irregular because the long vowel sound is unexpected in a closed syllable. Decoding alone is generally insufficient to arrive at the correct pronunciation for words with irregular spellings. For this reason, irregular words benefit more from top-down contextual information than regular words (Nation & Snowling, 1998) and thus, word regularity may interact with the effect of context on word learning. Once again, the hypotheses can be divided into two camps; those who predict that the success of reading irregular words in context will lead to better long-term gains by pairing phonology with orthography, and those who predict that relying on context will divert cognitive resources away from processing a word's orthography thus, leading to weaker consolidation in memory.

These two divergent hypotheses were examined over a series of studies. Wang and colleagues (2011, 2012, 2013) consistently report that children exhibit robust contextual

facilitation when they practice reading irregular nonwords in passages (e.g., broon pronounced as rhyming with run) but not for regular items. Yet, when these same children were asked to identify correctly spelled targets amongst foils or spell the words themselves, performance was similar after both context and isolation training (Wang et al., 2011). The exception to this was when the authors implemented an old-new task in which nonwords and foils were presented one at a time. Under this condition, a bootstrapping effect of context was found exclusively for words with irregular pronunciations. The authors argued that context facilitated orthographic learning, but only when children could not rely on regular spelling-sound mappings during training. In contrast, others have reported the benefits of context for regular words as well as a mix of word types (Cunningham, 2006; Cunningham et al., 2002; Martin-Chang, 2017; Nation & Snowling, 1998; J. S. H. Taylor et al., 2011). Thus, manipulating word regularity is an important factor to consider when designing training studies.

With adults, Taylor and colleagues (2011) also detected a supportive effect of semantic information on orthographic learning. In Experiment 2, they exposed 32 adults to a series of pseudowords written with an artificial orthography containing nonalphabetic symbols. Further, the vowels varied in regularity and frequency. During training, nonwords were either paired with semantic information and pronunciations or with pronunciations only. In this study, training only took place in isolation. Nonetheless, when pseudowords were accompanied by definitions (semantic information), reading accuracy was boosted compared to the phonology-only condition, especially for the irregular items (J. S. H. Taylor et al., 2011). In an old-new task like that of Wang et al. (2011), pairing pronunciations with definitions during training improved adults' ability to recognize trained target items. However, the authors collapsed across regularity, so it is unknown whether the added semantic information moderated orthographic learning to a similar extent for regular and irregular items. It leaves the question unanswered

whether these same effects would be observed for real words read in a naturalistic passage-level context.

A final methodological difference between the studies reviewed here is the number of words participants were asked to learn. Memory research suggests that the number of items to be learned affects the chances of retaining and retrieving chunks of information (Cowan & Alloway, 2009). It stands to reason that learning to read and spell a few words (e.g., Cunningham, 2006; Nation et al., 2007; Ouellette & Fraser 2009; Ricketts et al., 2011; Wang et al., 2011) compared to large set of words (e.g., Martin-Chang, 2017; Martin-Chang & Levy, 2005, 2006; Martin-Chang et al., 2007) does not carry the same burden. Researchers agree that context is drawn upon when reading is hindered (Stanovich, 1984). Challenges may arise from a variety of sources, including long, low-frequency words (Perfetti & Roth, 1981; Stanovich & West, 1981), irregular words (Nation & Snowling, 1998), degraded text (Stanovich & West, 1979), and weak reading skills (Nicholson, 1991; Stanovich et al., 1981). If training a large set of words increases task difficulty— and increasing difficulty augments reliance on context—it follows that the number of words to be learned might impact orthographic processing. Thus, mixed findings regarding the benefits of contextual facilitation for word learning could extend beyond outcome measures and in part be attributed to the nature and number of words to be learned. Yet, this methodological aspect has rarely been discussed as a point of interest within study designs (cf. Martin-Chang, 2017) and has never been directly tested.

Experiment 1

The self-teaching hypothesis suggests that when readers successfully decode a word, it functions as a mechanism to acquire orthographic representations (Share, 2011). With reading practice, more words are established in memory by connecting a word's spelling with its pronunciation and meaning(s) (Ehri, 2014; Perfetti & Hart, 2002; Share, 2011). Further, partial decoding attempts are resolved more often with the support of context than without. In isolation,

partial decoding attempts often lead to incorrect or incomplete pronunciations. In this scenario, each failed decoding attempt is a missed opportunity to bind the three key identities of word knowledge. Yet, mounting evidence suggests reading in context may draw cognitive resources away from processing the graphemes within words, resulting in weaker lexical representations in memory (Landi et al., 2006; Martin-Chang & Levesque, 2013, 2015; Miles & Ehri, 2017).

In Experiment 1, we regularized a large set of real words by transcribing them into phonetic symbols (Moats, 2003). This permitted examining whether contextual facilitation would be observed in a regular orthography. Training phonetic symbols exposed adults to a novel orthographic system, emulating the process of a novice reader mapping familiar phonemes onto unfamiliar graphemes, a situation akin to children in the partial alphabetic phase (Ehri, 2020). In this case, the target words were part of participants' oral vocabulary, and they had partial knowledge of the orthographic system (J. S. H. Taylor et al., 2015). This permitted decoding to proceed. For example, the standard English grapheme *b* and the phonetic symbol for the /b/ sound are identical and is represented in the onset of the word *bat*. At the same time, the lax vowel sound in *bat* is represented by the phonetic symbol /æ/ and is an unfamiliar grapheme. To further manipulate task demands, participants trained either a small or large set of unique words, in and out of context.

Method

A 2 (Contextual information: context vs. isolation) × 2 (Set size: small vs. large) mixed-subjects design was implemented. Contextual information was the within-subject manipulation; every participant trained target words in both the context and isolation condition. Stimulus set size was the between-subject manipulation. Participants were randomly assigned to either the small or large set size condition. The condition order was counterbalanced. The dependent variables were word reading accuracy during training and retention, spelling recognition, and spelling production. All mean comparisons were carried out with Sidak adjustment.

Participants

Forty-four participants were recruited through campus and social media postings. The materials were written at a Flesch-Kincaid Grade Reading Level of Grade 8-9. To ensure that participants could read and understand the material they completed the reading and spelling subtests of the WRAT4 (Wilkinson & Robertson, 2006) and the Nelson-Denny Reading Test (Brown et al., 1993). Two participants were excluded due scores below a Grade 8 equivalent and one more participant was excluded due to familiarity with the phonetic symbols. Thus, 41 participants (32 women) were included in the analyses. To ensure that participants in the set size groups (24 vs. 86) were similar in terms of age and literacy skills a series of *t*-test were run. All *ps* were nonsignificant (see Table 1). Thus, any differential outcomes in orthographic learning following the practice phase cannot be attributed to pre-existing differences on any of these reading related skills. The majority ($n = 35$) of participants reported English as their strongest language while six others reported English as their second strongest language. In all, eight participants were English monolinguals. Nine participants attained at least a college degree whereas a large majority held either an undergraduate ($n = 16$) or graduate degree ($n = 13$). Three participants held a PhD degree.

Table 1

Experiment 1: Descriptive Statistics

	Stimulus set size		<i>t</i> (39)	<i>p</i>
	Small 24 <i>M</i> (<i>SD</i>)	Large 86 <i>M</i> (<i>SD</i>)		
Age	32.90 (7.87)	30.00 (5.86)	1.34	.19
WRAT4 Reading ^a	63.95 (3.22)	62.81 (4.07)	.99	.30
WRAT4 Spelling ^a	48.60 (3.77)	48.81 (3.04)	.20	.85
ND Vocabulary ^a	67.20 (9.41)	66.71 (9.71)	.16	.87
ND Comprehension ^a	59.60 (9.81)	57.33 (8.33)	.80	.43
	<i>n</i> = 20	<i>n</i> = 21		

Note. WRAT4 = Woodcock-Johnson Reading Achievement Test-Fourth Edition. ND = Nelson-Denny. ^aRaw scores.

Materials

A total of 172 unique target words were selected (the majority were from Martin-Chang and Levesque, 2015). They were transcribed into phonetic symbols (Moats, 2003). Target words were between 2–to–13 phonemes ($M = 5.81$, $SD = 2.22$). They were nouns ($n = 90$), verbs ($n = 46$), and adjectives ($n = 36$), but not function words (e.g., prepositions) or other high-frequency words (e.g., determiners; Miles & Ehri, 2017). Further, target words did not overlap in root morphemes (e.g., fear and fearing).

The pool of 172 target words were divided into two sets of 86 words: Word Set A and Word Set B. In doing so, the part of speech (i.e., verbs, nouns, adjectives) were equally distributed across word sets. Further, the target words varied by word frequency and were balanced across word sets ($p = .78$). To create the small word set, a subset of 24 words were selected from each word set. These two sets of target words also varied by word frequency and were balanced across word sets ($p = .73$).

Two context stories were edited from Martin-Chang & Levesque (2015) to include the two word sets described above. Story A had 320 words and a Grade 8.9 reading level. Story B had 310 words and a Grade 8.6 reading level. The target words (either 86 or 24) appeared written with phonetic symbols (Moats, 2003) while, the remaining text was in standard English alphabet. The word sets were counterbalanced across all training conditions and set size groups.

Design and Procedure

A 2 (Contextual information: context vs. isolation) $\times$ 2 (Set size: small vs. large) mixed-subjects design was implemented. Contextual information was the within-subject manipulation; every participant trained target words in both the context and isolation condition. Stimulus set size was the between-subject manipulation. Participants were randomly assigned to either the small or large set size condition. The condition order was counterbalanced. The dependent

variables were word reading accuracy during training and retention, spelling recognition, and spelling production. All mean comparisons were carried out with Sidak adjustment.

Experiment 1 took place during the covid-19 pandemic, at a time when in-person activities were highly restricted. Thus, testing took place online, over three sessions. Stimuli presentation and responses were recorded with SuperLab Pro 5.0 (Cedrus Corporation, 2014). During the first session, participants completed the reading and spelling subtest of the WRAT4 and three training trials (i.e., reading in context and isolation). During the second session, participants completed three more training trials and the Nelson-Denny Reading Test. During the third and final session, participants completed testing phase (see below). Each session took place 3-days apart.

Training Phase. Participants were asked to read as quickly and accurately as possible. To measure learning, each story and list was read six times. The list was randomized across participants. The story appeared in its logical narrative format. The target words were written in bolded phonetic symbols (Moats, 2003). During the isolation condition (i.e., list reading) words were presented for 4 s. The experimenter scored accuracy on a keyboard. A fixation cross appeared for 500 ms before the next target word appeared. During the context condition (i.e., story reading), participants read the story from beginning to end. If a participant lingered on a target word, at 4 s the experimenter asked them to move on. This was done to maintain a parallel with training in isolation.

Testing Phase. Participants completed three tasks. They read a retention list, made old-new decisions (spelling recognition task), and spelled target words. For the retention task, participants read all the target words from both the context and list training conditions, randomized in a new list. Thus, participants who trained the large set size read 172 target words and those who trained in the small set size, read 48 target words. For the spelling recognition task, target words and never-before-read foils were randomized and presented one at a time.

Participants were asked to make old-new distinctions by determining whether the word was seen in training—thus they had to recognize the correct spelling of a target word and reject foils (misspellings). They were presented with the 48 correctly spelled words from the small word set (half of these were originally trained in context and the other half in isolation) and 48 corresponding misspellings. Foils consisted of either: (1) phoneme substitutions in which a single grapheme was replaced by a visually similar grapheme, (2) transpositions in which two adjacent graphemes were transposed, or (3) an articulation swaps in which a phoneme was replaced with another from a different of place of articulation but with the same manner of articulation (e.g., *t* swapped for *p*). Critically, no one responded yes to all trials; participants complied with the task parameters as demonstrated by the range of foils selected of zero to ten. Finally, participants spelled the same 48 words using phonetic symbols.

Results and Discussion

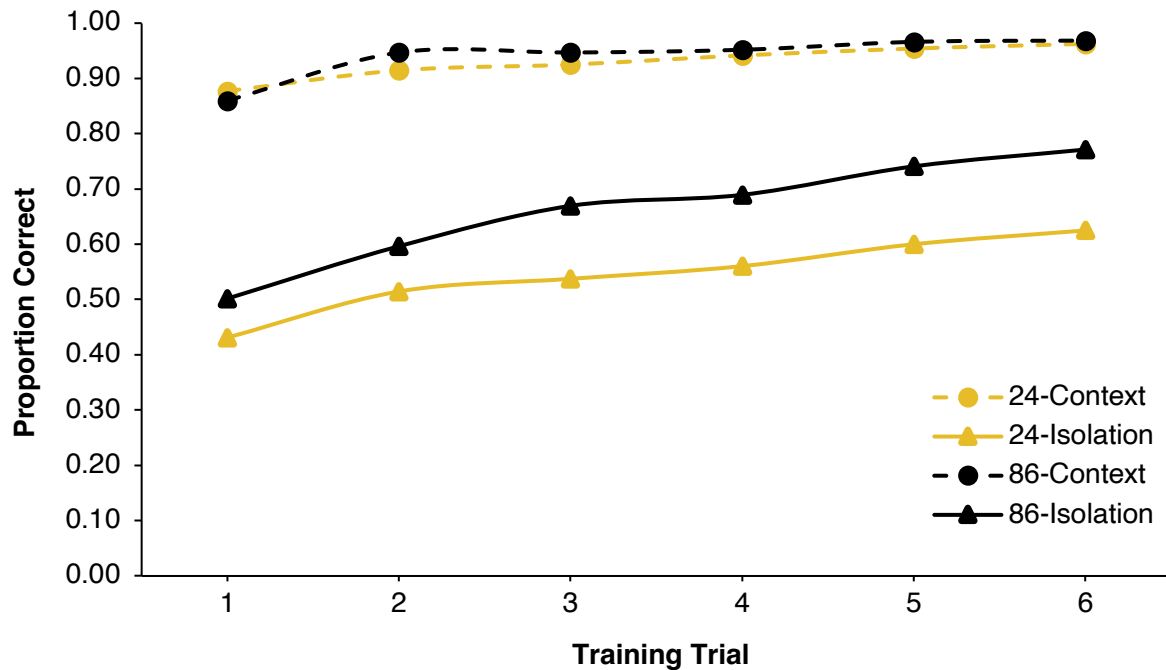
Word Reading Accuracy during Training

To determine whether adults were able to teach themselves to read words written with new symbols without explicit instruction or feedback, we conducted a 2 (Set size: 24 vs. 86 words) $\times$ 2 (Training condition: context vs. isolation) $\times$ 6 (Training trials 1–6) mixed ANOVA on word reading accuracy during training (see Figure 1). The main effect of set size, $F(1, 39) = 5.29$, $p = .03$, $\eta_p^2 = .61$, was significant with a large effect. This indicates that across training condition and training trials, participants attained better reading accuracy scores when they trained with large set size of 86 words compared to small set size of 24 words. This signals that number of words to be learned (set size) plays a role in reading accuracy of regularized words. There was a main effect of training condition, $F(1, 39) = 276.89$, $p < .001$, $\eta_p^2 = .88$, with a large effect. This indicates that across set size and training trials participants read more words accurately in context than in isolation. The main effect of training trials, $F(5, 195) = 200.35$, $p <$

.001, $\eta_p^2 = .84$, was also significant with a large effect. Thus, across set size and training condition, reading accuracy increased over successive trials. These findings reinforce the notion that lexicalization, at least in its most rudimentary form (i.e., good enough for reading), is facilitated by training in context (Martin-Chang, 2017; Martin-Chang et al., 2017).

Figure 1

Experiment 1: Proportion of Words Read Correctly during Training



Note. Scores are proportions out of 24 or 86.

The Set size $\times$ Training condition interaction was significant with a small effect, $F(1, 39) = 7.84$, $p = .01$, $\eta_p^2 = .17$. To determine whether this interaction remained at the last trial, we conducted a follow-up analysis with Trial 6 as the outcome variable. The pattern was corroborated, $F(1, 39) = 9.06$, $p = .01$, $\eta_p^2 = .19$. Further, pairwise comparison revealed that participants benefited from the large set size when reading in isolation $F(1, 39) = 8.86$, $p = .01$, $\eta_p^2 = .19$, but, not in context ($p = .67$; see Table 2). As expected, participants read more words accurately in context compared to isolation in both the small set size, $F(1, 39) = 102.27$, $p <$

.001, $\eta_p^2 = .72$, and large set size, $F(1, 39) = 36.63$, $p < .001$, $\eta_p^2 = .48$. This interaction is the second sign suggesting that set size plays a role in learning regularized word. Despite being at a disadvantage while reading in isolation, those who trained more words fared better during self-teaching. Reading accuracy was, nonetheless, best for words trained in context; likely because reading in context inherently increases the probability of correctly pronouncing a word while visually analyzing its printed form.

Continuing with interaction analyses, the Set size $\times$ Training trial interaction was statistically significant with a small effect, $F(5,195) = 5.55$, $p < .001$, $\eta_p^2 = .14$, showing that across training condition participants benefited from the larger set size on successive trials. Finally, the Training condition $\times$ Training trial interaction was statistically significant with a medium effect, $F(5,195) = 29.98$, $p < .001$, $\eta_p^2 = .43$, suggesting that across set size participants' reading accuracy was better in context on successive trials. Crucially, the same pattern of results was obtained when comparing participants in the large and small set size when limiting the analyses to the same set of 24 words in context and isolation.

Word Reading Accuracy during Retention

The benefit of increasing the strength of lexical representations in memory is that readers can access them under any circumstance, so that words can be fluently read without resorting to laborious decoding or requiring the scaffolding from context. As such, we asked our participants to read all the targets words they had previously trained in a new list. Reading accuracy during retention was examined with a 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA (see Table 2). The main effect of set size, $F(1, 39) = 4.52$, $p = .04$, $\eta_p^2 = .10$, was significant with a small effect. This indicates that across training condition, participants' word reading accuracy during retention was better for target words originally trained with the large set size compared to the small set size. These results echo the training

data. There was main effect of training condition, $F(1, 39) = 19.59$, $p < .001$, $\eta_p^2 = .33$, with a medium effect. This indicates that across set size, participants read more words accurately that were originally trained in context compared to isolation. Indeed, between Trials 1 and 6, participants were consistently more successful at decoding words in context compared to isolation. So, this increased self-teaching during training may have provisioned reading accuracy during retention.

Table 2

Experiment 1: Proportion Correct for Conditions and Tasks

Tasks	Stimulus set size			
	Small 24		Large 86	
	<i>M</i> (<i>SE</i>)	95% CI	<i>M</i> (<i>SE</i>)	95% CI
Training trial 1–6				
Isolation	.54 (.03)	[.48, .61]	.66 (.03)	[.60, .72]
Context	.93 (.01)	[.91, .95]	.93 (.01)	[.91, .96]
Training trial 6				
Isolation	.62 (.03)	[.55, .70]	.77 (.03)	[.70, .84]
Context	.96 (.01)	[.94, .98]	.97 (.01)	[.95, .99]
Retention in list				
Isolation	.64 (.03)	[.57, .71]	.78 (.03)	[.71, .85]
Context	.81 (.02)	[.77, .86]	.82 (.03)	[.77, .86]
Spelling recognition ^a				
Isolation	.96 (.01)	[.95, .98]	.97 (.01)	[.96, .99]
Context	.85 (.02)	[.81, .89]	.87 (.02)	[.84, .91]
Spelling production				
Isolation	.30 (.04)	[.21, .40]	.34 (.04)	[.26, .43]
Context	.28 (.04)	[.20, .37]	.29 (.04)	[.21, .36]

Note. Scores are proportions out of 24 or 86. ^aData is missing for one participant due to software failure.

The Set size $\times$ Training condition interaction was significant, with a small effect, $F(1, 39) = 7.78$, $p = .01$, $\eta_p^2 = .17$. Pairwise comparisons revealed that participants in the large set size outperformed those in the small set size for words originally trained in isolation, $F(1, 39) = 7.74$, $p = .01$, $\eta_p^2 = .17$, but, not in context ($p = .85$). And, participants read more words accurately that

had been trained in context compared to isolation in the small set size, $F(1, 39) = 25.41$, $p < .001$, $\eta_p^2 = .39$, but, not in the large set size ($p = .25$). Again, the same pattern of results was obtained when limiting the comparison to the subset of 24 words from context and isolation. Doing so ensured that the results were not unduly influenced by the larger denominator of the large set size.

When considering reading accuracy during retention for the large set size of 86 words, our findings corroborate those of Martin-Chang and colleagues (Martin-Chang & Levy, 2006; Martin-Chang, 2017). They reported similar reading accuracy for words that children initially trained in and out of context when later presented in isolation during a transfer task. However, in our study, when adults trained with the smaller set size of 24 words, the words that they originally trained in context were then read with better accuracy during retention. This suggests that the adults who had fewer opportunities to map orthographic representations onto pronunciations, learning in context was more beneficial than isolation. It seems that the bootstrapping effect of context was particularly relevant when learning occurred over fewer words (24 words). Recall that all participants read their training sets six times. Yet when learners had more decoding opportunities due to training 86 words, as much learning was observed in isolation as in context.

Spelling Recognition

We were also interested in whether learners were able to lay down orthographic representations during training that were precise enough to identify correct spellings. Thus, we conducted a 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA (see Table 2). The main effect of set size was not significant ($p = .27$). The main effect of training condition, $F(1, 38) = 44.69$, $p < .001$, $\eta_p^2 = .55$, was significant with a medium effect. This indicates that participants correctly identified more words that were originally trained in

isolation compared to words that were trained in context. The Set size $\times$ Training condition interaction was not significant ($p = .64$). Thus, these results contrast with the pattern for reading accuracy. Here, learners correctly recognized more words that they had initially trained in isolation compared to context. Further, the success rate on the recognition task was not moderated by training set size. A notable difference between the reading transfer and old-new decision task is that participants did not have to produce correct pronunciations to succeed but, required specified enough orthographic representations to reject foils and accept target words as correct.

Spelling Production

Spelling is an outcome measure that that is often shown to be a good indicator of lexical quality (Conrad, 2008; Conrad et al., 2019; Hersch & Andrews, 2012; Rossi et al., 2019; Veldre & Andrews, 2014). Thus, to evaluate whether orthographic knowledge differed on a spelling task, we conducted a 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA. The main effect of set size was not significant ($p = .69$). There was a main effect of training condition, $F(1, 39) = 4.03$, $p = .05$, $\eta_p^2 = .09$, indicating that across set size participants spelled more words correctly that were trained in isolation compared to context (see Table 2). The Set size $\times$ Training condition interaction was not significant ($p = .30$). Overall, spelling production was best for words that were initially trained in isolation—echoing the results of spelling recognition. This pattern of results may indicate that when adults could focus just on the orthography during isolation training, it came as an advantage for subsequent spelling production.

Experiment 2

What might explain the interaction between training condition and set size for word reading accuracy? As expected, in Experiment 1 reading accuracy was highest when reading

took place in context compared to isolation. What we add to the literature is that the difference in reading accuracy between context and isolation was smaller for those who trained with the large set size. Critically, this pattern remained stable when controlling for the stimulus set size, that is when only considering the same set of 24 words trained in context and isolation. Thus, even when limiting our comparison to the same set of words across groups, participants who had a richer experience during training closed the gap between isolation and context.

A good deal of orthographic learning occurs at the word level (Ehri, 2014; Perfetti, 2007; Share, 2011), yet this does not explain why participants who trained in isolation with the large word set fared better than those who trained with the small word set—a difference that widened as training progressed and remained significant during retention. The data shown here invite us to speculate whether learning propagated across words resulting in sublexical (general) orthographic learning (Conrad, 2008; Conrad & Deacon, 2016; Conrad et al., 2019; Seidenberg & McClelland, 1989; Zhao et al., 2018). When participants read a large set of words in isolation, they had more opportunities to crack the alphabetic code by matching graphemes to phonemes without assistance from top-down processes. In contrast, set size did not moderate reading accuracy in the context condition. The contextual facilitation observed in both the small and large set size may have increased performance to the point that it obscured any potential interaction. Thus, context may have rendered decoding less necessary, which could explain why reading an extra 62 words in the large set size condition boosted performance, but only for words trained in isolation.

This supposition was underpinned by the pattern of results that emerged for spelling recognition and production. Although stimulus set size was not a reliable predictor, training in isolation was the most beneficial for orthographic learning. This raises the question of why isolation training led to better spelling even though training in context resulted in better reading accuracy (Landi, 2013; Martin-Chang et al., 2017) and why training with the large set of words

did not increase performance. Scores on the spelling recognition task were very high for words trained in isolation; thus, this could simply be due to a ceiling effect. If this was the case, a more sensitive measure, such as reaction time, might register an interaction between set size and training condition. Turning to spelling production, the best performance was in isolation. Thus, spelling production resulted in a similar pattern of results, but the scores were much lower. Producing accurate spellings of words is more difficult than reading or recognition, thus a more sensitive measure that taps smaller units might shed some light on the development of sublexical orthographic learning.

Therefore, the purpose of Experiment 2 was to validate the interaction between stimulus set size and training condition for word reading accuracy as well as the isolation advantage for spelling. The experiment was replicated using a different set of target words. We also collected reaction times for spelling recognition to increase measurement sensitivity. Lastly, a pretest and post-test measure of individual graphemes was included. This permitted a more direct evaluation of sublexical orthographic knowledge.

Method

Participants

Forty-two participants were recruited through campus and social media postings. Two participants did not complete the study due to scheduling conflicts. Of the 40 remaining participants (24 women), the majority ($n = 29$) reported English as their strongest language and 11 others reported English as their second strongest language. Five participants attained at least a college degree whereas a large majority held either an undergraduate ($n = 19$) or graduate degree ($n = 14$). One participant held a PhD degree; another failed to report. For the WRAT4 (Wilkinson & Robertson, 2006), participants scored at least a Grade 11 reading equivalent and a Grade 10 spelling equivalent. To ensure that participants in the small set size and large set size groups were similar in terms of age and literacy skills, a series of *t*-tests were

run. All *ps* were nonsignificant (see Table 3). Thus, any differential outcomes in orthographic learning following the training should not be attributed to pre-existing differences on these reading-related measures.

Table 3

Experiment 2: Descriptive Statistics

	Stimulus set size		<i>t</i> (38)	<i>p</i>
	Small 24 <i>M</i> (<i>SD</i>)	Large 86 <i>M</i> (<i>SD</i>)		
Age	29.30 (6.19)	32.85 (6.98)	1.70	.10
WRAT4 Reading ^a	66.60 (1.90)	65.85 (3.42)	0.86	.40
WRAT4 Spelling ^a	50.55 (3.28)	49.25 (3.71)	1.17	.25
Grapheme pretest	.52 (.07)	.52 (.10)	0.35	.73
	<i>n</i> = 20	<i>n</i> = 20		

Note. WRAT4 = Woodcock-Johnson Reading Achievement Test-Fourth Edition. ^aRaw scores.

Materials

Literacy Skills. Reading and spelling skills were indexed by the Wide Range Achievement Test - Fourth Edition (WRAT4; Wilkinson & Robertson, 2006). Knowledge of the phonetic symbols (Moats, 2003) was assessed by presenting 39 graphemes individually in a randomized order. Participants were asked to provide a sound and a word that was representative of the grapheme. For example, when presented with /ð/, a correct response would be pronouncing the voiced /th/ and providing the word *there*. However, it proved difficult to pronounce individual phonemes, thus one point was accorded for each word containing a correct example. The maximum score was 39.

Training Stimuli. To carry out the replication with a different set of words, 83 of the original 172 target words were substituted. For the small word set, 45 of 48 target words were substituted. This permitted discounting potential idiosyncrasies of the target words implemented in Experiment 1 as a significant contributor to the pattern of results. Overall, target words were shorter compared to Experiment 1. Here, they were between 2–to–8 phonemes (*M* = 4.81, *SD* =

1.39). Target words were composed of 91 nouns, 44 verbs, and 37 adjectives. Like Experiment 1, function words were not included as target words. We also controlled the graphemes themselves by equally distributing each grapheme in both word sets (e.g., in the large word set, 11 instances of /ε/ appeared in both word sets). Further, the stories were piloted with 36 undergraduates using a cloze test. The findings were used to control for predictability of the target words. This resulted in stories where each target word was provided less than 10% of the time in the cloze task. Lastly, the target words varied by word frequency and were balanced across word sets (large set size, $p = .68$; small set size, $p = .92$). All parameters (e.g., stories, word frequency, part of speech) were similar to Experiment 1.

Procedure

Participants completed the same procedure as in Experiment 1 with two modifications. First, the third session occurred in the lab. Working with the participants in person allowed us to collect reaction times for the spelling recognition task. Participants were instructed to press a *yes* key for word they recognized and a *no* key for words they did not recognize. The keypress triggered the item to disappear which in turn triggered the following trial. SuperLab Pro 5.0 (Cedrus Corporation, 2014) recorded the delay between the item presentation and a participant's keypress. Second, participants completed the individual grapheme test, once prior to training and then again after training.

Results and Discussion

Word Reading Accuracy during Training

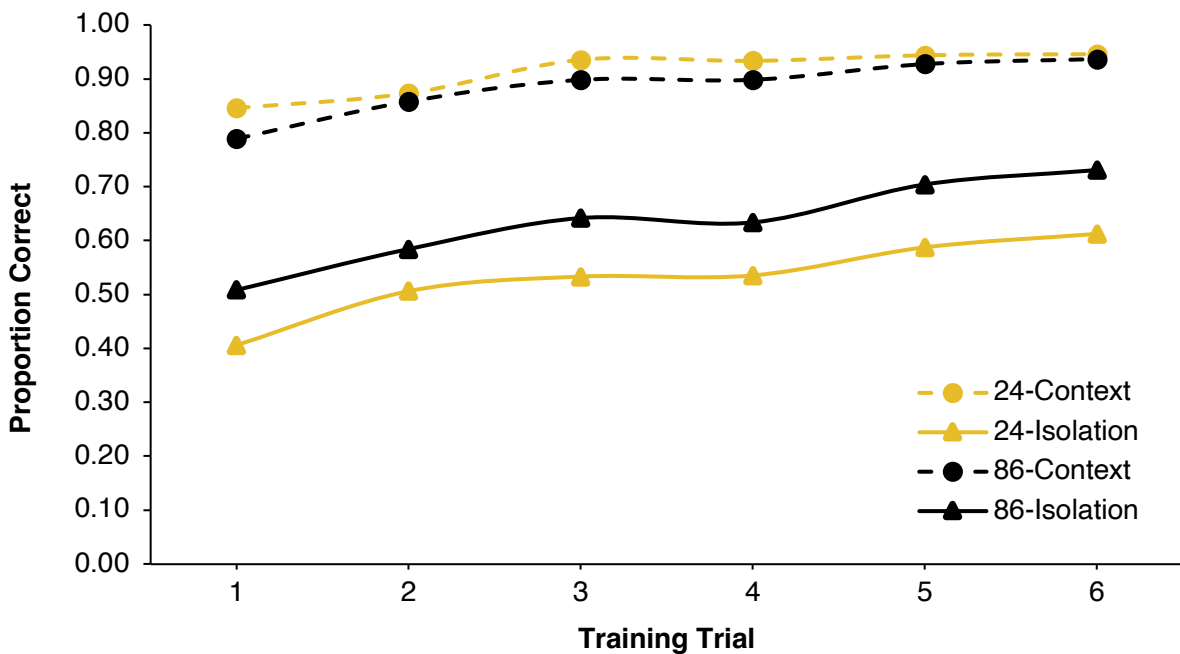
A 2 (Set size: 24 vs. 86 words) $\times$ 2 (Training condition: context vs. isolation) $\times$ 6 (Training trials 1–6) mixed ANOVA was conducted on the reading accuracy data (see Figure 2). All mean comparisons were carried out with Sidak adjustment for multiple comparisons. Unlike Experiment 1, the main effect of set size was not significant ($p = .18$). However, the main effect

of training condition, $F(1, 38) = 448.80, p < .001, \eta_p^2 = .92$, was significant with a massive effect. Thus, we replicated that across set size and learning trials, participants read more words accurately in context than in isolation. This replication supports the notion that contextual facilitation occurs for regularized words when reading is made challenging. Once again, the main effect of training trials, $F(5, 190) = 137.96, p < .001, \eta_p^2 = .78$, was significant with a large effect; showing that reading accuracy increased over successive trials. The Set size $\times$ Training condition interaction was significant with a medium effect, $F(1, 38) = 20.69, p < .001, \eta_p^2 = .35$.

To determine whether this interaction remained at the last trial, we carried out a follow-up analysis at Trial 6, $F(1, 38) = 12.57, p = .005, \eta_p^2 = .25$. Replicating Experiment 1, pairwise comparisons confirmed that participants benefited from the large set size when reading in isolation $F(1, 38) = 6.97, p = .01, \eta_p^2 = .15$, but there was no difference in performance in context ($p = .56$). Furthermore, participants read more words accurately in context compared to isolation in both the small set size, $F(1, 38) = 171.72, p < .001, \eta_p^2 = .82$, and large set size, $F(1, 38) = 65.47, p < .001, \eta_p^2 = .63$. Continuing with interaction analyses, the Set size $\times$ Training trial interaction was at trend ($p = .06$). Finally, the Training condition $\times$ Training trial interaction was significant with a small effect, $F(5, 190) = 14.01, p < .001, \eta_p^2 = .27$. This suggests that from Trial 1 to 6, participants improved more in isolation relative to context.

Figure 2

Experiment 2: Proportion of Words Read Correctly during Training



Note. Scores are proportions out of 24 or 86.

Experiment 2 therefore replicates the finding that the number of words to be learned (set size) moderates the association between contextual facilitation and word reading accuracy during training. Using a different set of stimuli, we again found that while reading in isolation, participants who trained with more words performed better compared to those who trained with fewer words. From inspecting Figure 2, it is apparent that contextual facilitation operates right at the outset of training.

A novel result from Experiment 2 is that at Trial 1, participants who trained with the large set size had better word reading accuracy than those who trained with the small set size. Because this was a between-subjects comparison, it could be that participants in one group were stronger readers relative to the other group. However, two things discount this possibility. The first is that scores on the reading and spelling task of the WRAT4 were similar between the two set size groups. In addition, if one group inadvertently contained better readers, we would

expect to see better scores from that group in both conditions. However, here participants in the large set size read more words correctly in isolation compared to those in the small set size whereas, it was the other way around in context.

It could be that the target words in the small set size (which are a subset of the target words in the large set size) were more challenging than the additional words in the large set size. However, this notion was ruled out because we obtain a similar pattern of results when constraining the comparison to the overlapping 24 words for participants in the large and small set sizes. Thus, it is more likely that orthographic learning was occurring at the beginning of Trial 1 and participants applied the acquired knowledge to decipher other words later within the same trial. For example, in the first half of the Trial 1 participants may have self-taught the symbol /ɛ/ by reading *bɛst* and apply it later in the trial to decode *gaten*. It seems that learning occurred within the first trial in addition to between trials. This was evident in isolation for the large set size; participants gained an advantage and maintained the advantage throughout training.

Word Reading Accuracy during Retention

Reading accuracy during retention was examined with a 2 (Set size: 24 vs. 86 words) × 2 (Training condition: context vs. isolation) mixed ANOVA with Sidak adjustment (see Table 4). Unlike Experiment 1, the main effect of set size was not significant ($p = .36$). Similar to Experiment 1, the main effect of training condition, $F(1, 38) = 23.38$, $p < .001$, $\eta_p^2 = .38$, was significant with a small-to-medium effect. Once again, the Set size × Training condition interaction was significant with a small effect, $F(1, 38) = 9.09$, $p = .005$, $\eta_p^2 = .19$. Participants read more words accurately that were trained in context compared to isolation in the small set size, $F(1, 38) = 30.82$, $p < .001$, $\eta_p^2 = .45$, but not in the large set size ($p = .21$). When comparing performance between the set sizes, participants who trained with large set size outperformed

those from the small set size for words originally trained in isolation, $F(1, 39) = 4.32$, $p = .04$, $\eta_p^2 = .10$, but not in context ($p = .70$). These results corroborate those of Experiment 1 despite using a different set of stimuli.

Table 4

Experiment 2: Proportion Correct for Conditions and Tasks

Tasks	Stimulus set size			
	Small 24		Large 86	
	<i>M (SE)</i>	95% CI	<i>M (SE)</i>	95% CI
Training trial 1–6				
Isolation	.53 (.03)	[.47, .58]	.63 (.03)	[.58, .69]
Context	.92 (.01)	[.89, .94]	.88 (.01)	[.86, .91]
Training trial 6				
Isolation	.61 (.03)	[.55, .68]	.73 (.03)	[.67, .79]
Context	.95 (.01)	[.92, .97]	.94 (.01)	[.91, .96]
Retention in list				
Isolation	.65 (.03)	[.59, .71]	.74 (.04)	[.68, .81]
Context	.79 (.03)	[.73, .86]	.77 (.03)	[.71, .84]
Spelling recognition				
Isolation	.87 (.04)	[.79, .94]	.90 (.04)	[.83, .98]
Context	.81 (.03)	[.76, .87]	.82 (.03)	[.77, .87]
Spelling recognition RT				
Isolation	1200 ms	[1025, 1375]	1373 ms	[1202, 1543]
Context	1359 ms	[1144, 1573]	1584 ms	[1375, 1793]
Spelling production				
Isolation	.25 (.04)	[.18, .32]	.27 (.04)	[.20, .34]
Context	.21 (.03)	[.15, .26]	.22 (.03)	[.16, .28]

Note. Scores are proportions out of 24 or 86. RT = Reaction time.

Here we asked which training condition would support self-teaching so that participants could later read the target words without scaffolding from context. At transfer, target words were presented individually for up to 4 s and only whole word responses were accepted as correct. Despite these stringent conditions, the success that participants cultivated when training in context translated to better word retention during retention for those in the small set size. Whereas, when participants trained with the large set size, performance between isolation and

context was similar during retention. This may seem obvious, that training more words results in more learning. However, it has never been tested empirically and this effect has practical implications (Colenbrander et al., 2019; Nation, 2017).

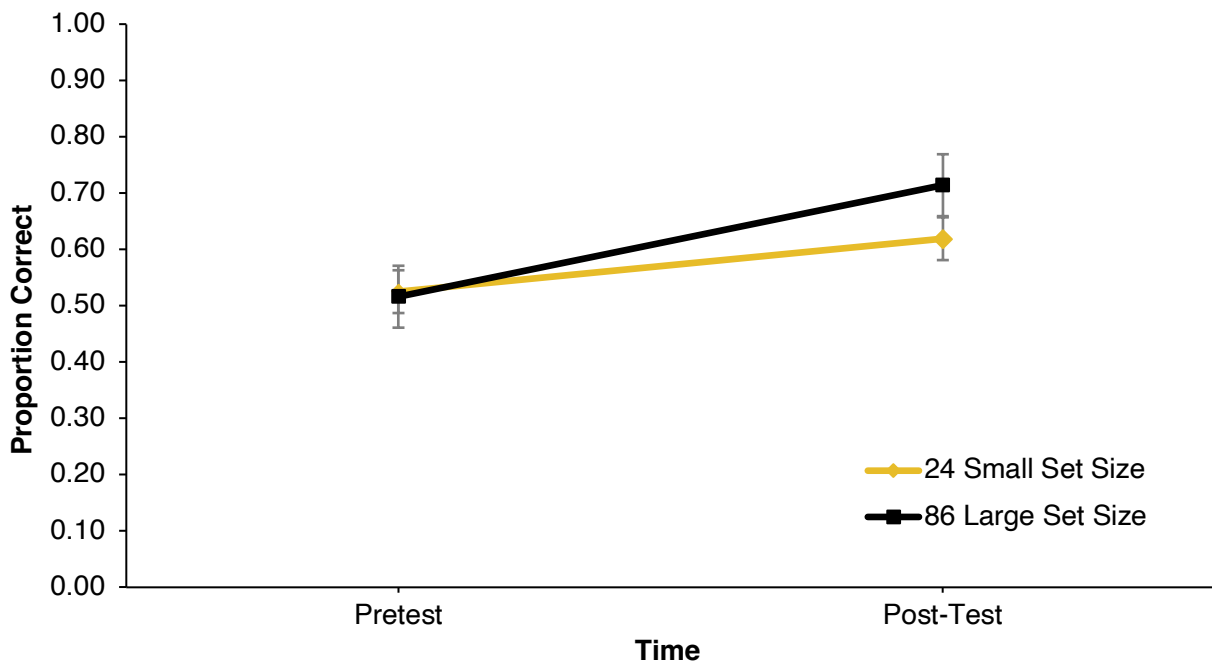
Individual Grapheme Test

In both Experiment 1 and 2, when training in isolation, participants in the large word set fared better than those who trained with the small word set—a difference that widened as training progressed and remained significant during retention. As outlined previously, it is possible that training with the large set size fostered orthographic learning at the sublexical level. The grapheme post-test sheds light on whether set size influenced participants' learning of the phonetic symbols implemented in the study.

A 2 (Time: pre vs. post) $\times$ 2 (Set size: 24 vs. 86) mixed ANOVA on word example was carried out with Sidak adjustment (see Figure 3). The main effect of time, $F(1, 38) = 51.08, p < .001, \eta_p^2 = .57$, was significant with a medium-to-large effect. This confirms that across set size, participants were more successful at providing examples after training. The main effect of set size was not significant ($p = .11$). However, the Time $\times$ Set size interaction was significant, $F(1, 38) = 6.46, p = .01, \eta_p^2 = .14$, with a small effect. Pairwise comparison confirmed that participants in both the small and large stimulus set size improved from pretest to post-test, $F(1, 38) = 10.61, p = .002, \eta_p^2 = .22$; $F(1, 38) = 46.94, p < .001, \eta_p^2 = .55$, respectively. At post-test, participants who trained with the large word set outperformed those who trained with the small word set, $F(1, 38) = 6.04, p = .02, \eta_p^2 = .14$. This indicates that participants who trained with the large set size learned more grapheme-to-phoneme mappings compared to those in the small set size.

Figure 3

Experiment 2: Individual Grapheme Test



Note. Scores are proportions out of 24 or 86.

Spelling Recognition

To evaluate whether participants could identify correctly spelled words on an orthographic decision task, we conducted a 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA with Sidak adjustment. Three participants were excluded; two because they did not follow instructions and another due to a technical failure. The analyses were performed on the remaining 37 participants. As shown in Table 4, the main effect of set size was not significant ($p = .63$). The main effect of training condition was significant, $F(1, 35) = 10.15$, $p = .003$, $\eta_p^2 = .22$, with a small effect. The Set size $\times$ Training condition interaction was not significant. Thus, similar to Experiment 1, adult learners identified more words that they initially trained in isolation relative to those trained in context.

Reaction time. The following analyses were conducted with the 37 participants included in the spelling recognition task. The data were then trimmed by removing 142 trials that were either faster than 300 ms or skipped due to technical failures. This resulted in 3.7% of the reaction time data being discarded from the analyses. Lastly, only items that were correctly recognized as a trained target word were used to compute participant's mean reaction time ($n = 3702$ trials). A 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA was performed with Sidak adjustment. The main effect of set size was not significant ($p = .11$). The main effect of training condition, $F(1, 35) = 30.70$, $p < .001$, $\eta_p^2 = .47$, was significant with a medium effect. This suggests that participants responded more quickly to words that were originally trained in isolation relative to context. The Set size $\times$ Training condition interaction was not significant ($p = .43$).

Taking the spelling recognition and reaction time data together, it suggests that the word-form representations acquired during isolation training were specified enough to support proficient recognition at test. By the end of training, our participants correctly read 27.5% more words in context compared to in isolation yet, we found that words participants read in isolation transferred better to this spelling-related task.

Spelling Production

To evaluate whether orthographic knowledge differed on a spelling task, we conducted a 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA with Sidak adjustment. One participant was excluded due to a z-score score more than 3 *SD* away from the mean. The main effect of set size was not significant ($p = .64$). Looking at the difference between training conditions, spelling production was numerically higher for words that were originally trained in isolation relative to context, but this difference was not significant, $F(1, 37) = 3.51$, $p = .07$ (all other $ps > .16$). Thus, regardless of whether words were read as part of the

large or small word set, or in context or isolation, they were all spelled with the same success rate. And although training words in isolation was not superior to context, it is still rather surprising that participants were reading so well in context but failing to transfer that accuracy into their spelling.

Experiment 3

In Experiments 1 and 2 we found evidence that training in isolation with the large stimulus set size helped participants learn more about the phonology and orthography of the phonetic symbols than training with a small stimulus set. This was evident in the reading measures including the grapheme post-test. Here, we asked whether engaging with a large number of words might lead to orthographic knowledge that generalizes to new, untrained words.

Beginning readers cannot be explicitly taught every word they will encounter. There are far too many words to learn on an individual basis. Rather children and adult learners are taught phonics in order to self-teach (via decoding) new words that contain familiar orthographic patterns (Ehri, 2020). A skilled reader can extrapolate through decoding new words that follow typical spelling-sound patterns (i.e., regular words). In experimental terms this is generalization. We did not measure generalization in Experiments 1 and 2, but we would expect participants who trained with the large set size to outperform those from the small set size.

For Experiments 3, we again created a large and small word set of target words, but this time, set size was a within-subjects manipulation. This permitted us to refine our investigation on the effect of set size on orthographic learning. To do this, five target graphemes were presented 18 times each across 55 different target words, thus creating the large set size. Another five target graphemes were presented 6 times each across 22 different target words, thus creating the small set size. We then asked whether reading more words lead to better generalization when reading and spelling new words.

Method

Participants

Fourteen participants (12 females) were recruited through campus and social media ($M_{\text{age}} = 28.07$, $SD = 6.71$) and completed the reading and spelling subtests of the WRAT4 (Wilkinson & Robertson, 2006) and the Nelson-Denny Reading Test (Brown et al., 1993). Participants scored at least a Grade 12 reading and spelling equivalent on the WRAT4 reading ($M_{\text{raw}} = 62.21$, $SD = 2.89$) and spelling subtests ($M_{\text{raw}} = 48.79$, $SD = 2.55$). For the Nelson-Denny Reading test, all participants scored above a Grade 12 equivalent (vocabulary: $M_{\text{raw}} = 62.43$, $SD = 9.84$; comprehension: $M_{\text{raw}} = 53.43$, $SD = 9.72$). All but one participant reported English as their strongest language. Two participants attained a high school degree and were currently pursuing a college degree. Five participants attained at least a college degree, three held an undergraduate degree, and four held a graduate degree.

Materials

Training Stimuli. To create the experimental conditions, graphemes were manipulated within the target words. To create the large set size, five different phonetic symbols (Moats, 2003) occurred 18 times each across 55 target words (phonetic symbols: /i/ like eagle, /æ/ like apple, /u/ like ooze, /č/ like chat, and /z/ like zoo). To create the small set size, five mutually exclusive phonetic symbols occurred six times each across 22 target words (phonetic symbols: /ɛ/ like echo, /e/ like ape, /ʌ/ like up, /š/ like she, /j/ like juice). Together, the 77 target words comprised a word set. Another word set was created with the same 10 symbols, but those that occurred 18 times instead occurred six times and vice versa. All other symbols within target words were controlled across word sets.

The target words were 2–to–7 graphemes long ($M = 4.43$, $SD = 1.20$). Target words were nouns, verbs, and adjectives, but not function words (e.g., prepositions) or other high-frequency words (e.g., determiners; Miles & Ehri, 2017). Nouns, verbs, and adjectives were

equally distributed across both word sets of 77 target words. Further, word frequency of the target words was balanced across two words sets ($p = .14$). The word sets were counterbalanced across participants.

Design and Procedure

This experiment employed a fully within-subjects design. Participants trained in isolation and completed six trials. In each trial participants were exposed to all 77 target words (55 large set size target words and 22 small set size target words). The study took place online and delivered with SuperLab Pro 5.0 (Cedrus Corporation, 2014). During the first session, participants completed the reading and spelling subtest of the WRAT4 and three training trials in isolation. During the second session, three more training trials were completed followed by the Nelson-Denny Reading Test. During the third session, participants read and spelling old, trained and new, untrained target words. Each session took place 3-days apart.

The new, untrained words were composed with graphemes from both the small and large set size. This was done to determine whether trained graphemes would generalize to new, untrained words. Participants spelled 46 target words using the phonetic symbol from the training phase. For each participant, half of these were previously trained, and the other half were new, untrained words. Each target grapheme, regardless of set size, was spelled six times. The dependent variables were word reading accuracy during training, word reading accuracy during retention, and spelling production. Analyses were carried out with Sidak adjustment for multiple comparisons.

Results and Discussion

Word Reading Accuracy during Training in Isolation

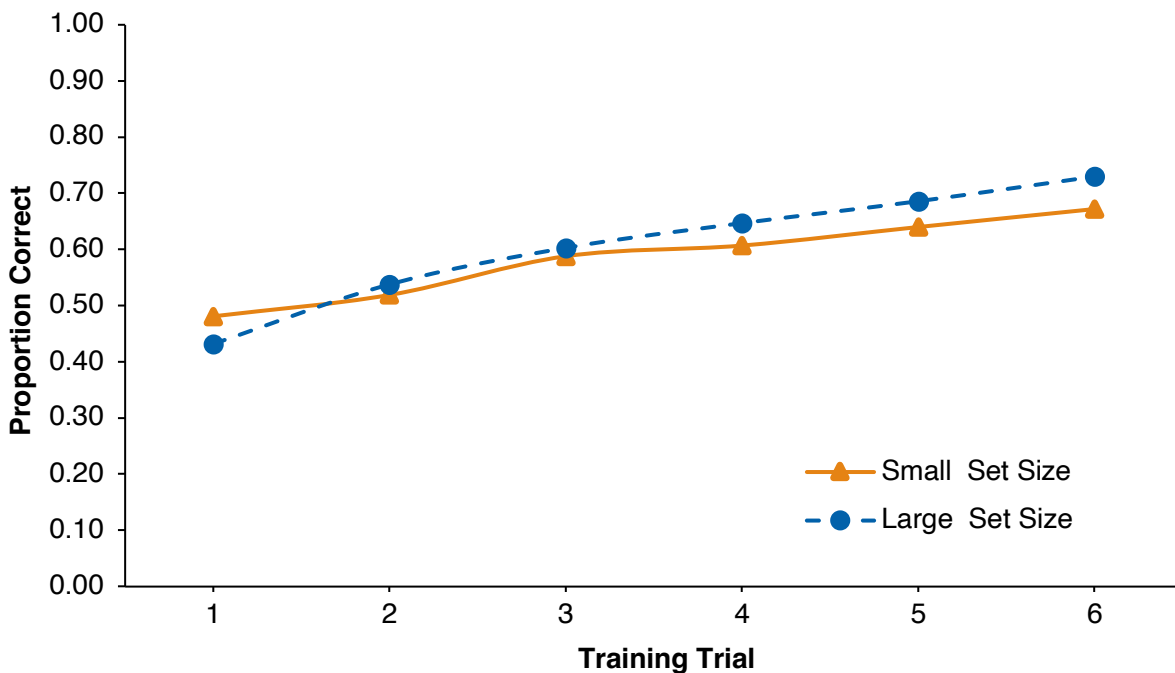
To examine whether word reading accuracy in isolation differed between the two set sizes (small vs. large) and training trial (1–6), we conducted a 2×6 repeated measures ANOVA. The main effect of set size was not significant ($p = .49$). The main effect of training trial,

$F(5, 65) = 41.85, p < .001, \eta_p^2 = .76$, was significant with a medium-to-large effect. This suggests that across set size, accuracy increased on successive trials (see Figure 4). The interaction between set size and trial was significant with a small effect, $F(5, 65) = 4.61, p = .001, \eta_p^2 = .26$.

Of greater interest, we evaluated whether an effect of set size was present at Trial 6. There was a reliable difference between set size, $F(5, 12) = 4.71, p = .05, \eta_p^2 = .27$, with a small effect. Keeping in mind that the effect size is small, this suggests that reading accuracy was best for the large set size compared to the small set size.

Figure 4

Experiment 3: Proportion of Words Read Correctly during Isolation Training



Note. Scores are proportions out of 22 or 55.

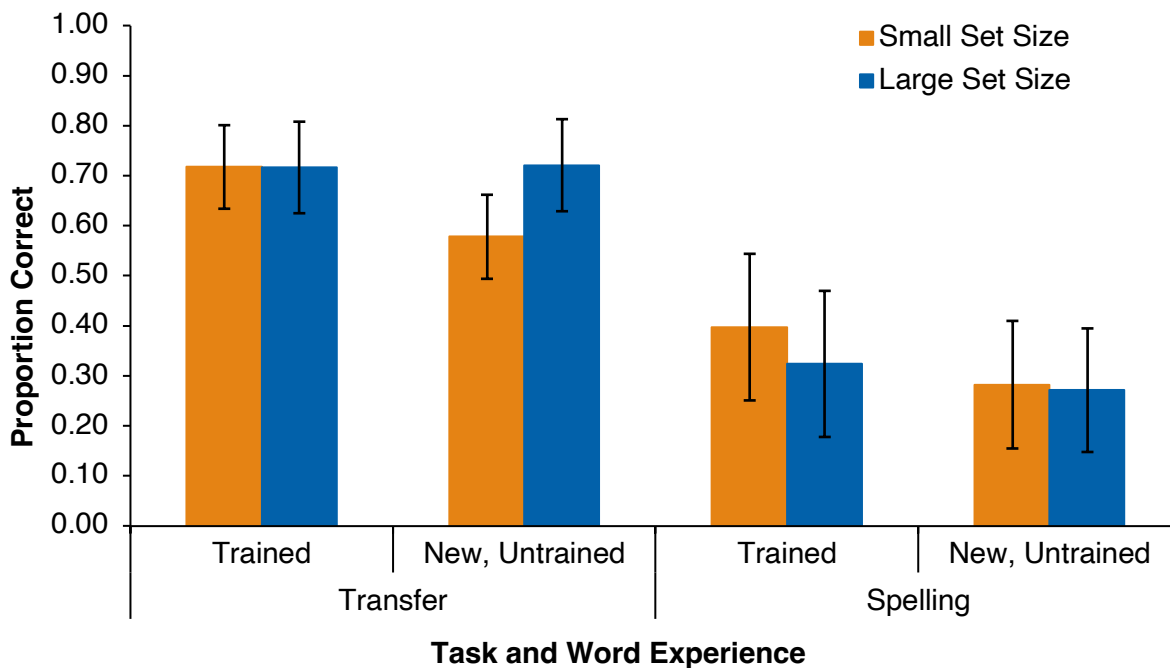
Word Reading Accuracy at Transfer

Participants subsequently read trained target words and new, untrained words in a new list (154 words total; see Figure 5). To evaluate whether word reading accuracy differed at transfer, we conducted a 2 (Set size: small vs. large) $\times$ 2 (Word experience: trained vs. new)

repeated measures ANOVA. The main effect of set size, $F(1, 13) = 7.98, p = .01, \eta_p^2 = .38$, was significant with a small-to-medium effect. This indicates that across word experience, participants accurately read more words from the large set size compared to the small set size. The main effect of word experience, $F(1, 13) = 6.89, p = .02, \eta_p^2 = .35$, was also significant with a small-to-medium effect; indicating participants read trained words more successfully compared to new, untrained words, regardless of set size.

Figure 5

Experiment 3: Proportion Correct for Set Size for Transfer and Spelling Production



Note. Scores are proportions out of 22 or 55.

The interaction was also significant, $F(1, 13) = 21.59, p < .001, \eta_p^2 = .62$, with a medium-to-large effect. Pairwise comparisons revealed that for trained target words, there was no difference between the small and large set size ($p = .98$). Of note, performance from the last training trial to transfer remained stable ($M = .73$ vs $.73, p = .46$) for the large set size, whereas words from the small set size improved from another chance at reading the target words at

transfer despite it being in a new list; performance increased ($M = .67$ vs $.72$, $p = .003$). Thus, an extra trial was required for target words from the small set size to catch up to the large set size. To stress this point, the same participants trained all 77 target words at the same time yet, set size played a role in orthographic learning.

Looking at generalization, participants accurately read more new, untrained words from the large compared to the small set size, $F(1, 13) = 17.11$, $p = .001$, $\eta_p^2 = .57$. Self-teaching was thus optimized during training by practicing with the large set size. The large set size had a greater breadth of words that resulted in more exposures to each target grapheme. The small set had the same number of graphemes but, participants saw them less frequently. Overall, participants had fewer opportunities to decode the graphemes in the small set size.

Returning to the comparison, we found that word reading accuracy was higher for trained compared to untrained words for target words from the small set size, $F(1, 13) = 27.39$, $p < .001$, $\eta_p^2 = .68$, whereas participants were equally successful for reading all words from the large set size ($p = .91$). This pattern reinforces the interpretation that more is better for learning and extends to generalization. Fewer opportunities to decode resulted in orthographic knowledge that was good enough for remembering trained words but not consolidated enough to generalize to new, untrained words. In this case, we would expect a similar pattern for spelling production.

Spelling Production

To evaluate whether orthographic knowledge differed on a spelling task, we conducted a 2 (Set size: small vs. large) $\times 2$ (Word experience: trained vs. new) repeated measures ANOVA. Participants spelled a total of 46 words (24 trained, 24 new). The main effect of set size was not significant ($p = .45$). There was a main effect of word experience, $F(1, 12) = 22.76$, $p < .001$, $\eta_p^2 = .65$, with a medium-to-large effect, indicating that participants fared better when asked to spell

trained target words compared to untrained target words. The Set size $\times$ Word experience interaction was at trend, $F(1, 12) = 4.25$, $p = .06$, $\eta_p^2 = .26$. Pairwise comparisons revealed that, unlike the transfer data, participants accurately spelled a similar number of new, untrained words from both the large and small set size (all $ps > .27$). However, similar to transfer, the data showed that spelling production was best for trained compared to untrained words in the small set size, $F(1, 12) = 28.62$, $p < .001$, $\eta_p^2 = .70$, whereas the difference was at trend for the large set size, $F(1, 12) = 4.42$, $p = .06$, $\eta_p^2 = .27$. Thus, set size was not a strong predictor for spelling. This suggests that establishing lexical representations in memory of novel words is a long process that is not complete or fully lexicalized after several trials (Ginestet et al., 2020; J. N. Taylor & Perfetti, 2016). Word-form representations precise enough for spelling take a longer time to consolidate. The different success rates between reading accuracy at transfer and spelling production also reflects the fact that lexical quality rests on a continuum (Rossi et al., 2019).

Experiment 4

In Experiment 2 we examined the orthographic learning that took place in isolation because this condition lent itself to noticing, processing, and encoding the graphemes within words. Encouraged by our findings, we wondered whether a similar pattern would be observed if reading took place with the top-down support of context. Based on our observations in Experiments 1 and 2, there were two competing predictions. (1) Set size would affect orthographic learning during context training and would be present at generalization but dampened. (2) The set size effect would be absent during training and at generalization. Thus, Experiment 4 replicated Experiment 2, but training took place in context. Based on results from Experiments 1 and 2, we anticipated that set size might only affect word reading accuracy and

that contextual facilitation during training would hamper developing lexical representations precise enough for spelling.

Method

Participants

Fourteen, different participants (12 females) were recruited through campus and social media postings ($M_{\text{age}} = 26.00$, $SD = 6.04$). For both the reading and spelling subtests of the WRAT4 (Wilkinson & Robertson, 2006), participants scored at least a Grade 12 reading and spelling equivalent (reading, $M_{\text{raw}} = 63.43$, $SD = 2.74$; spelling, $M_{\text{raw}} = 49.21$, $SD = 3.45$). For the vocabulary component of the Nelson-Denny Reading test (Brown et al., 1993), all participants scored at least a Grade 12 equivalent ($M_{\text{raw}} = 66.64$, $SD = 9.12$) and for the comprehension component they scored at least a Grade 12 equivalent ($M_{\text{raw}} = 63.00$, $SD = 6.69$). All but two participants reported English as their strongest language and these two reported that English was their second strongest language. Two participants attained a high school degree and were currently pursuing a college degree. Five participants attained at least a college degree, five held an undergraduate degree, and two held a graduate degree.

Materials

Training Stimuli. The two context stories from Experiment 1 were edited to include the target words outlined in Experiment 2. Story E included the target words from Word Set E. It was 453 words long with a 7.5 Flesch-Kincaid grade reading level. Story F included the target words from Word Set F. It was 465 words long with a 7.2 Flesch-Kincaid grade reading level. Again, only the target words were presented with phonetic symbol (Moats, 2003) and the remaining text was written in the standard English alphabet (see Appendix K).

Design and Procedure

The procedure was identical to Experiment 3 with one exception. Participants trained target words in context rather than in isolation. Again, the dependent variables were word

reading accuracy during training, word reading accuracy at transfer to list of train and new, untrained words, and spelling production of train and new, untrained words.

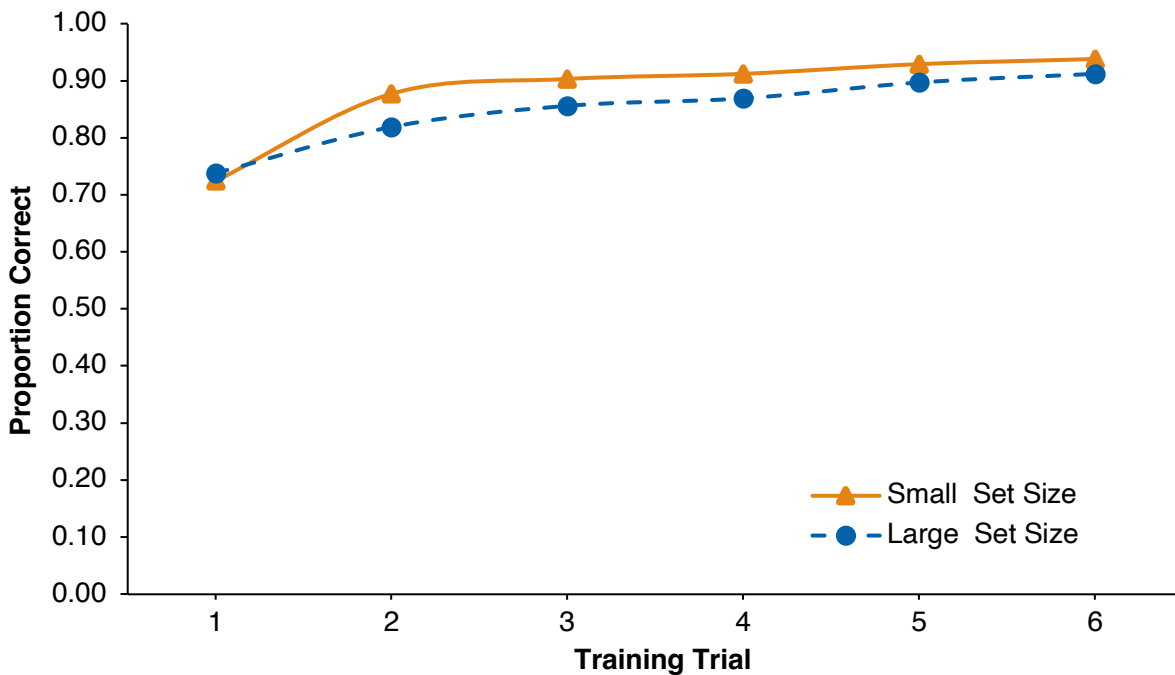
Results and Discussion

Word Reading Accuracy during Training in Context

Here, we examined whether word reading accuracy differed between set size (small vs. large) and training trials (1–6) by conducting a 2×6 repeated measures ANOVA. The main effect of set size was significant, $F(1, 13) = 5.21$, $p = .04$, $\eta_p^2 = .29$, with a small effect. Unlike the three experiments reported here, participants surprisingly read more target words from the small set size compared to the large set size. The main effect of training trial, $F(5, 65) = 42.46$, $p < .001$, $\eta_p^2 = .77$, was significant with a large effect. Thus, across set size, accuracy increased on successive trials (see Figure 6). The Set size $\times$ Training trial interaction was significant with a small effect, $F(5, 65) = 4.48$, $p = .001$, $\eta_p^2 = .26$. To determine whether the effect of set size remained at the last trial, we carried out follow up analyses with Trial 6 as the outcome variable; it was at trend ($p = .07$). When reading with the support of context, this suggests that by the end of training set size no longer affected participants' word reading accuracy. It is possible that contextual facilitation supported reading accuracy to the point that it superseded set size. Indeed, word reading was near ceiling by Trial 3 which may explain why this pattern of results differs from those observed of participants in Experiment 2 who trained in isolation.

Figure 6

Experiment 4: Proportion of Words Read Correctly during Context Training



Note. Scores are proportions out of 22 or 55.

Word Reading Accuracy at Transfer to List

Although training took place in context, transfer occurred in isolation in order to measure reading gains without the facilitatory effect of context. For the transfer task, participants read both trained target words and new, untrained words in a list. To evaluate whether word reading accuracy differed between conditions, we conducted a 2 (Set size: small vs. large) $\times$ 2 (Word experience: trained vs. novel) repeated measures ANOVA. The main effect of set size was not significant. The main effect of word experience, $F(1, 13) = 48.38$, $p < .001$, $\eta_p^2 = .79$, was significant with a large effect; indicating that across set size, participants were more successful at reading trained words ($M = .80$, $Cis [.73, .86]$) compared to new, untrained words ($M = .59$, $Cis [.50, .68]$). This again shows that training in context fostered lexical representations that are *good enough* to support learning; if participants had not learned, performance would have

remained unchanged from Trial 1 to 6 (Martin-Chang, 2017; Martin-Chang et al., 2017). The interaction was not significant ($p = .46$). These results are in line with the Trial 6 training data indicating that set size did not moderate orthographic learning when training occurred with the support of context. Generalization was not affected by this factor either. A reason might be that contextual facilitation during training allowed participants to arrive at the correct pronunciation without closely analyzing the orthography (Staub, 2015; Tiffin-Richards & Schroeder, 2020).

Spelling Production

To evaluate whether spelling production differed, we conducted a 2 (Set size: small vs. large) $\times$ 2 (Word experience: trained vs. untrained) repeated measures ANOVA. Recall that participants spelled 46 words using the phonetic symbols learned during training; half of these were originally trained in context and the other half were new, untrained words, and each were drawn from small and large set sizes. The main effect of set size was not significant ($p = .25$). There was a main effect of word experience, $F(1, 13) = 12.89$, $p = .003$, $\eta_p^2 = .50$, with a medium effect, indicating that participants fared better when asked to spell trained target words ($M = .41$, CIs [.30, .52]) compared to untrained target words ($M = .32$, CIs [.20, .44]). This suggests that lexical representations were created during context training. The interaction was not significant. This data seems to corroborate the idea that the effect of context overrides the effect of stimulus set size.

General Discussion

We investigated whether contextual facilitation would interact with stimulus set size and, in turn, affect orthographic learning. We had five goals for this study. We asked whether contextual facilitation during training occurs for regular words when reading is challenging. The results were conclusive; it does. Second, we asked whether contextual facilitation during training would support later reading when transferring to isolation. Once again, our data were

conclusively positive. Third, we asked whether the effect of context would support or undermine spelling recognition and production. Like Martin-Chang et al. (2017), we found that context training was not as beneficial as isolation training when it came to spelling. Fourth, we asked whether training a large set of words would have a negative impact on reading accuracy and thus orthographic learning. Contrary to our supposition, training a large set of words improved performance in isolation for word reading accuracy. In response to this, our final question asked whether training more words would lead to generalization. Our results were once again positive. We found that adults were better at learning phonetic symbols, and thus reading and spelling new words with trained graphemes from the large set size. Thus, we add to the literature by showing that the advantages of training in isolation is most prominent when participants are given ample training reading words in the large set size.

Regular Words and Contextual Facilitation during Training and Retention

There is a longstanding debate over whether context bolsters learning of regular words (e.g., Nation & Snowling, 1998; Wang et al., 2011, 2012, 2013; for review, see Colenbrander et al., 2020). Here, we examined contextual facilitation during and after training in a shallow orthography using phonetic symbols. Participants read nearly 30% more words correctly at Trial 1 in context compared to isolation. In our view, this attests to the power of contextual facilitation. Indeed, this large discrepancy in favour of context was observed in three quarters of the sample. Viewed differently, not one participant across the experiments met with greater success in isolation during training. Moreover, despite the regular orthography, the effect of contextual facilitation remained evident until the very last training trial.

Importantly, we do not support the claim that reading is a by-product of guessing from context (for discussion, see Stanovich, 1984). As discussed in the introduction, contextual facilitation is a secondary self-teaching mechanism that scaffolds decoding rather than replacing it (Share & Stanovich, 1995). If participants were guessing from context during training, we

would expect very little retention when words were presented in isolation at transfer. Indeed, when adults trained with the small set size, performance was best for words that were trained in context relative to isolation (performance was equally high in the large set size). This finding argues against the notion that reading in context does not lead to learning (Landi, 2013; Rastle & J. S. H. Taylor, 2018).

Nevertheless, some reasons for arguing that learning does not take place in context might rest on the fact that that performance declines when participants read the words in isolation for the first time at transfer. However, if we use word reading accuracy in isolation at Trial 1 as a proxy baseline measure, then we can see that the same participants accurately read approximately 40-50% of words (across set size) in isolation at Trial 1, and accurately read 77-79% of context words (across set size) when they saw them for the first time in isolation at transfer. From this point of view, if learning had not occurred during context training, reading scores at transfer would necessarily be similar to the baseline accuracy scores in the isolation condition at Trial 1.

Our participants maintained their performance from training to transfer in the isolation conditions—no decline in word reading accuracy from Trial 6 to transfer was observed (also see Archer & Byrant, 2001; Landi et al., 2006). Thus, what participants learned during isolation training was retained. At the same time, performance was lower compared to the context condition because, even with the decline in reading accuracy experienced from context training to transfer, participants did not learn as many words in isolation relative to context training. This resulted in a context advantage at transfer across set size and especially so, for the small set size. Taken together, the greater degree of improvement in context relative to isolation shows that the orthographic representations formed during training in context were *good enough* to support reading (Martin-Chang, 2017; Martin-Chang et al., 2017).

Nevertheless, it raises the question of why our findings contrast those of other studies that did not find contextual facilitation for regular words (e.g., Wang et al., 2011, 2012, 2013). First, our design was likely more sensitive due to having more target items. This may have permitted us to capture the learning taking place. Even within the small set size, our adult participants trained 24 words compared to previous studies where 4 words were trained. Here, reading accuracy varied between 40% and 80% on the first trial; in contrast, reading accuracy was close to ceiling at Trial 1 in both context and isolation in Wang and colleagues' study (2011). Under these circumstances learning would be difficult to observe.

Second, it is generally acknowledged that reading within the partial alphabetic phase is heavily influenced by the first letter of a word (for review, see Ehri, 2020). The target words trained in previous work each started with a different letter (e.g., vack, smope, praif, broon; e.g., Wang et al., 2011), therefore it is possible that the children's learning was heavily based on the first letter and that they subsequently relied on memory for the rest of the pronunciation. Whereas in the current study, the vast majority (81 of 86) of words shared an onset with at least one other target. On average, each onset was used four times (e.g., kæmʊflɔʃ, kɪtɪn, kʊd, kʌntrɪz). Using items with overlapping initial graphemes reduced the chances that participants were memorizing words based on the onset. Thus, even when implementing a regular orthography, decoding difficulty was higher in the current study.

Lastly, we tested adult participants while Wang and colleagues worked with children (Wang et al., 2011, 2012, 2013). Adults have accumulated more experience laying down orthographic representations compared to children (Nation, 2017). Yet, our experimental design was created to mitigate this issue by implementing an orthography where participants were familiar with only some of the symbols before training. In this situation, participants were artificially placed in the partial alphabetic phase, where they could rely on their partial alphabetic

knowledge to self-teach over successive trials. Future work should explore whether a similar pattern of results is observed with children.

Stimulus Set Size

The contention that stimulus set size impacts self-teaching has been raised before (see Martin-Chang, 2017). But it was not studied programmatically, and thus was left as an open question. We hypothesized that training more words would increase contextual facilitation during training because the task would be more difficult. This was rooted in the contextual facilitation literature showing that when reading is effortful, top-down support is recruited to arrive at the word's correct pronunciation (Simpson et al., 1989; Stanovich et al., 1985; Stanovich & West, 1983). Surprisingly, our data showed that performance was equally high in the two set size groups in context throughout¹. As indicated by the high reading scores, we were unsuccessful at sufficiently increasing task difficulty in context. However, for isolation training we can decisively conclude that reading more words did not increase task difficulty. Indeed, individuals who read the large word set fared better throughout training and at transfer, indicating that they profited from the large set size (Experiments 1, 2, and 3).

In Experiments 1 and 2 the stimulus set size variable was a between-subjects comparison, thus participants in the large set size might be better readers than those selected to for the small set size group. However, we equated the two groups on measures of vocabulary, reading, comprehension, and spelling abilities, which makes this unlikely. Furthermore, the pattern was extended in Experiment 3 where a within-participant design was employed. If further evidence is required, in Experiments 1 and 2, post hoc comparisons simulated a within-subject design by examining the same 24 words learned in the small and

¹ Experiment 3 was replicated in context (Experiment 4) with a new sample and further corroborates the pattern that the effect of context overrides the effect of set size. No comparisons reached statistical significance and thus, are not discussed.

large word sets. Here again, words learned within the large stimulus set size were read with greater accuracy in isolation during training and retention.

Because training with more words led to better word reading accuracy in isolation, we wondered whether there would be evidence of more general orthographic learning in this condition (Conrad, 2008; Conrad et al., 2019; Seidenberg & McClelland, 1989; Zhao et al., 2018). Thus, in Experiment 2 we measured how stimulus set size shaped participants' ability to read individual graphemes before and after training. We found that participants who trained with the larger number of words were able to provide more exemplars of words containing phonemes than those who trained with the smaller number words. Critically, participants from both the large and small set size conditions were exposed to each target word six times over six trials, thus it is not that participants from the large set size had more practice on individual words but, instead they could match and test the correspondence between grapheme and phoneme in a larger number of unique words. These results were extended with the generalization data (Experiment 3). Participants read more new, untrained words with the large set size relative to the small set size.

Transfer to Spelling

It is tempting to conclude that reading in context is a panacea for word learning. However, our data suggest that lexicalization during isolation training is more likely to support fluent spelling recognition and accurate spelling production. It seems that the benefits from reading in context are offset by the lower likelihood that readers notice and consequently encode features of the orthography (Miles & Ehri, 1997). It is the noticing, processing, and encoding that lay down high-quality orthographic representations in memory, and these representations in turn support spelling. This productive struggle (Aiken et al., 2021; Kapur, 2016) resulting from a higher cognitive load while decoding in isolation may be driving attention, which in turn helped participants build up orthographic knowledge that was more precise.

This reflects the supposition by others (Landi et al., 2006; Martin-Chang et al., 2017) that participants formed higher quality orthographic representations when training words in isolation, which lead to better spelling performance several days after training. Lastly, set size did not interact with training condition for the spelling measures in Experiments 1 and 2, however, in Experiment 3 there was a hint of evidence that the isolation advantage for spelling was present when participants practiced reading a larger set of diverse words. Coupling an activity that privileges bottom-up processes, with more opportunities to self-teach, seems ideal for building high-quality lexical representations that support accurate spelling. More work is needed to test this idea.

Limitations and Future Directions

Here, we asked adults to teach themselves to read two fairly large sets of words that were written with an unfamiliar orthography. In so doing, we were the first to examine the moderating effect of set size on the association between contextual facilitation and orthographic learning. Although, we are able to conclude that training in isolation lead to advantages for spelling recognition and accuracy compared to training in context, we are unable to comment how focused word study would match up against reading in isolation. In an educational setting, children are often encouraged to find smaller words in longer words (e.g., *end* in *friend*) or are taught word families (e.g., *mock*, *clock*, *dock*). We imagine these techniques would lead to superior spelling than simply reading in isolation. Future work should investigate the differential benefits of incidental learning and focused word work.

We utilized a counterbalanced design where participants were randomly assigned to the training and set size conditions. This design permitted for differences on outcome measures to be attributed to the variables that were manipulated and not participant characteristics. We posit that including English speakers from diverse language backgrounds strengthened the external

validity of the experiments. Still, assessing individual difference is an interesting avenue for future research to pursue.

Finally, our participants were adults with diverse education background, but the majority held at least an undergraduate degree. Our participants performed at least a Grade 12 equivalent on standardized measures. Thus, it is unclear if similar effects would be detected with individuals with less formal education, or in participants who were still learning to read (i.e., children).

Conclusion and Implications for Practice

Findings from the present study carry implications for reading skill acquisition and instruction. Educators of children and instructors of adult learners should consider word type and set size, and how reading practice takes place. Learning to read individual words is a multifaceted process that involves forming orthographic, phonological, and semantic identities as well as connections between these identities in memory. Our data showed that the training method that facilitated the formation of initial orthographic representations good enough to support word reading may not be precise enough to support spelling.

In planning lessons, teachers need to consider the purpose of the activities they offer to students. When reading acquisition is the goal, practicing in context is time well spent. To refine lexicalization, practicing in isolation is fruitful. Importantly, both methods are likely to benefit regular and irregular word learning (for review on irregular words, see Colenbrander et al., 2020). At the same time, we showed that increasing the number of words to be learned boosted word acquisition in such a way that it compensated for the effects of contextual facilitation.

Our findings tease apart the seemingly contradictory conclusions drawn from the self-teaching literature in context (Share, 2011) and the paradoxical outcome of isolation training (Landi et al., 2006). We add to literature by demonstrating that when word reading accuracy is the outcome of interest, context is superior when readers practice with fewer words (viz., small

set size) however, context and isolation training can lead to similar refinements of lexical representations in memory when exposure takes place over a fairly large number of words. The most obvious place context operates is reading books. Be it fiction or non-fiction, in or out of school, reading provides endless opportunities to process and link pronunciation to meanings and spelling patterns. Once reading acquisition is rooted in the alphabetic principle, engaging in independent reading builds lexical representations (Erhi, 2020). In essence, every time a word is read, its representation is reactivated in memory increasing its lexical quality (Perfetti & Hart, 2002; Perfetti, 2007). A diversity of input akin to print exposure may provide a foundational context to build lexical representations in memory through multiple and repeated exposures (Cunningham & Stanovich, 1990; Nation, 2017; Stanovich & West, 1989; J. N. Taylor & Perfetti, 2016). Complimentary to reading in context is focusing practice at the word level, which may be a boon for words students have difficulty spelling.

Conceptual Replication with Children Participants

The findings of Study 1 revealed that both reading and spelling varied as a function of stimulus set size and training condition, but not in the same ways. Training in context better supported later word reading accuracy whereas, training in isolation better supported later spelling recognition and production. This pattern of results provides clarity with regards to the “...paradoxical relation between reading and learning” (Landi et al., 2006, p. 114). Although participants read and retained many more words from the context condition compared to the isolation condition, the words that they learn from the isolation condition seem to be learned in more detail. Theory suggests that the added detail (i.e., higher quality lexical representations) permitted recognizing as well as producing correct spellings at post-test. This is consistent with the idea that decoding is a productive struggle (Ehri, 2022; Landi et al., 2006; Share, 1995), which may be more salient while training in isolation relative to context.

Without focusing on the orthography, lexicalization may be weaker than expected (Ehri, 2014; Chambrè et al., 2020; Ocal & Ehri, 2017). Some have argued that when reading in context, decoding is partially subverted by top-down processes (Landi et al., 2006). If this is the case, it may explain why training in context did not lead to forming high quality lexical representations (Perfetti & Hart, 2002). So even though participants read fewer words accurately during isolation training, what they did learn served them better than context training when it came fluent spelling recognition and accurate spelling production.

Study 1 also revealed that training more words does not hamper learning. In the case of isolation training, the participants who trained with the large stimulus set size outperformed those who trained with the small set size. This effect was replicated with a within-subjects design; more words from the large stimulus set size were read accurately during training and better remembered at post-test. However, the set size effect proved to be somewhat elusive for the spelling measures. The data hint toward the idea that training a large diversity of words

supports better orthographic learning. The single grapheme test supports this tentative conclusion. Participants provided a word that included the sound that represented the individual grapheme presented to them. The data from this task demonstrated that those who trained with the large set size overtook participants who trained with the small set size. This seems to indicate that training with the large set size bolstered learning grapheme-to-phoneme mappings compared to the small set size.

Exposing adults to phonetic symbols rather than the standard English alphabet accomplished two things. It permitted us to examine whether contextual facilitation would occur during training for regularized words, and it emulated the process of learning how to decode. Despite the insights gained from Study 1, it left open for debate whether similar results could be obtained with children who are in the process of learning how to read. The literature on contextual facilitation suggests that children will likely show a link between context training and reading accuracy, and a link between isolation training and spelling production (e.g., Martin-Chang et al., 2017; Wang et al., 2011). However, there is some uncertainty around the effect of stimulus set size. In most cases, adults will have garnered a lot more reading experience relative to young children, and evidence suggests that reading experience yields dividends in lexical quality (Perfetti & Hart, 2002; Nation, 2017). Moreover, the reciprocal link between reading experience and reading skills gains more traction in adolescence and adulthood (Mol & Bus, 2011; van Bergen et al., 2021). Thus, it is possible that children will not extract the same benefits as adults when training with a large set of words. We also extended our question of whether an effect of context is present for regular words by directly contrasting words with regular and irregular spelling patterns in and out of context. Lastly, because an interaction between training condition and stimulus set size was elusive for spelling, we assessed spelling production before and after training. This permitted a closer look at the trajectory of learning.

Differential Effects of Context and Isolation Training on Orthographic Learning

Monyka L. Rodrigues & Sandra Martin-Chang

Department of Education

Concordia University, Quebec, Canada

Abstract

Across two studies, children in Grade 2 read different sets of words under two experimental training conditions: context and isolation. Training was carried out over 10 trials and was followed by spelling and reading post-tests. In Experiment 5 target words varied in terms of regularity and word length. In Experiment 6, all target words were regular, but children trained with either a small or large stimulus set size. Across both experiments, reading in context boosted reading accuracy during training as well as 1 week later. A different pattern was noted for spelling production; children spelled more words correctly from the isolation training condition. We also found that the greatest advantage was for exception words (Exp. 5). Lastly, we noted that the number of words to be learned (stimulus set size) was a significant moderator; when children had fewer opportunities to map orthographic representations onto pronunciations, training in context was more advantageous than training in isolation. However, reading in isolation seemed to create orthographic representations that were higher in quality, especially for children who trained with the large stimulus set size.

Differential Effects of Context and Isolation Training on Orthographic Learning

For young readers, sounding out words can be an effortful process. However, it is an essential tool for learning how to read (Ehri, 1995; 2020). That is, to become a good reader, a child must understand that graphemes (a letter or string of letters) map onto phonemes (speech sounds). Supporting this view, the self-teaching hypothesis (Share, 1995) states that decoding words grapheme-by-grapheme helps children associate a word's orthography (spelling) with its phonology (sound) and semantics (meaning). Through this process of orthographic learning, children graduate from decoding word-by-word to fluent reading, and words become specified lexical representations in memory (Castles & Nation, 2006; Perfetti, 1992). Once children establish this self-teaching mechanism, orthographic learning proceeds through a combination of oral vocabulary, decoding skills, and frequent reading (Share & Stanovich, 1995).

When children read connected text, they are likely to encounter new and challenging words that require decoding. Contextual facilitation occurs when decoding attempts are validated by seeing the word in context (Stanovich, 1984; Stanovich et al., 1981; West & Stanovich, 1978). Reading in context supports decoding by reducing the number of possible words that fit with the semantic information from the text (Martin-Chang, 2017; Nation & Snowling, 1998; Simpson et al., 1989; Stanovich et al., 1985; Stanovich & West, 1983; Tunmer & Chapman, 1998, 2002). The effect of context is especially salient for novice and struggling readers (Landi et al., 2006; Stanovich et al., 1984; Stanovich et al., 1981; Wang et al., 2011; Wong & Underwood, 1996). Even children reading at grade level benefit from the effects of context when faced with irregular words that violate expected letter-to-sound correspondences (e.g., comb, flood, gross; Nation & Snowling, 1998; Wang et al., 2013; J. S. H. Taylor et al., 2011). However, few studies have examined (1) whether contextual facilitation extends beyond irregular words to other challenging words, such as long words (e.g., scrape, manicure, crocodiles), (2) how the benefits of context compare to reading words in isolation for tasks that

require orthographic learning beyond training, such as for improving later reading and spelling, and (3) how the number of words to be learned impacts orthographic learning. Over two experiments we examined the effects of self-teaching in and out of context. We implemented words of varying length, regularity, and stimulus set size, and measured reading accuracy and spelling production.

Decoding is a necessary skill that all readers must acquire to propel reading development. It is also the skill that established readers fall back on when they encounter unfamiliar or challenging words (Ehri, 2014; Oral & Ehri, 2017). Moreover, learners must train a flexible reading system to be able to read so-called exception or irregular words. These words do not adhere to consistent or frequently encountered letter-sound correspondences (Coltheart & Leahy, 1992; Seidenberg et al., 1984; Waters & Seidenberg, 1985; Strain et al., 2002). Languages with deep orthographies, such as English, include a large number of exception words. Due to the tricky nature of irregular words, decoding them seldom leads to the correct pronunciation (Wang et al., 2011, 2012, 2013). In this situation, readers must recruit other strategies to corroborate decoding attempts. Strategies may include reading by analogy, which involves matching the unfamiliar word seen in print (e.g., *thrown*) to a known word (e.g., *shown*; for review, see Ehri, 1992, 2005, 2014). Contextual facilitation is another secondary, compensatory strategy (Nation & Snowling, 1998). For example, when a reader come across the word *gross* for the first time in print, they might decode it as /gras/ (like, *cross*), then crosscheck that pronunciation against the contextual information (perhaps a story with a boy throwing boogers), which triggers access to the familiar spoken word that sounds like /gros/.

Orthographic Learning in and out of Context

A host of evidence demonstrates that children successfully self-teach novel words in context (Landi, 2013; Martin-Chang, 2009; Share & Stanovich, 1995). For example, in Share's (1999) seminal study second-grade children read nonwords in context and three days later,

completed a spelling recognition and spelling production task. For the spelling recognition task (i.e., orthographic choice task), participants were asked to identify target words amongst three foils. Results indicated that the children read the nonwords with a high degree of accuracy and successfully chose the correct spelling of the targets and avoided the foils. They were also able to accurately spell over half the target words. This provided direct evidence that children could self-teach nonwords from a single exposure during passage reading. In follow-up experiments, exposure to target words were restricted to brief visual exposure along with articulation suppression. Children repeated a nonword (*dubba*) to interfere with phonological processing and, once again orthographic learning was assessed. Blocking the phonological system impaired orthographic learning but did not entirely prevent it. Share (1999) concluded that decoding is a critical mechanism for forming lexical representations in memory and that this process can take place via self-teaching in context.

Within the self-teaching paradigm, others have contrasted context training to isolation training to gain a better understanding of the degree of incidental learning during independent reading (e.g., Cunningham, 2006; Landi et al., 2006; Martin-Chang & Levy, 2005, 2006; Martin-Chang, 2017; Martin-Chang et al., 2017; Nations et al., 2007). For example, Cunningham (2006) asked children in Grade 1 to read four short stories and four texts that had been scrambled to remove contextual support. As expected, word reading accuracy was best for targets in the context condition relative to the scrambled condition. Three days after reading practice, children completed a spelling recognition and spelling production task. The authors found that children could reliably recognize the trained targets but, no performance differences were found between the context and scrambled conditions. Overall, spelling accuracy was low, again with no effect of reading condition. The authors concluded that recognizing correct spellings was less challenging than spelling production, and thus may be a more sensitive measure of newly formed lexical representations. Indeed, across a series of studies Martin-

Chang and colleagues (Martin-Chang, 2017; Martin-Chang & Levy, 2005, 2006; Martin-Chang et al., 2007; Martin-Chang et al., 2017) have demonstrated that “...words learned in context are no more vulnerable to being forgotten than are words learned in isolation” (Martin-Chang et al., 2007, p. 52). These studies support the view that children can self-teach novel words and store, at least rudimentary information, in memory about the words’ orthography for later retrieval.

Still, some researchers have suggested that even though reading is more successful in context that, counterintuitively it may not be ideal for forming highly specified lexical representations (Landi 2013; Landi et al., 2006; Martin-Chang et al., 2017). Due to contextual facilitation, reading in context creates more opportunities for matching word orthography with phonology and semantics. Whereas, when reading out of context the reader must necessarily depend on their decoding skills because that is the only information that is available. As such, readers are more likely to focus more of their attention on the letter-order and the relation between graphemes and phonemes within words. This effortful bottom-up process involved in decoding is thought to be ideal for creating lasting orthographic representations in memory. The drawback, however, is that readers are successful less often when reading in isolation relative to context (Landi, 2013; Martin-Chang, 2017).

Archer and Bryant (2001) investigated the impact of contextual facilitation when words were later transferred to being read in isolation. That is, they asked children to read trained targets words in a list. Children in Grade 1 read a small set of four words in and out of context over three trials. The context was provided by orally presenting short sentence stems. As expected, the authors found that contextual facilitation occurred during training—reading accuracy was best for words practiced in context compared to words practiced in isolation. When considering both the immediate and delayed transfer to isolation task, no differences were found for words learned between the two training conditions. This lead Archer and Bryant

(2001) to argue that there was no difference between learning to read words in context and isolation.

In a similar vein, Landi et al. (2006) found that children had higher word reading accuracy rates in context relative to isolation. One-week later children read words they previously trained in context and isolation mixed in a new list. The raw numbers replicate Archer and Bryant (2001). However, when the authors conditionalized performance by dividing the number of correct responses during retention by the correct responses during the practice trials, they found that children had retained more words from the isolation reading condition. Landi et al. (2006) posited that reading in isolation helped children to direct their attention to the word's orthography, in turn fostering the link between orthography and phonology. They further proposed that when reading with the support of context, decoding is partially subverted leading to less orthographic learning. Of note, the context condition in both these studies (Archer & Bryant, 2001; Landi et al., 2006) consisted of a missing word at the end of a series of predictable sentences, which may have reduced the reader's attention to the print (Martin-Chang, 2017). This is a far cry from natural reading conditions. Nonetheless, these data signaled some of the first evidence that training with contextual support may not be ideal for forming precise lexical representations.

Using more naturalistic texts during training, Martin-Chang and colleagues (2017) also reported that when word reading accuracy was assessed in isolation one week after training, children's performance was similar between words trained in context and isolation. However, when spelling accuracy was assessed following training, children's performance was best for words trained in isolation relative to context. It is plausible that developing lexical representations that are robust enough to support spelling, requires more attention to the orthographic form of the word.

Word Regularity

Although the ability to read exception words and even regular nonwords words increases with experience and age (Coltheart & Leahy, 1996; J. N. Taylor & Perfetti, 2016; Perfetti & Hart, 2002; Schmalz et al., 2013), longer words are shown to continue to pose challenges, even when they follow typical grapheme-to-phoneme mappings (Martens & de Jong, 2006). Early work by Nation and Snowling (1998) demonstrated that contextual facilitation was greatest when decoding challenges arose from reading exception words. In this study, children between the ages of 7 and 10 read regular and exception words in and out of context. The exception words had irregular spelling patterns (e.g., aunt, brooch, colonel), whereas the regular words were spelled as expected (e.g., shave, glint, chrome). Using spoken short sentences as the context condition and list reading as the isolation condition, children read the same words in both conditions. Participants word reading accuracy was best in context relative to isolation and the effect was especially large for exception words.

In order to assess learning of novel words, Wang et al. (2011) implemented a training paradigm involving four trials. They investigated the impact of context training on the formation of orthographic representations. Children in Grade 2 trained a small set of four nonwords words in context or isolation. The nonwords had either regular or exception spelling patterns (e.g., *vack* trained to be pronounced regularly as /væk/ or irregularly as /vak/). During a preexposure phase, children become familiar with the pronunciation and meaning of the target words thus, adding the nonwords to the children's oral vocabulary. After training in or out of context, children completed a spelling production task and two spelling recognition tasks. During training, an effect of context was present for only exception words. Although the authors found null results for both the spelling production and the orthographic *choice* task, the orthographic *decision* task yielded a context effect, again for only exception words. The decision tasks presented the target words one at a time, and the children were asked whether they recognized the presented

nonword. The authors concluded that children especially benefited from the support of context when decoding is challenging and incomplete. These results are somewhat inconsistent with previous accounts of the context effect. To shed light on the different pattern of results throughout the field, in Experiment 5 we asked children in Grades 1 and 2 to read both regular and irregular words, in an out of context. After a retention period, we assessed word reading accuracy and spelling production.

Word Length

We also investigated word length as a potential moderator of the effect of context on orthographic learning for regular words. Longer words can be more challenging to recognize and read than shorter words (Barton et al., 2014). To our knowledge, no studies have systemically tested word length effects for regular words in and out context. For example, Stanovich et al., 1981 examined the interaction between word difficulty and context. Words that children found challenging to read greatly benefited from the presence sentence context, whereas easy words were read equally well with and without meaningful context. The authors proposed that decoding easy words happened so quickly that they were correctly identified before the slower top-down contextual processes could contribute. This contrasted to slower and more laborious reading of difficult words where bottom-up processes took longer to analyze the stimulus thus, permitting secondary, compensatory process of contextual facilitation to pitch in. However, it is possible that the word difficulty effect was driven by a frequency effect. Children read 30 low frequency words with a mean length of 5.7 letters and 30 high frequency words with a mean length of 4 letters. The high frequency words were 4 times more frequent than the low frequency words whereas, the long words were less than two letters longer than the short words (cf. Barton et al., 2014).

Relatedly, both lexical and naming studies report robust word length effects (Gagl et al., 2015; Martens, & de Jong, 2006, 2008; Rau et al., 2014; Barton et al., 2014; Tiffin-Richards &

Schroeder, 2015). For example, Martens and de Jong (2006) presented both real and nonwords to Dutch children. Participants were asked to identify the real words and reject nonwords and the speed of their responses were measured. This lexical decision task has some parallels to the orthographic decision task, except reaction time is also assessed. Participants were fourth-grade students with dyslexia who were matched with both reading age and chronological age controls. The reading age control group were children in Grade 2. Children were presented 40 real words and 40 nonwords in isolation, that ranged between three to six letters. Overall, the word length effect was most prominent for children with dyslexia as well as their reading age counterpart for words with four or more letters. Longer words are presupposed to be more challenging to read because they are being processed sublexically; that is, decoded grapheme by grapheme. In the absence of a word length effect, words are presumed to be processed lexically. That is, once a child is familiar with a word, all letters are processed in parallel—it becomes whole word recognition.

So far, evidence suggest that beginning readers have difficulty decoding both long and irregular words (Joseph et al., 2009; Martens & de Jong, 2006; Perfetti et al., 1979; Wang et al., 2011, 2012, 2013). There is also evidence suggesting that children are more successful at reading and learning exception words in context than in isolation (Nation & Snowling, 1998; Wang et al., 2011). Here we asked whether long regular words would also benefit from self-teaching in context.

Stimulus Set Size

It was recently suggested that experiments on the effect of context can yield different results depending on the stimulus set size—that is, number of words to be learned (Martin-Chang, 2017; Martin-Chang et al., 2017; Rodrigues et al., submitted). A large portion of the training studies thus far implemented sentences or very short passages containing very few target words. However, children in Grade 2 often read short chapter books (e.g., *The Magic*

Tree House, or Junie B Jones) that range in length from 5000 to 12000 words (<https://mascotbooks.com/not-by-the-book-blog/be-mindful/>). When considering the recommended independent reading level of 95% accuracy (Shanahan & North Central Regional Educational Lab, 2005; Tompkins, 2014), it places children learning to read between 20 and 60 new words each book; at the recommended instructional reading level of 90-94% accuracy, the number of words to learn increases further. Moreover, outside of direction instructions, reading provides endless opportunities to process and map spelling onto pronunciation and meanings. In fact, children who read 15-to-20 minutes a day will have processed 5.7 million words by the end of high school (Renaissance Learning, 2018). Once reading acquisition is rooted in the alphabetic principle, engaging in independent reading builds lexical representations (Ehri 2020; Share, 2011). Thus, it seems relevant to ask whether contextual facilitation is impacted by the number of words to be learned (i.e., stimulus sets size). In the present study, we extend the adult data presented in Study 1 (Rodrigues et al., submitted) by experimentally manipulating the number of words read by young readers in Grade 2.

Rodrigues et al. (submitted) systematically tested the effect of stimulus set size on various reading and spelling outcomes with adult participants learning to read words written in an unfamiliar orthography. It was plausible that learning to read and spell 10 or fewer words would not carry the same cognitive load as learning to read and spell a large set of 80 or more words. Rooted in the contextual facilitation literature, they expected that training more words would increase reliance on context because the task would be more difficult. Decades of evidence shows that when reading is challenging, top-down support from context is levied to supplement decoding attempts (e.g., Simpson et al., 1989; Stanovich et al., 1985; Stanovich & West, 1983). Instead, Rodrigues et al. found an effect of stimulus set size for words trained in isolation. The data showed that during training and when target words were retested in isolation after a delay (transfer-retention task), performance was best for words trained with the large set

size. No effect of stimulus set size was found for word trained in context. This may suggest that increasing the number of words does not moderate reliance on context for word identification. It also suggests that when participants could only rely on their decoding skills in isolation, they took advantage of the large number of unique words by matching and testing the correspondence between unfamiliar graphemes and phonemes. Our confidence in the interpretation that training with the large stimulus set size was beneficial for learning is bolstered by the generalization data presented by the same authors (Rodrigues et al., submitted). They found that participants read more new, untrained words that contained graphemes from the large set size relative to the small set size on a transfer task.

The Present Study

The objectives of the present study were to assess whether reading words that varied in regularity, length, and stimulus set size, with and without context, differentially affected gains in children's reading and spelling accuracy. To create a situation in which target words were part of children's oral vocabulary, children were presented with a set of real words, which they knew verbally, but less likely to know in writing. In Experiment 5, three word-types were used: regular-short, regular-long, and irregular. Regular words followed expected spelling patterns and could be decoded. Exception words, on the other hand, had at least some graphemes that violated conventional spelling patterns. In Experiment 6, only regular words were implemented, but we manipulated the number of words to be learned (i.e., stimulus set size).

We hypothesized that the exception words would benefit the most from context training because they would be challenging to read. We expected regular-short words to benefit the least from context training, and that regular-long words would fall somewhere in between. Based on the findings of previous studies (Snowling & Nation, 1998; Wang et al., 2011), retention was expected to be equivalent for context and isolation conditions for regular words but not exception words. Based on the work presented by in Study 1 (Rodrigues et al.,

submitted), we expected that children who trained with the large stimulus set size would close the gap between context and isolation training. Whereas children who trained with the small stimulus set size would demonstrate a context effect for word reading accuracy during and after training. With regards to spelling, it was expected that exception words would be the most difficult to spell, followed by long words, and short words, and that spelling production would be best for words train in isolation for all word types. We also expected an isolation advantage in Experiment 6 but, that this advantage would be washed out in the large stimulus set size.

Experiment 5

Method

Participants

Forty children (female = 22) were recruited. The mean age was 7 years 7 months (ranging from 6 years 10 month to 8 years 5 months). The Wide Range Achievement Test–Third Edition (WRAT3; Wilkinson, 1993) demonstrated that children’s reading ability were at or above grade level ($M_{raw} = 29.35$, $SD = 5.08$).

Design

A 2 (contextual information: context vs. isolation) $\times$ 3 (word type: irregular vs. regular-short vs. regular-long) within-subjects design was implemented. There were two main dependent variables. The first dependent variable was word reading accuracy measured at pretest, during training (over 10 trials), and a post-test after an 8-day delay. The second dependent variable was spelling production measured at pretest and at post-test 1 day after training.

Materials and Procedure

A total of 60 target words were selected. Twenty were irregular words selected from Strain, Patterson, and Seidenberg (2002) with a mean length of 4.95 letters ($SD = 0.83$). Twenty regular words were selected to match the number of morphemes (1 or 2). The mean letter

length of the regular-short words was 4.05 ($SD = 0.60$). The last 20 target words were regular-long words with a mean of 6.70 letters ($SD = 0.98$). Mean comparisons confirmed the word length manipulation between short and long words ($p < .001$). Critically, the exception words were deemed difficult, not due to their length, but because they do not follow expected or common grapheme-to-phoneme correspondence. For each word type, target words varied based on how challenging they were to read and write.

Table 5

Experiment 5: Pretest performance on word reading accuracy and spelling production

	Training condition			
	Context		Isolation	
	<i>M (SE)</i>	95% CI	<i>M (SE)</i>	95% CI
Reading				
Short regular	.74 (.05)	[.64, .84]	.73 (.05)	[.64, .83]
Long regular	.67 (.06)	[.56, .79]	.66 (.06)	[.55, .77]
Exception	.44 (.04)	[.35, .53]	.41 (.04)	[.32, .50]
Spelling				
Short regular	.57 (.04)	[.51, .67]	.59 (.04)	[.51, .66]
Long regular	.27 (.04)	[.18, .36]	.28 (.04)	[.21, .35]
Exception	.18 (.03)	[.11, .25]	.18 (.03)	[.11, .24]

Note. Scores are proportions out of 10.

The pool of 60 target words were divided into two sets of 30 unique words and each set contained 10 words from each regularity and length category (word set G and H; see Appendix L). Two context stories were created with naturalistic passages: story G was created to include word set G and story H was created to include word set H. Story G had 561 words and a Grade 3.2 Flesch-Kincaid reading level. Story H had 489 words and a Grade 2.8 Flesch-Kincaid reading level (see Appendix M). To minimize the impact of potential idiosyncrasies of the material, the target word sets were counterbalanced across context and isolation conditions.

Preexposure Phase. Participants were exposed to the meaning and pronunciation of the 60 target words prior to training. This allowed children to familiarize themselves with words

that may have been previously unknown to them (e.g., mauve, flaunt, and cot). During this phase, a target word was represented by an image. The image (e.g., picture of a formal dress) was concurrently presented orally (e.g., gown) along with a definition (e.g., a gown is a fancy dress; see Appendix N). Each target was presented one at a time, in a fixed random order on the classroom's smart board.

To assess learning, one week later, participants completed a multiple-choice vocabulary test. Thirty (of the 60) words were identified as the most difficult or unfamiliar during the preexposure phase and were selected for the vocabulary test (e.g., scarce, gauge). The vocabulary test asked participants to match a dictated target word to an image. Each target word was accompanied by two foils that were unrelated to the meaning of the target word. Under each image was a letter (i.e., A, B, C), and participants were asked to record their answer by circling the corresponding letter on their response sheet. Images were presented on the classroom smart board in a fixed, random order. No orthographic representations (spelling) were included on the board or response sheet. Participants were then asked to select a sentence that correctly implemented the target word. Two sentences were dictated and designated A and B. One sentence correctly conveyed the meaning of the target word, while the other misused it. Participants were asked to select the letter indicating the correct usage of the target word on their response sheet. Both tasks were administered in a group setting and took 25 minutes.

Pretest. Before the onset of training participants completed the reading subtest of the WRAT3, and a spelling and reading pretest of the target words. The reading pretest took place individually. Children were presented with a list of the 60 target words in a fixed, random order and asked to read them from top to bottom. No feedback was provided, and it took approximately 5 minutes to administer. In a group setting in the classroom, participants spelled dictated target words. Each word was accompanied by a disambiguating sentence and then

was repeated (see Appendix O). The spelling test included the same 60 target words along with four easy words (can, help, his, run) to maintain confidence. The words were dictated in a fixed, random order.

Training Phase. The training phase began two weeks later. During this phase, children were asked to read aloud. To evaluate the effect of contextual information within self-teaching, no feedback was provided. If a child struggled to read a word after a two second delay, they were instructed to move on. During training, children were presented the target words for the first time in print, with half appearing in context and half appearing in isolation. Over five sessions, each target word was presented 10 times, two repetitions per session. Training took place individually in a quiet room at school. The condition order was counterbalanced. Children trained one training condition at a time. Over a two-week period, half the participants trained context and the other half in isolation. Once complete, participants then trained in the condition that they had yet to complete.

During training in isolation, words appeared one at a time in a fixed, random order on a computer screen. Between each word a fixation cross appeared. During each session, participants read the list twice and this took approximately five minutes. For training in context, each target word appeared twice in a story. Context training occurred on paper, with target words printed in red, bolded, and underlined, whereas the rest of the text was printed in black. A shared reading paradigm was employed where the researcher read the story and paused at each target word to allow the child to read it. This took approximately 10 minutes.

Post-tests. One day following the last training trial, participants spelled the target words. Administration followed the same procedure as the spelling pretest except it was administered in small groups. Recall that participants trained one contextual condition at a time, therefore, each participant completed two spelling post-tests—one after context training and another after isolation training. Word reading accuracy was assessed one week following the spelling post-

test (i.e., eight days after the last training trial). This retention task was identical to training except that children read the list and read the story once. This task was administered individually.

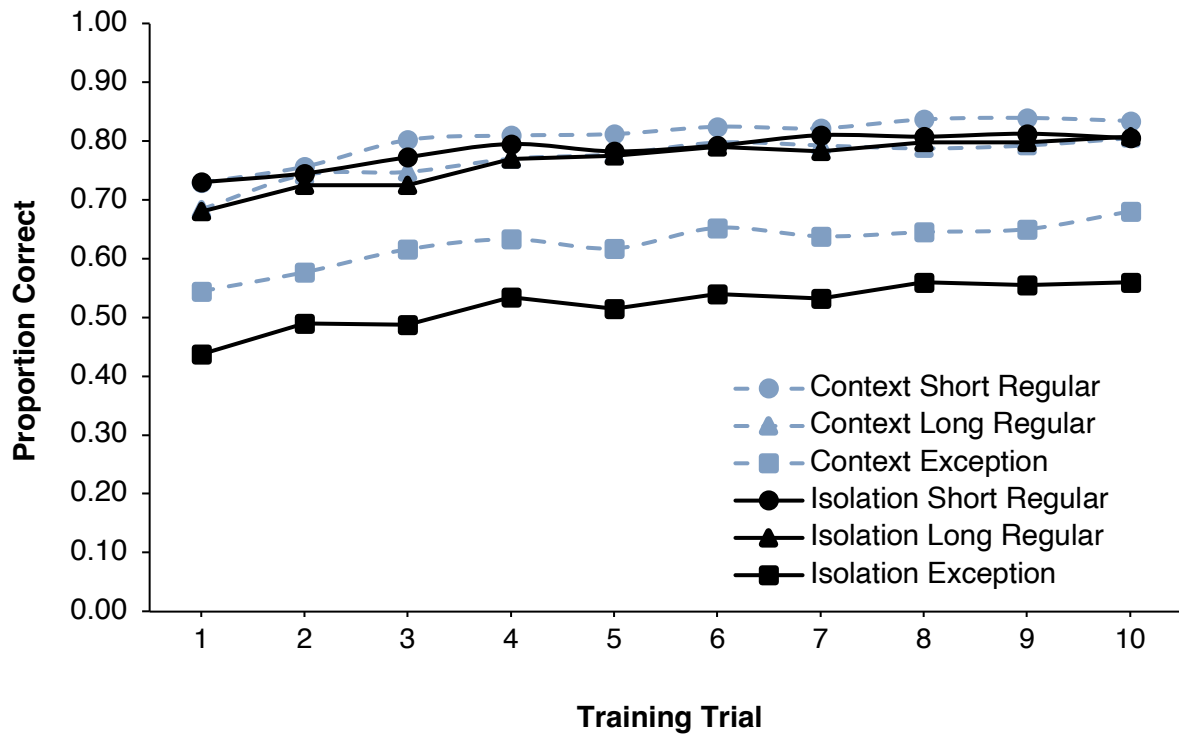
Results and Discussion

Word Reading Accuracy during Training

To determine whether children extracted whole word phonology from exposure to the orthographic forms of whole words without explicit instruction or feedback, a 2 (Training condition: context vs. isolation) $\times$ 3 (Word type: regular-short, regular-long, exception) $\times$ 10 (Training trial: 1–10) repeated measures ANOVA was conducted. All three main effects were significant with small to medium effect sizes: training condition $F(1, 39) = 8.93, p = .005, \eta_p^2 = .19$; word type $F(2, 78) = 72.49, p < .001, \eta_p^2 = .65$; trial $F(9, 351) = 37.12, p < .001, \eta_p^2 = .49$ (see Figure 7). Therefore, across word type and training trials, participants attained better reading accuracy scores when they trained in context relative to isolation. Across training conditions and training trials, participants' word reading accuracy scores were lowest for exception words. Across training condition and training trials, reading accuracy increased over successive trials.

Figure 7

Experiment 5: Proportion of Words Read Correctly during Training



Note. Scores are proportions out of 10.

There was an interaction between training condition and word type, $F(2, 78) = 6.48$, $p = .002$, $\eta_p^2 = .14$, with a small effect size. All other $ps > .52$. To determine whether this interaction remained at the last trial, we carried out a follow-up analysis with Trial 10 as the outcome variable. Indeed, the pattern was corroborated, $F(2, 72) = 7.17$, $p = .001$, $\eta_p^2 = .16$. Further, pairwise comparison revealed that a context effect was present for exception words only. That is, participants accurately read more exception words in context compared to isolation $F(1, 39) = 17.10$, $p < .001$, $\eta_p^2 = .31$, with a small-to-medium effect size. Whereas short and long regular words were read equally well in both context and isolation training conditions ($ps > .28$). Looking within context, participants read more short and long regular words correctly compared to exception words, $F(2, 38) = 14.18$, $p < .001$, $\eta_p^2 = .43$, with a medium effect size. This was also

the case for isolation, $F(2, 38) = 38.74$, $p < .001$, $\eta_p^2 = .67$, with a large effect size. No other comparisons were significant (all $ps > .46$). See Table 6 for means and confidence intervals.

Table 6

Experiment 5: Word Reading Accuracy for Training Trial 10

	Training condition			
	Context		Isolation	
	<i>M (SE)</i>	95% CI	<i>M (SE)</i>	95% CI
Short regular	.83 (.04)	[.76, .91]	.81 (.04)	[.72, .89]
Long regular	.81 (.04)	[.72, .89]	.81 (.05)	[.71, .91]
Exception	.68 (.03)	[.62, .74]	.56 (.04)	[.47, .65]

Note. Scores are proportions out of 10.

As is the case with null results, we cannot firmly conclude whether regular words benefit from context training. However, under these circumstances, we did not see differences between context and isolation. Like Wang et al. (2011), we observed that regular words were read with a high degree of success in both training conditions. By Trial 2, children in Wang's study could read regularized nonwords equally well in context and isolation. In the present study, at Trial 1, word reading accuracy for regular words was approximately 70% in both context and isolation and ended just above 80% by Trial 10.

This raises the question as to why these results contrast with the adult training data from Study 1 (Rodrigues et al., submitted). Although the target words were regularized in Study 1 (i.e., participants could rely on decoding grapheme-phoneme correspondences), learning the artificial orthography might have proved challenging enough for adults to activate the support from context. For example, in Experiment 1 isolated word reading accuracy started at 50% for the small stimulus set size of 24 words. This translates to a learning potential of 12 target words, whereas the children in the current study started training at 70% accuracy which left room for learning approximately three words for the stimulus set size of 10 words per category. Indeed, here we sought to train words that were neither at ceiling for reading nor at floor for spelling

prior to training. Despite best efforts, reading performance for regular words was high at pretest, which may have limited our ability to detect an effect of context; perhaps the words needed to be longer (Cohen et al., 2003; Spinelli et al., 2005). Another difference with the experiments presented in Study 1 is the stimulus set size, it may have needed to be larger to be sensitive to change (Guitard et al., 2018).

Our also data demonstrate that children benefited from context for exception words, replicating findings by Wang et al. (2011) and Nation and Snowling (1998). Likely due to the tricky nature of exception words, context supported partial decoding attempts. For example, a participant in Experiment 5 misread the exception word *comb* as *kalmb* in the sentence, “She started to comb my hair...”, but corrected the pronunciation to *kom* once they read *my hair* (MacKinnon, 2015). When children partially decode a word, contextual information redirects them to the correct pronunciation of a word that is part of their oral vocabulary.

Word Reading Accuracy during Retention

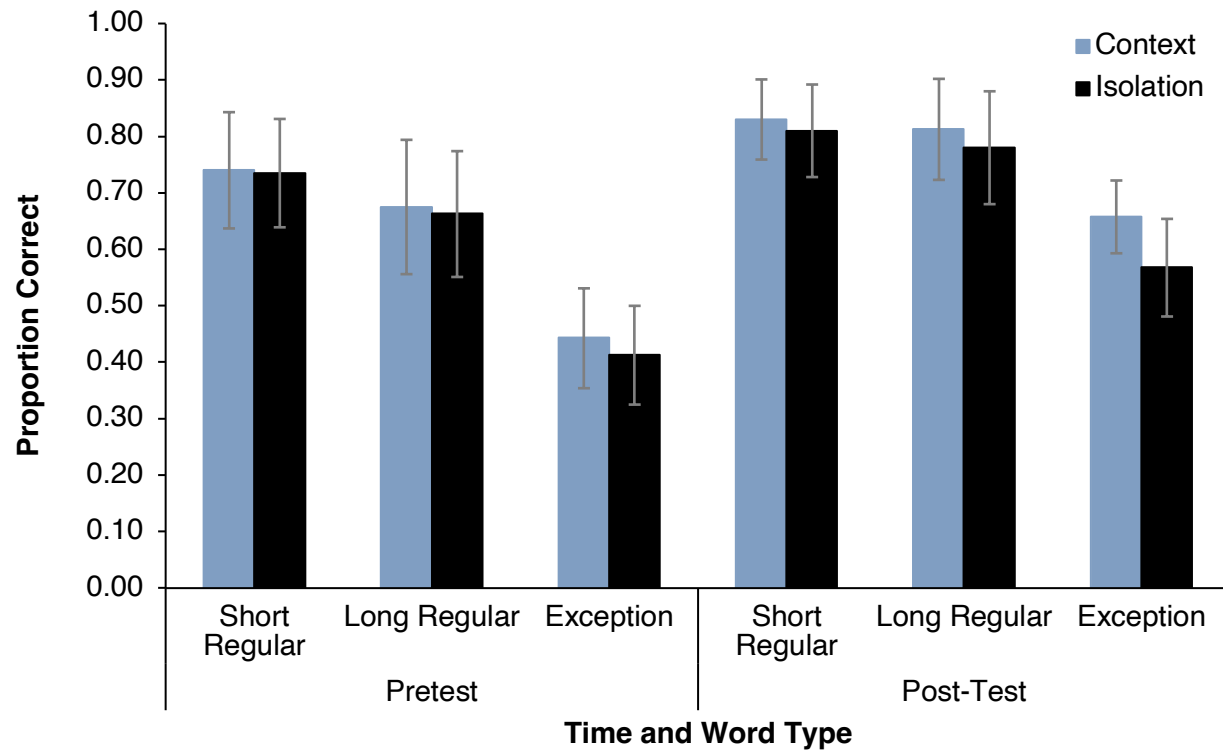
Based on previous research we expected that exception words would continue to display an effect of context beyond training and be measurable during retention. We also expected that long regular and exception words would be read less well than short regular words, as well as read with more success in context relative to isolation. Thus, we ran a 2 (Time: pretest vs. post-test) $\times$ 2 (Training condition: context vs. isolation) $\times$ 3 (Word type: short, long, exception) repeated measures ANOVA (see Figure 7). All three main effects were significant with small to large effect sizes, time $F(1, 39) = 43.09, p < .001, \eta_p^2 = .53$, training condition $F(1, 39) = 5.06, p = .03, \eta_p^2 = .12$, and word type $F(1, 39) = 100.26, p < .001, \eta_p^2 = .72$. These main effects reflect improved performance from pretest to post-test, increased word reading accuracy for words trained in context relative to the isolation condition, and that performance differed between the word types. There was a significant Word type $\times$ Time interaction, $F(1, 39) = 8.19, p < .001, \eta_p^2 =$

.17 with a small effect size. No other interactions were significant (all $ps > .09$). To answer our research questions, we then turned to pairwise comparisons. When focusing on the post-test data, there was no effect of context for short and long regular words ($ps < .20$). However, children read more exception words correctly after context training compared to the isolation, $F(1, 39) = 9.46, p = .004, \eta_p^2 = .20$. Lastly, children read short and long regular words equally well ($ps > .61$), whereas children were more successful reading short ($ps < .001$) and regular long words ($ps < .001$) than exception words in both context and isolation.

When considering the effectiveness of training, it becomes clear that children's word reading accuracy improved for each word type after both context and isolation training. This diverges from Landi et al. (2006) who found that words trained in context were poorly retained relative to words trained in isolation. Whereas we found that short and regular long words were read equally well after a retention period, but that exception words trained in context were retained better compared to those read in isolation. A methodological difference between the two studies is that, although both studies implemented a one-week retention period, our task used the same materials that the children had read during training, whereas Landi et al. (2006) tested the trained target words in a combined list. Thus, in Landi's study target words trained in context were transferred to list whereas, words trained in isolation were tested once again in isolation. This discrepancy in outcome measures may explain why both regular and exception words were retained well in our study; they were both trained and tested in the same contextual material.

Figure 8

Experiment 5: Word Reading Accuracy during Retention by Time, Word type, and Condition



Note. Pretest took place in isolation. The training conditions (context and isolation) represents how the target words were trained. Scores are proportions out of 10.

Spelling Production

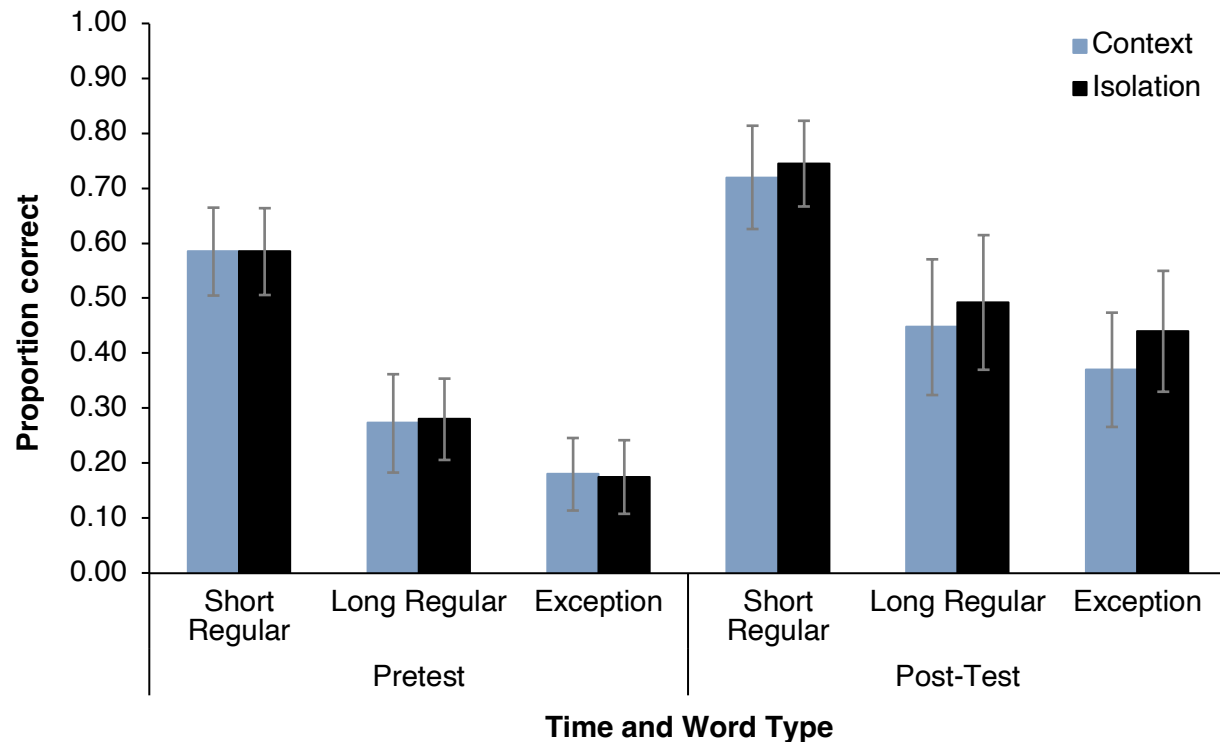
We were also interested in whether children were able to create orthographic representations during training that were precise enough for later spelling production and whether children would acquire more orthographic knowledge incidentally when reading in isolation relative to context. Figure 9 depicts the spelling performance at pretest and post-test, as a function of training condition and word type. A 2 (Time: pretest vs. post-test) $\times$ 2 (Training condition: context vs. isolation) $\times$ 3 (Word type: short, long, exception) repeated measures ANOVA was conducted with spelling scores. All three main effects were significant with small to large effect sizes. Time $F(1, 39) = 64.13, p < .001, \eta_p^2 = .62$, training condition $F(1, 39) = 5.00, p = .03, \eta_p^2 = .11$, and word type $F(1, 39) = 187.18, p < .001, \eta_p^2 = .83$. These main effects reflect

improved performance from pretest to post-test, higher spelling accuracy for words trained in isolation compared to the context condition, and differentiated performance based on word type. Two of the three two-way interaction were significant: Time $\times$ Training condition, $F(1, 39) = 5.77$, $p = .02$, $\eta_p^2 = .13$, indicating that at post-test spelling performance was best for words training in isolation (but was similar at pretest); Time $\times$ Word type, $F(1, 39) = 4.10$, $p = .02$, $\eta_p^2 = .10$, indicating that performance improved from pretest to post-test and performance differed between word type. No other interactions were significant ($ps > .25$).

Pairwise comparisons at post-test revealed that children spelled more exception words correctly after isolation training $F(1, 39) = 8.27$, $p = .006$, $\eta_p^2 = .18$ whereas, regular short and long words were spelled equally well after context and isolation training ($ps > .09$). When comparing word type within training condition, it becomes apparent that isolation training permitted exception words to be learned as well as long regular words. Although, short words remained easier to spell relative to both long regular and exception words. That is, at pretest spelling performance was ranked from highest to lowest from short, long, to exception words ($ps < .02$). At post-test after context training, word type could still be ranked from short, to long, to exception words ($p < .01$). After isolation training, however, spelling performance for exception words rose to meet regular long words and thus, at post-test, only the performance of regular short words differed from regular long and exception ($ps < .001$).

Figure 9

Experiment 5: Spelling Production by Time, Word type, and Condition



Note. Pretest took place in isolation. The training conditions (context and isolation) represents how the target words were trained. Scores are proportions out of 10.

Children's spelling accuracy improved for each word type after both context and isolation training. Although this seems self-evident that performance should improve after training, it should be highlighted that our participants read independently without feedback. Moreover, the children were not alerted to the fact that they would be asked to spell the words after training. And yet, enough incidental learning occurred to transfer from reading practice to improve spelling performances. When looking at the various word types, improvement was best for long regular and exception words. Similar to Wang et al. (2011) we did not find a training effect for short regular four-letter words. In contrast, however, we did find that for irregular words, isolation training was more effective in boosting spelling accuracy than context training. Once again, it is possible we were able to detect an effect of isolation due to using a larger stimulus

set, which perhaps created a more sensitive measure. Another methodological difference that may have led to our results is the number of training trials in the present study; children saw the target words six additional times during training compared to the children in Wang et al.'s study (2011). Perhaps this gave children more chances to encode the orthography of the words trained in isolation. Likewise, Martin-Chang et al. (2017), who also implemented 10 training trials and fairly large stimulus word set of 25 words, found that spelling performance became more precise more efficiently from isolation training compared to context training. This motivates us to conclude that when training a fairly large set of words, isolation training may be better than training in context for developing lexical representations precise enough for spelling.

Experiment 6

Method

Participants

Twenty-five children (female = 13) in Grade 2 were recruited from an English elementary school in a bilingual community. Two participants were excluded from analyses due to extended absenteeism during training. The mean age was 8 years 1 months (ranging from 7 years 7 month to 8 years 6 months). To confirm that participants' receptive vocabulary, word reading, and spelling skills were age-appropriate two standardized tests were administered. The Wide Range Achievement Test—Third Edition (WRAT3; Wilkinson, 1993) demonstrated that children's reading ability were slightly above grade level ($M = 29.46$, $SD = 5.32$) and their receptive vocabulary scores on the Peabody Picture Vocabulary Test—Fourth Edition (Dunn & Dunn, 2007) were just below grade level ($M = 118.92$, $SD = 13.94$). To ensure that participants in the two conditions (small stimulus set size vs large stimulus set) were similar in terms of literacy skills, two *t*-test were run and were nonsignificant ($ps < .41$). Thus, differential outcomes in orthographic learning following the practice phase should not be attributed to pre-existing differences vocabulary knowledge or word reading ability.

Design

A 2 (contextual information: context vs. isolation) $\times$ 2 (set size: small vs. large) mixed-subjects design was implemented. Contextual information was the within-subject manipulation; every participant trained target words in both the context and isolation condition. Stimulus set size was the between-subject manipulation. Participants were randomly assigned to either the small or large set size condition. The condition order was counterbalanced. The dependent variables were word reading accuracy during training and retention, and spelling production. A priori power analysis was conducted based on Experiment 1 with adults (Study 1) in G*Power (Faul et al., 2007). It suggested a sample size of 80 participants to achieve a medium effect size $f(1.25)$, $\eta_p^2 = .61$, with 80% power, and a total of 40 cells in the design (2 Set size: 24 vs. 86 $\times$ 2 Training condition: context vs. isolation $\times$ Trial 1-10). However, as indicated in the Participant section, this sample size was not obtained. This study occurred at the tail-end of the pandemic when social distancing measures were gradually being decreased. Thus, when the schools were once again available to us, we were eager to implement classroom-based research, but also understanding that not all teachers or parents would be comfortable having their children work with individuals from outside of the school community.

Materials and Procedure

A total of 168 decodable target words were selected. That is, they followed common grapheme-to-phoneme correspondences. Target words were between 3–to–10 letters ($M = 5.65$, $SD = 1.56$) with a maximum of four syllables ($M = 1.59$, $SD = .67$). They were nouns, verbs, and adjectives, but not function words (e.g., prepositions) or other high-frequency words (e.g., determiners). Further, target words did not overlap in root morpheme (e.g., *place* and *placed*).

The pool of 168 target words were divided into two sets of 84 unique words: word set I and word set J. The target words varied by word frequency and were balanced across word sets ($p = .30$). To create the small word set, a subset of 24 words were selected from each word set I and word set J. These two sets of target words also varied in word frequency and were balanced across word sets ($p < .42$). The target words parameters are available in Appendix P.

Two context stories were created to mimic naturalistic passages (Story I and J), and these stories included Word Set I and J, respectively. Story I had a total of 168 words and a Grade 3.6 reading level. Story J had a total of 187 words and a Grade 4.0 reading level (see Appendix Q). To minimize the impact of potential idiosyncrasies of the target word sets or stories, the target word sets were counterbalanced across contextual information conditions.

Pretest. Participants completed a spelling pretest of the 48 target words from the small set size. In a group setting in the classroom, participants spelled dictated target words. Each word was accompanied by a disambiguating sentence and then was repeated. The words were dictated in fixed, random order and took approximately 30 minutes to administer.

Training Phase. The training phase began five days later. During training, children read aloud, and no feedback was provided. If a student struggled to read a word after a two second delay, they were told they could skip it and keep reading. Over five sessions, each target word was presented 10 times, two repetitions per session. Training took place individually in a quiet room in the students' school. Participants were presented with both training conditions (i.e., context and isolation) during each session.

During training in isolation, participants in the small stimulus set condition were presented with a list of 24 words typed in rows on a sheet of paper. Participants in the large stimulus set condition, were presented the same page in addition three more pages with 20 target words listed in the same format on each. During each session, participants read the list twice. Context training also occurred using a paper format. Participants read the story from

beginning to end. Children in the small stimulus set condition, read the first paragraph of the story (which contained 24 target words), whereas participants in the large stimulus set condition read the entire story, which contained 84 target words.

Post-tests. Immediately following the last training trial, participants spelled the target words. Administration followed the same procedure as the spelling pretest except it was administered individually. One week following the spelling post-test, participants read all the target words randomized in a new list (the transfer-retention task). Participants who trained the large set size, read 168 target words. Those who trained in the small set size, read 48 target words.

Results and Discussion

Word Reading Accuracy during Training

To examine performance across conditions and trials, we conducted a 2 (Set size: 24 vs. 86 words) $\times$ 2 (Training condition: context vs. isolation) $\times$ 10 (Training trials 1–10) mixed ANOVA. Mean comparisons were carried out with Sidak adjustment for multiple comparisons. The main effect of training condition, $F(1, 21) = 21.35$, $p < .001$, $\eta_p^2 = .50$, was significant with a medium effect. This indicates that across set size and training trials participants read more words accurately in context than in isolation. The main effect of training trials, $F(9, 189) = 25.26$, $p < .001$, $\eta_p^2 = .55$, was also significant with a medium effect. Thus, across set size and training condition, reading accuracy increased over successive trials (see Figure 11). This provides further evidence children can more easily read with the support of context than without (Martin-Chang, 2017; Martin-Chang et al., 2017). All other F s were not significant ($p > .16$).

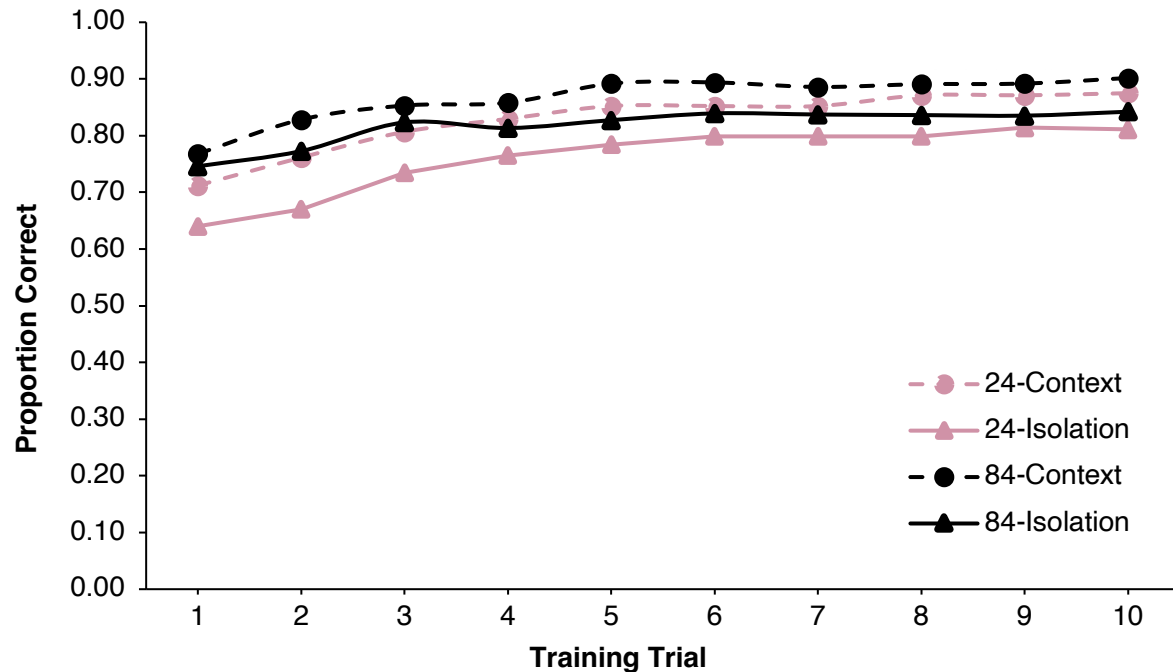
What may explain the null finding for stimulus set size? According to G*Power (Faul et al., 2007), the experiment was under-powered due to insufficient sample size. A hint of a stimulus set size effect was present at Trial 1. Children's word reading accuracy was between

70 and 80% for all conditions except the small stimulus set size in isolation, which was at a 64% success rate. Then, by Trial 10 the means between the two set size groups were very close in value. So, perhaps there is a reason other than a power issue for the lack of a stimulus set size effect.

There are a number of differences between the experiments presented in Study 1 and this experiment. Here, children read very short passages whereas, adults read target words written in an artificial orthography embedded in normal text. To more closely emulate the methodology of the adult study, we could have used a shared reading paradigm, where children would have been asked to read only the target words and the experimenter would have read the remainder of the text. This would have more closely replicated Martin-Chang (2017) where context training involved children reading 85 words within stories containing 300-400 words. However, because of our interest in stimulus set size, we elected to keep the stories as similar as possible to the stimulus set size from which they were derived. A secondary benefit to asking children to read the story in its entirety, as opposed to employing shared reading, was that the manipulation was more akin to print exposure whereby some children consume more print than others. Finally, because our target sets involved reading real words, it is possible that the children were familiar with their orthographic forms. All of these differences in combination make comparing the results between Study 1 and Study 2 more difficult.

Figure 10

Experiment 6: Proportion of Words Read Correctly during Training



Word Reading Accuracy during Retention

To assess whether participants could access the target words fluently without resorting to laborious decoding (recall that they were only given two seconds to read each word) or requiring scaffolding from context, participants read all the targets words they had previously trained in a new list (isolation). We conducted a 2 (Set size: 24 vs. 86) $\times$ 2 (Training condition: context vs. isolation) mixed ANOVA with Sidak adjustment. The main effect of training condition, $F(1, 21) = 4.32$, $p = .05$, $\eta_p^2 = .17$, was significant with a small effect (see Table 7). This denotes that across set size, participants' word reading accuracy was best for target words originally trained in context. This echoes the context effect found during training, that is, reading performance was best for words trained in context regardless of set size. Moreover, the fact that word reading performance during retention was best for words trained in context suggests that self-teaching during context training provisioned reading accuracy during retention task. The Set

size $\times$ Training condition interaction was at trend with a small effect, $F(1, 21) = 3.92$, $p = .06$, $\eta_p^2 = .16$. Pairwise comparisons revealed that participants read more words accurately that were originally trained in context compared to isolation in the small set size, $F(1, 21) = 7.89$, $p = .01$, $\eta_p^2 = .27$ but, not in the large set size ($p = .94$). All other F s were not significant.

Table 7

Experiment 6: Proportion Correct for Conditions and Tasks

Tasks	Stimulus set size			
	Small 24		Large 84	
	<i>M (SE)</i>	95% CI	<i>M (SE)</i>	95% CI
Training trial 1–10				
Isolation (list)	.76 (.04)	[.68, .85]	.82 (.04)	[.73, .90]
Context (story)	.83 (.03)	[.76, .90]	.87 (.03)	[.80, .94]
Training trial 10				
Isolation (list)	.81 (.04)	[.73, .89]	.84 (.04)	[.77, .92]
Context (story)	.87 (.03)	[.81, .94]	.90 (.03)	[.84, .96]
Retention in list				
Isolation (list)	.81 (.04)	[.74, .89]	.86 (.04)	[.78, .93]
Context (story)	.86 (.04)	[.77, .93]	.86 (.04)	[.78, .93]
	<i>n</i> = 11		<i>n</i> = 12	

Note. Scores are proportions out of 24 or 84.

This pattern of results partially replicates the findings reported in Experiments 1 and 2 of Study 1 (Rodrigues et al., submitted). Rodrigues et al. found two pieces of converging evidence pointing to the effect of context and the moderating effect of stimulus set size. Here we found only one. When comparing word reading accuracy between the large and small stimulus set size, Rodrigues et al. found that participants in the large set size outperformed those in the small set size for words originally trained in isolation but, not in context. They also reported that when comparing word reading accuracy between context and isolation, participants read more words accurately that had been trained in context compared to isolation in the small set size

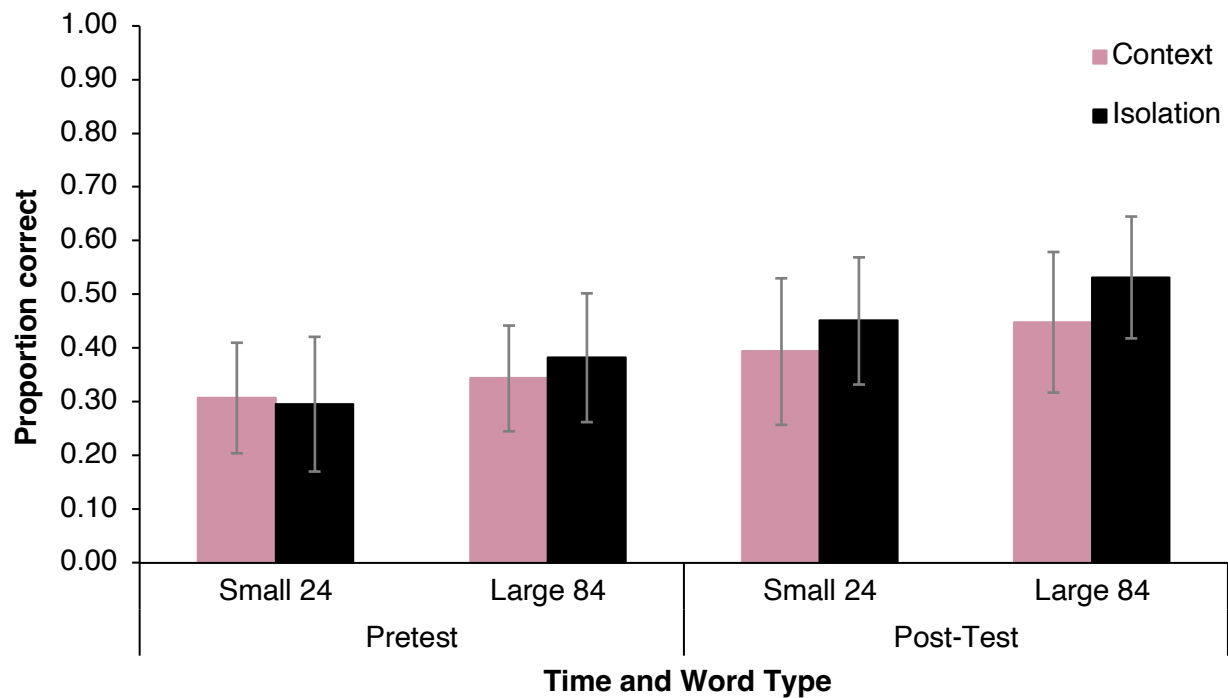
but, not in the large set size. Here, in Experiment 6 with children participants, mean differences reached statistical significance for only the latter pattern, namely there was an effect of context for the small stimulus set size and not the large set size. Thus, our results still point to the notion that when a reader has fewer opportunities to map orthographic representations onto pronunciations, training in context is more advantageous than training in isolation.

Spelling Production

To examine whether orthographic knowledge differed on a spelling task after training, we conducted a 2 (Set size: 24 vs. 86) $\times$ 2 (training condition: context vs. isolation) $\times$ 2 (Time: pre and post) mixed ANOVA. There was a main effect of time, $F(1, 21) = 44.16$, $p < .001$, $\eta_p^2 = .68$, indicating that across set size and training condition, participants spelled more words correctly after training (see Figure 11). The Time $\times$ Training condition interaction was at trend with a small effect, $F(1, 21) = 4.24$, $p = .06$, $\eta_p^2 = .16$. All other F s were not significant ($p > .41$). When probing the interaction, it can be seen that from pretest to post-test the effect size for words trained in isolation is more than twice as large, $F(1, 21) = 72.29$, $p < .001$, $\eta_p^2 = .77$, than that of the words trained in context, $F(1, 21) = 11.99$, $p = .002$, $\eta_p^2 = .36$. This suggest that there was more improvement from Time 1 to Time 2 in isolation compared to context. In practical terms, it may suggest that children more readily self-taught orthographic representations during isolation training relative to context training. Ultimately, however, when looking at only the post-test scores, there was no difference between isolation and context, nor was there an effect of stimulus set size. Numerically speaking, spelling accuracy was best for children who trained with the large set size thus, it may be a power issue.

Figure 11

Experiment 6: Spelling Production by Time and Condition



General Discussion

The aim of the current study was twofold. Experiment 5 explored self-teaching in and out of context to examine whether the effect of context interacted with word length and word regularity. Experiment 6 employed only regular words and examined whether the effect of context interacted with stimulus set size. The findings from the two experiments suggest that young children read target words more accurately when they were presented within meaningful contexts compared to when they were presented in isolation, and yet the effect of context on orthographic learning through self-teaching was moderated by word regularity and stimulus set size.

Experiment 5 showed that assistance from contextual information was most relevant for exception words, whereas Experiment 6 showed that context was more beneficial for

participants training with the small stimulus set size. To measure orthographic learning, two tasks were conducted after a delay: word reading accuracy and spelling production. In Experiment 5, where the target words to be learned varied in terms of regularity and length, a facilitatory effect of context for reading accuracy was observed for exception words during retention but, there was an isolation advantage for spelling production. The results from Experiment 6, in which the target words to be learned were all pronounced in a regular fashion, showed a similar pattern; children who trained with the small stimulus set demonstrated better reading accuracy during retention for words originally trained in the context condition but, an isolation advantage for spelling production was supported by more improvement from pretest to post-test.

What becomes apparent, then, is that a different picture emerges depending on whether the outcome of interest is reading or spelling. For reading accuracy there were benefits of context. For spelling accuracy, more improvement was observed after children practiced reading words in isolation. Where the results diverge from Experiment 5 and 6, is the context effect for words that follow expected or common pronunciation conventions, that is, regular words. Experiment 5 pitted regular and exception words against one another, whereas Experiment 6 only had regular words. What might explain the context effect for regular words in Experiment 6 and the absence of one in Experiment 5? A critical difference between the two experiments is the preexposure phase. In Experiment 5, participants were pre-exposed to the meaning and pronunciation of the target words. Our intent was to familiarize children with words that may not have previously unknown. Moreover, before training started, participants completed both a reading and spelling pretest of the target words. Doing so may have increased familiarity with the words to the point that the support from context was no longer relevant. We sought to address this possibility in Experiment 6 by still selecting words within children's developmental zone, but not teaching the target words ahead of time. This methodological

difference may account for the observed context effect during training. What is more, in Experiment 6 all target words were tested in isolation whereas, in Experiment 5 target words were tested with the same materials. Thus, it is unknown whether the words trained in context in Experiment 5 would have also been read with a greater accuracy than words trained in isolation when transferred to list at post-test.

Nonetheless, the findings from Experiment 5 are in line with previous studies contrasting context and isolation training on the development of orthographic representations. We found that support from context was not critical for orthographic learning for words with regular spellings (Cunningham, 2006; Nation et al., 2007; Nation & Snowling, 1998; Wang et al., 2011). Exception words, on the other hand, are more difficult to read for novice readers (Coltheart & Leahy, 1996); as a result, children drew upon context and reading accuracy improved without receiving corrective feedback. These gains are consistent both with the body of work showing the benefits of repeated exposures and the efficacy of self-teaching (e.g., Ricketts et al., 2011; Zawoyski et al., 2015).

In Experiment 6, the benefits of context training were present for regular words. Specifically, word reading accuracy was best for words trained in context relative to isolation during and after training. While keeping in mind that this experiment was underpowered and that the interaction was at trend, we found some indications that the effect context was limited to the small set size of 24 words. But when children trained with the large set size, no training effect was observed. With more data, this might extend the pattern of results reported with adults in Study 1 (Rodrigues et al., submitted). Again, our data showed that increasing the number of words to be learned does not increase reliance on context.

With that said, there remains a notable difference between the work presented here and Landi et al. (2006); words children trained in context were remembered just as well or better compared to words trained in isolation. A context advantage for exception words was found with

a retention task where children re-read the same materials one week later (Experiment 5) and in the retention task where children read all regular target words in a mixed list one week later (Experiment 6). Indeed, when applying the conditionalization formula implemented by Landi et al. (2006), we find that the lowest retention score to be 95% (Experiment 6, context retention score divided by Trial 10 score yields this equation, $.86/.90 = .95$). What might explain the differing pattern of results observed between our findings and those of Landi et al. (2006)? We implemented 10 training trials, whereas Landi et al.'s design consisted of either one trial (Experiment 1) or three trials (Experiment 2)—it is possible that having at least seven more opportunities to learn a word's orthographic representation leads to more consolidation in memory. Moreover, the material making up the contextual information may have been richer in our current work. We implemented stories that mimicked naturalistic texts common for children in Grade 2, whereas Landi et al. (2006) asked children to read individual predictable short sentences. This may have reduced or limited the activation of contextual information.

The benefits of contextual facilitation can be far reaching. When reading is effortful, children can use context to support decoding attempts. For example, when children achieve partial decoding and pronounce *island* like *izland*, a story about a sea turtle's voyage to lay her eggs can guide them to the correct pronunciation. Differences in word reading accuracy between context and isolation condition were largest early at the beginning of training. Indeed, past research corroborates this robust initial advantage of context (Archer & Bryant, 2001; Landi et al., 2006; Martin-Chang, 2017; Nation & Snowling, 1998; Nicholson, 1991; Wang et al., 2011). This initial boost may affect novice readers in a positive manner. It is plausible that the greater success experienced when reading in context might fuel young readers and motivate them persevere even when reading challenging (Becker et al., 2010; Nation, 2017; Mol & Bus, 2011). Once children successfully decode a new word, subsequent encounters may require less effort (Ehri, 2014), which in turn may support the development of positive reading experiences

(Allington, & McGill-Franzen, 2021; Martin-Chang et al., 2017; Merga, 2017; Mol & Bus, 2011; Sparks et al., 2014). Thus, reading with the support of context may have a positive cascade. Indeed, decades of research suggests that greater reading accuracy is a critical component of fluency and supports comprehension which make reading more pleasurable (Mol & Bus, 2011; van Bergen et al., 2018, 2021). This in turn fosters a desire to read more, bolstering reading skills.

Turning to spelling accuracy, differences between regular and exception words were notable once again (Experiment 5). Participants made gains across all word categories however, the greatest improvement was for exception words. This echoes Wang et al. (2011) who found that exception words had the lowest spelling accuracy, demonstrating that exception words were the most challenging to read and spell accurately. Unsurprisingly, short regular words were the easiest for participants to spell, followed by long regular words, and exception words. When contrasting training conditions, we found an advantage for words trained in isolation for exception words (Experiment 5) and regular words (Experiment 6).

Our results corroborate Martin-Chang et al. (2017) who found better spelling results for words trained in isolation than in context. They used a large stimulus set size akin to ours as well as target words that varied in length, regularity, and frequency. An isolation training advantage might indicate that when attention is focused on decoding (without relying on top-down processes for word identification) it benefits orthographic learning, which in turn advantages subsequent spelling production. With respect to stimulus set size (Experiment 6), numerically speaking, the children's data is similar to the adult's data from Study 1. Keeping in mind that the analyses for Experiment 6 were underpowered, there may be an isolation advantage, especially when participants trained with the large stimulus set size of 84 words. If this were the case, it would encourage the view that a diversity of input during training bolsters orthographic learning. Regardless, the findings presented here indicate that participants'

spelling skills improved via self-teaching via reading, and training in isolation was particularly fruitful.

Limitations and Future Directions

Across two experiments we asked children to self-teach the orthography of large sets of words. First, we examined the moderating effect of word regularity on contextual facilitation. Second, we examined the moderating effect of stimulus set size on contextual facilitation. Overall, we can conclude that reading in isolation leads to advantages for spelling accuracy compared to reading in context. This was especially evident in performance growth from Time 1 to Time 2 for regular long words and exception words and when children trained with 30 words or less. However, it leaves unanswered how focused word study matches up against reading in isolation. For example, educators might ask children to sort words that start with the letter *c* (e.g., cat, cot, cell, circus, clue, crayon) based on the sound it makes (i.e., /s/ or /k/). Further instruction and sorting can take place to demonstrate that the letter *c* is pronounced /k/ when it is followed by consonants and the letters *a*, *o*, *u* whereas, it is pronounced /s/ when it is followed by the letters *i*, *e*, *y*. We suspect that this technique might lead to superior spelling accuracy than simply reading in isolation. Future work should investigate the differential benefits of incidental learning versus focused word work.

Experiments 5 and 6 utilized counterbalancing for the within-subjects and mixed-subject designs of. Participants were randomly assigned to the training and stimulus set size conditions which varied across target word sets. Target words were systematically manipulated within and across experiments to target regularity, length, and stimulus set size. Doing so increased the probability that differences on the outcome measures can be attributed to the variables that were manipulated and not participant characteristics. By including English speakers within francophone communities, we strengthened the external validity of the experiments—bilingualism is more and more common. Still, assessing individual difference is an interesting

avenue for future research. Indeed, some findings suggest that reading accuracy of poor readers and English learners may greatly benefit from context (Nation & Snowling, 1998; Nicholson, 1991; Stanovich & Stanovich, 1995; Wong & Underwood, 1996). As such, it would be especially interesting to evaluate spelling as an outcome.

To ensure that word reading accuracy was also fluent, we encouraged children to skip the word if decoding laboured beyond 2 s. It is possible that increasing the allotted time for decoding words would foster deeper learning, especially during the first few trials when self-teaching is more salient. Finally, participants were asked to read to the best of their abilities and once a response was provided, they may have stop analyzing the word's orthography. Making use of eye tracking may be a solution to evaluate how long each word is visually inspected.

Conclusion and Implications for Practice

Self-teaching is argued as a mechanism for creating orthographic representations in memory. The data reported here support and extend this notion. Critically, if we had only measured reading accuracy, we may have missed the fact that reading in isolation supports the development of orthographic representations in memory that are precise enough for spelling. Likewise, if we had only measured spelling production, we may have missed that reading in context supports the development of orthographic representations in memory that are good enough for reading. Thus, our findings suggest that educators should consider not only word type and stimulus set size, but the type of literacy outcome they are trying to foster.

In planning lessons, teachers need to consider the purpose of the activities they offer to students. When reading acquisition is the goal, practicing in context is time well spent. To further refine lexicalization, practicing in isolation is also time well spent. Importantly, both methods are likely to benefit regular and irregular word learning. At the same time, we showed that increasing the number of words to be learned boosted word acquisition in such a way that it compensated for the effects of context and isolation training.

The most obvious place context operates is reading books. Be it fiction or non-fiction, in or out of school, reading provides endless opportunities to process and link pronunciation to meanings and spelling patterns. Once reading acquisition is rooted in the alphabetic principle, engaging in independent reading builds lexical representations (Erhi 1992, 2005, 2020; Share, 1995, 2011). In essence, every time a word is read, its representation is reactivated in memory increasing its lexical quality (Perfetti & Hart, 2002; Perfetti, 2007). A diversity of input akin to print exposure may provide a foundational context to build lexical representations in memory through multiple and repeated exposures (Cunningham & Stanovich, 1990; Nation, 2017; Stanovich & West, 1989; J. N. Taylor & Perfetti, 2016). Complimentary to reading in context is focusing practice at the word level, which may be a necessary step for words students have difficulty spelling.

Dissertation Discussion

Most would agree that the majority of reading takes place in context—be it, storybooks, novels, educational textbooks, technical manuals, or your first vegetarian recipe. This is in part why researchers have spent a great deal of time and resources investigating the effects of reading in context in the controlled settings of the laboratory. In the late 1970s, Perfetti (Perfetti et al., 1979, Perfetti & Roth, 1981) as well as Stanovich (West & Stanovich, 1978; Stanovich & West, 1979) and their respective colleagues uncovered what is now known as contextual facilitation. Reading in context relative to isolation, nonsensical, and scrambled texts leads to superior word reading accuracy as well as quicker reading times. Decades later, researchers used correlational research to coin the term *print exposure* (Cunningham & Stanovich, 2001; Stanovich, & West, 1989), which encapsulates the long-term benefits of reading in context. The print exposure literature suggests that habitual reading practice has cognitive benefits that go beyond obtaining meaning from text, and that the benefits of reading (and consequences of lack thereof) are reciprocal and additive in nature (Cunningham & Stanovich, 1997; Mol & Bus, 2011; Share & Stanovich, 2005; Stanovich, 1984). Namely, children and adults who frequently read novels for pleasure possess superior vocabulary, verbal fluency, comprehension, spelling skills as well as cultural and declarative knowledge compared to those who read less (Allington, & McGill-Franzen, 2021; Black & Barnes, 2021; Grolig et al., 2020; Mar & Rain, 2015; Martin-Chang et al., 2021; Tremblay et al., 2020; van Bergen et al., 2021).

Yet, despite this research, some evidence suggests that context is not ideal for developing lexical representations in memory (Landi, 2013; Martin-Chang et al., 2017). It has been suggested that bottom-up decoding is superseded when reading in context. Theory suggests that contextual information diverts attention from processing a word's orthography, which results in forming weaker lexical representations in memory (Landi, 2013). In those cases where orthographic learning benefits were observed for context training, they were found

exclusively for irregular words (Wang et al., 2011, 2012). When reviewing the literature on both contextual facilitation and print exposure in line with my own research, I suggest certain qualifiers should be added.

The aim of my dissertation was twofold. First, I sought to understand how stimulus set size and word regularity interact with the effect of context and isolation, and second, I hoped to tease apart differential impacts based on various reading and spelling outcome measures. Indeed, the focus on understanding the impact of the number of words to be learned (stimulus set size) and the effect of isolation training on spelling recognition and production was particularly novel. In the first study, I investigated whether adults could extract whole word phonology from exposure to the orthographic forms of words without explicit instruction or feedback. In this study I asked adults to read different sets of regularized words written in a false orthography in and out of context, and I employed multiple delayed post-tests of reading and spelling. In the second study, I examined whether children could also self-teach (i.e., learn without explicit instruction or feedback) in and out of context. In this study I asked children to read multiple sets of words written in standard English orthography but that varied based on word regularity, word length, and stimulus set size. Like the adult participants, children were asked to read and spell target words after a delay.

Theoretical Implications

The findings of this dissertation contribute to the growing body of literature on the formation of orthographic representations resulting from self-teaching in context and isolation. Although contextual facilitation has been studied in the past (e.g., Martin-Chang, 2017; Wang et al., 2011), the independent variable of interest here, namely, stimulus set size, was overlooked. Previous studies also failed to consider the impact of outcome variables when interpreting their findings. By systemically manipulating the number of words to be learned (stimulus set size) and carefully outlining the constructs that each dependent variable measured, the findings

of this dissertation broadened the field's understanding of (1) how the various outcome measures tap the continuum between orthographic representations that are good enough for reading and precise enough for spelling, (2) the contribution of context and isolation training on forming lexical representations, and (3) how key factors, like stimulus set size, can moderate the formation of representations in memory.

In alphabetic languages, like English, learning to read words rests upon decoding. Learning how letters (graphemes) map onto sounds (phonemes) is critical for developing the skills required to determine what a word sounds like from what it looks like in print. Like many skills, decoding takes time to improve and requires direct instruction and practice. Initially, decoding attempts are partially correct and effortful but, through the process of applying knowledge of grapheme-phoneme relations, decoding skills improve and reading becomes progressively more fluent. According to the phase theory of reading development (Ehri, 1995, 2005), decoding is a connection-forming process; translating words from print to sound provide opportunities for forming word-specific representations in memory. These include pronunciation, meaning, and spelling representations. The three facets of the representation become integrated in memory and support fluent word recognition (Ehri, 1995; Perfetti & Hart, 2002; Share, 1995).

Based on this theory, I created a situation in which participants could translate graphemes to phonemes of target words over a series of trials. Multiple exposures increased the number of opportunities participants were given to make connections between print and sound, and thus in theory, more opportunities for creating lexical representations in memory. Although there is evidence to suggest rudimentary representations can form after just one exposure (Share, 2004), I was interested in assessing repeated exposures—a situation more akin reading practice in the classroom. Thus, I experimentally created a situation where both adults and children were asked to self-teach words. Adults trained with words written in an

artificial orthography whereas, children train words written in standard English appropriate for their grade level. This accomplished two things. First, adults were faced with some graphemes they had little to no prior experience with; this experimental design allowed me to assess orthographic learning. Whereas children read words they likely knew the definitions of, but had little practice with in writing. Second, the target words were regularized (adults) or regular (children)—there was a one-to-one correspondence between grapheme and phoneme. This experimental design permitted me to assess contextual facilitation for regular words.

Word learning from independent reading is explained by Share's self-teaching hypothesis (1995). But for self-teaching to occur readers must possess at least some rudimentary decoding skills. The reader must be able to decode at least some grapheme into their phonological form to build word-level orthographic representations in memory. Through repeated encounters, a word becomes lexicalized, reducing the reliance on decoding for its identification. The lexical representation will then be available on future encounters.

To remain within the self-teaching paradigm, participants were instructed to read as best they could without external feedback from an expert. It also allowed us to retain some face validity. Educators cannot explicitly teach every word a child will encounter in text. Likewise, adults, learn most new words independently. They may hear a professor pronounce the word but, beyond that, they are responsible for learning how to read and spell new words. Employing the self-teaching paradigm also permitted me to contrast the benefits of training in context and isolation without feedback. Self-teaching in context differs from self-teaching in isolation. In context, word reading accuracy can be scaffolded by meaningful text. Whereas, in isolation word reading accuracy rests upon bottom-up skills. Contrasting the two allowed examining which cognitive mechanism is best for forming lexical representations for later reading and spelling.

The formation of lexical representations is indexed as per the lexical quality hypothesis (Perfetti & Hart, 2002). Quality is defined as the degree to which a word's representation in memory is detailed by its sound, meaning, and spelling. A word that has a high quality representation in memory contains phonological, semantic, and orthographic information that are tightly bound together. Not only that, a word with a high quality representation has fully specified identities that are “more stable and less context-bound than those of lower quality” (Nation, 2019, p. 2). As a result, they support accurate and quick (i.e., fluent) word reading, which frees cognitive resources for comprehension (Perfetti, 2007). In practical terms, this means that “there are words we know well and others we know less well” (Nation, 2019, p. 2). This continuum can be measured using tasks that tap different levels of quality. In this dissertation, I included word reading accuracy, word recognition with foils, and spelling production tasks—each tapped a different level of lexical quality. Having these outcome measures also allowed examining which self-teaching condition (context vs. Isolation) would support the formation of lexical representations and to what degree of quality.

Self-teaching in Context and Isolation and its effect on Word Reading Accuracy

In line with previous studies, I found that contextual facilitation was present during training (e.g., Cunningham, 2006; Landi et al., 2006; Martin-Chang, 2017; Wang et al., 2011). That is, participants were more successful reading in context than in isolation. One robust novel contribution to the field that my dissertation adds—one which was replicated across each experiment—is that the context advantage observed during training was transferred to the post-test when participants read trained words in new situations without the benefit of context.

These findings provide evidence against the notion that reading in context does not lead to learning (Landi, 2013; Rastle & J. S. H. Taylor, 2018). Self-teaching studies that contrasted context against isolation reported that performance declined when participants read trained target words for the first time in isolation (e.g., Archer & Byrant, 2001; Landi et al., 2006). In

Study 1 presented here, Trial 1 in isolation can be considered a baseline measure of reading accuracy; it is the first time participants read the words, and they read them without external feedback or support from context. Comparing this baseline measure to the transfer task, we see that despite the drop in word reading accuracy from the last trial in context training to post-test transfer (of the words trained in context), the scores tended to remain significantly above baseline performance—this implies that learning took place during context training. From this point of view, if learning had not occurred during context training, reading scores during retention would be similar to the baseline accuracy scores in the isolation condition at Trial 1. Similar evidence can be seen with the child data. Unique to Experiment 5 (Study 2), children did a reading pretest. Reading accuracy for words trained in context improved between 9% and 22%, depending on the word type, from pretest to post-test. Thus, it seems clear that reading words in context leads to learning and develops lexical representations in memory that are at least *good enough* for later reading.

At the same time, it is important to note that retention was best when words were trained in isolation. This might sound contradictory to what was argued above, but it is not. Across the experiments presented in Study 1 and 2, participants maintained their reading performance from isolation training to the transfer task. This suggests that the words participants learned during isolation training were retained and accessible at post-test (also see Archer & Byrant, 2001; Landi et al., 2006). There is another critical difference between words trained in context and isolation, performance tends to be lower when training in isolation compared to context. That is, participants read fewer words correctly in isolation compared to context. Therefore, even when considering the decline in reading accuracy from context training to transfer, participants could read more words after context training compared to training in isolation. On average, participants read almost 20% more words in context compared to isolation. The combination of these factors resulted in a context advantage during retention.

Self-teaching in Context and Isolation and its Effect on Spelling Production

Had I not included spelling as an outcome measure, I might have erroneously concluded that reading in context is ideal for word learning. However, when reviewing the spelling recognition and production data, it becomes clear that isolation training was also advantageous. For the spelling recognition task, adult participants more readily identified correctly spelled target words that had been trained in isolation compared to context. Although the effect did not perfectly replicate within each experiment, my data does points to idea that isolation advantage translated to better spelling production several days after training. For example, with children spelled more words that had been trained in isolation than in context. This reflects the supposition that training words in isolation is more likely to leads to the formation of precise orthographic representations that can support producing accurate spellings (Landi et al., 2006; Martin-Chang et al., 2017). The reaction time data from the spelling recognition task provides more weight to this interpretation; adult participants were quicker in recognizing correctly spelled words that had been trained in isolation relative to context. It is also in line with lexical decision tasks data demonstrating that participants who are better spellers more rapidly access orthographic representations with less interference from similarly spelled neighbours (Andrews & Bond, 2008; Andrews & Hersch, 2010).

The Effect of Word Regularity and Word Length on Orthographic Learning

When assessing contextual facilitation in relation to word regularity, the data collected with adults and children diverge to some extent. When adults and children trained with regularized and regular target words, respectively, an effect of context was detected for word reading accuracy during and after training. However, when children trained both regular and irregular words, there was no statistical difference between context and isolation for regular words. Although there are a number of strengths with the methodologies used, differences in materials may partly explain the discrepancy in the pattern of results.

For example, adults trained with an artificial orthography whereas, children trained with the standard English alphabet. Recall that I implemented an artificial orthography for two reasons. First, to be able to measure orthographic learning within an expert population, adults had to be exposed to new symbols that were part of an orthographic writing system. But the target words also had to be part of their oral vocabulary. In choosing to use real words with real definitions, two of the three lexical identifies, phonology and semantics, were very likely already stored in participants' memory. That is, participants most likely knew how to pronounce the target words and knew their meaning. In terms of the artificial orthography, the symbols/graphemes were partly novel to participants and thus, to success during training, they had to lean on their reading skills to map graphemes onto phonemes, and this was done through trial-and-error over six trials (no feedback was provided). Children arguably partook in the same process but with the standard English alphabet. We asked children to train target words that were likely within their developmental zone to be able to strike a balance between words they could decode and learn to spell.

Second, the artificial orthography regularized the target words. This means adult participants could consistently rely on presented graphemes to sound the same way in different words that used the same graphemes. In fact, a majority of participants treated training as a solvable puzzle; many were visibly motivated and implemented acquired phonemic information through analogy. That is, participants seem to pick up on the fact that they could rely on the words they had correctly decoded as a starting point to decode other words with overlapping graphemes. For example, using skubΛ (scuba) to decode fud (food) and frΛnt (front). Regularizing the target words therefore permitted us to approach the debate of whether participants lean on context to learn regular words (Nation & Snowling, 1998; Wang et al., 2011, 2012, 2013; for review, see Colenbrander et al., 2020). It has been posited that regular words can be reliably deciphered through decoding thus, the secondary compensatory mechanism of

contextual facilitation is not activated. In Study 1, though, participants were exposed to a reliable orthographic system and still, participants leaned on the contextual information provided by the text to support word reading accuracy. In most cases, an effect of context was still present at Trial 6 and remained during retention.

This data suggests that it may not be word regularity itself that moderates the effect of context but, an element of challenge. Otherwise, I would have expected the effect of context to dissipate before Trial 6 or during retention, because reading would have been easy enough that the effect of context would not have time to contribute to word identification. Thus, I would speculate that a context effect would be present if adults were asked to read regular but appropriately challenging words written in the standard English alphabet. This interpretation is indirectly reinforced by the data presented in Study 2. Although we hypothesized that there would be a context effect for long regular words during training (Experiment 5), it became clear that most children did not find the long regular words challenging to read. Word reading accuracy at pretest was on average 67% and increased slightly at Trial 1 to approximately 70%. This translates to three errors out of ten target words to be learned. At the same time, a context effect was present for exception words and theory informs that this effect was present because exceptions words are more difficult to learn than words with regular spellings (Nation & Snowling, 1998; Wang et al., 2011). Thus, I would speculate that a context effect would be present if children were asked to read regular words that were challenging to read—pretest might have to be as low as 30% to measure learning and detect an effect of training condition (similar to spelling pretest scores). Indeed, we see a hint of this in Experiment 6, where all target words were regular, and a main effect of context was present during training.

The Effect of Stimulus Set Size on Orthographic Learning

Martin-Chang (2017) noted that the data that led researchers to conclude that context training is not ideal for forming lexical representations in memory were based on very small

stimulus set sizes. Wang et al. (2011) "...concluded that when decoding is difficult, as it is for irregular words, context helps children read words accurately. However, like Cunningham (2006), Wang et al. used a small learning set (four words per condition). Therefore, it is unclear whether their findings would generalize to reading longer passages" (Martin-Chang, 2017, p. 2). In my dissertation, I systematically tested whether stimulus set size could impact word reading accuracy and ultimately, orthographic learning. I originally posited that training more words would increase the reliance on context because the task would be more taxing. Decades of research demonstrates that when reading is challenging, top-down support from context is more likely to be recruited to determine a word's correct pronunciation (e.g., Simpson et al., 1989; Stanovich et al., 1985; Stanovich & West, 1983). Counterintuitively, our data showed that training more words did not increase reliance on context but, instead boosted reading performance for words that adults trained in isolation (during and after training).

Specifically, when comparing training performance between the large and small stimulus set size, participants in the large set size overtook those in the small set size for words trained in isolation but, not in context. When comparing performance between context and isolation, participants read more words accurately in context compared to isolation in the small set size but, not in the large set size. In other words, the context-isolation gap that is commonly reported (e.g., Martin-Chang, 2017) disappeared in the large set size because participants were a lot more successful in isolation than expected. But for participants who trained with the small set size, the context-isolation gap remained. The conceptual replication with children hints toward this pattern as well. When comparing word reading accuracy between context and isolation during retention, there was an effect of context for the small stimulus set size but not the large set size. For a more definitive conclusion, the mean difference for children between isolation performance in the small and large set size needed to reach statistical significance. Taken together, it nonetheless suggests that when a reader has fewer opportunities to decode, training

in context remains more advantageous than training in isolation. With a larger sample of children to increase power, we may also see that increasing set size boosts learning in isolation.

A reasonable conjecture is that seeing the same graphemes in many different words accelerates the building of lexical representations in memory (Cunningham & Stanovich, 1990; Nation, 2017; Stanovich & West, 1989; J. N. Taylor & Perfetti, 2016). Support for this interpretation comes from an eye tracking study that implemented a training design and manipulated exposure to the orthographic, phonological, and semantic information of rare words. They concluded that “more experience means more exposure to more words and thus to spellings, pronunciations, and meanings” (p. 1099, Taylor & Perfetti, 2016). Our results motivated also the idea that training with the large set size might be cultivating orthographic learning at the sublexical level. The grapheme post-test was used to determine whether stimulus set size affected participants’ learning of the phonetic symbols they read during training. Participants were asked to provide a word that was representative of the grapheme displayed on the screen. We found that participants who trained with the large stimulus set size provided more correct responses. We found the same effect of stimulus set size when we asked participants to read new, untrained words. Participants read more new, untrained words that contained graphemes from the large set size relative to the small set size. Future work should increase sample size and evaluate whether this effect would be present for words trained in context.

In the Introduction we motivated the need to also assess spelling. As noted above, participants were able to spell more words they trained in isolation compared to context—the opposite of what was observed for reading. However, the effect of stimulus set size was in the same direction for reading and spelling. That is, the isolation advantage for spelling was most evident for participants who practiced reading with large set size. As is, it seems appropriate to

conclude that pairing a reading activity that privileges decoding with a large amount of practice is well suited for building up lexical representations.

Practical Implications

The immediate implication of my dissertation is to encourage students to benefit from contextual facilitation for self-teaching. When considering the findings presented here along with the literature on reading in context, I am conformable drawing parallels between experimentally reading in context and reading books for pleasure. To be clear, I am not suggesting we put children in a resource room with a storybook and hope for the best. Instead, the benefits of reading in context should be leveraged within a structured pedagogy that is founded on the five pillars of literacy: vocabulary, phonemic awareness, phonics, fluency, and comprehension (Shanahan & North Central Regional Educational Lab, 2005). As a student becomes rooted in developing decoding skills, my work suggests that several classroom activities that involve reading in context can be effectively optimized to support immediate skill development and longer-term gains for reading developing.

Equally important is considering the limits of reading in context. Within my dissertation I present evidence showing that Isolation training is also constructive for word learning. When we asked adult and child participants to express their orthographic knowledge either through recognition or spelling production, we found that these tasks were easier for words they had trained in isolation. At this juncture I can only speculate, however it seems that reading words in isolation increases the likelihood that readers notice and consequently encode features of the orthography (Landi et al., 2006; Martin-Chang et al., 2017). In order to learn a word, students must notice, process, and encode the orthographic details of that word. It is these same steps that are necessary to lay down high quality lexical representations in memory (Perfetti, 2007). This productive struggle (Aiken et al., 2021; Kapur, 2016) resulting from a higher cognitive load while decoding in isolation may be driving attention, which in turn helped participants to build up

orthographic knowledge that was more precise than those formed during training context. As such, there are a number of classroom activities involving working with words out of context that may boost orthographic knowledge (Atkinson et al., 2014; Gentry & Ouellette, 2019; Moats, 2005; Williams et al., 2009; Zutell, 1998). This also fits with the practice of examining words out of context that is suggested in programs such as The Daily Five (Boushey & Moser, 2014) and the Orton-Gillingham approach (n.d.).

Furthermore, my data puts forth that educators of children and instructors of adult learners should consider that word type and set size might impact skill development. In planning lessons, teachers should determine the purpose of an activity before offering it to students. When reading accuracy is the goal, reading in context, (i.e., storybooks or novels) is time well spent. When spelling is the goal, word activities in isolation is also time well spent; it is especially beneficial for irregular words that do not adhere to consistent or frequently encountered letter-sound mappings. Although none of the participants read the same word in context and isolation, it seems highly plausible that having both experiences would be reciprocal in nature. My data also showed that increasing the number of words to be learned did not hamper learning. In fact, training with more words boosted word acquisition in such a way that it compensated for the trade-off between context and isolation.

Immediate Benefits of Reading in Context

The Matthew effect, postulated by Stanovich (1986), is borne out when children who achieve early successes in decoding build on their abilities through reading experiences at a fast rate. These same children are more likely to develop the habit of reading outside of school (Cunningham & Stanovich, 1997; Sparks et al., 2014). Meanwhile, those who struggle with cracking the alphabetic code fall behind their more-abled peers and read less often. This can lead to a downward spiral in reading development. To counteract the negative spiral, early reading intervention for struggling readers continues to be recommended (for meta-analysis,

see Wanzek et al., 2018) as well as developing positive reading habits (McGeown et al., 2015; Tremblay et al., 2020; van Bergen et al., 2021; Zare et al., 2023). Crucially, positive reading habits emanates from an interest in reading (Martin-Chang et al., 2021; Schiefele et al., 2012; van Bergen et al., 2021).

For these reasons correctly leveraging contextual facilitation is relevant to classroom instruction. Across my studies, both children and adults enjoyed more success when reading within the context of a story compared to reading words in a list devoid of semantic and syntactic cues. Beyond developing orthographic representations, the success of reading in context can be encouraging and motivating. Indeed, a number of studies have demonstrated that motivation is an integral part of developing positive reading habits and plays reciprocal role (Locher et al., 2019; McGeown et al., 2016; Schaffner & Schiefele, 2016; Walgermo et al., 2018).

At the same time, for students to effectively harness the bootstrapping effect of context, educators and instructors should carry out bite sized lessons focused on teaching vocabulary and phonics prior to independent reading. Although this might sound like common sense and is in fact supported by the science of reading (Ehri, 2020), it is critical to take note that even with decades of research, the educational field is still grappling with how to teach children to read in and out context. The Right to Read inquiry report, released February 2022 by the Ontario Human Right Commission, demanded an end to guessing words using clues from context and shift instruction to relying on phonological awareness and decoding skills. To be clear, I do not endorse the use of guessing from context for word identification or as a strategy for developing reading skills—a method repeatedly demonstrated to not work (e.g., Alexander, 1998; Nicholson, 1991; Tunmer & Chapman, 2002; Stanovich, 1984). Instead, my work suggests that readers can supplement slow and effortful decoding with contextual information to identify and pronounce words correctly. Thus, reading in context should be coupled with phonics instruction.

Long-term Benefits of Reading in Context

The immediate benefits of reading in context can lead to long-term gains. Decades of research demonstrates that reading works of fiction is an effective method to accrue language and reading abilities (Cunningham & Stanovich, 1998; Sparks et al., 2014; van Bergen et al., 2021), declarative knowledge (e.g., Stanovich & Cunningham, 1993), and social cognition skills (for a review, see Kozak & Recchia, 2019). What is more, a meta-analysis demonstrated the intricate link between print exposure and performance on vocabulary, word recognition, reading comprehension, and spelling measures (Mol & Bus, 2011). And the strength of the correlations increased within the years spanning primary and middle school and was strongest in middle school.

Some have found that the link between print exposure and reading-related skills remains in adulthood. Stanovich and West (1989) asked adults to perform a series of tasks tapping phonological awareness, orthographic knowledge, and decoding skills (collectively referred to as word processing skills). They found that the group who read for pleasure but had below average reading comprehension skills, demonstrated superior word processing abilities when compared to adult who rarely read for pleasure but had above average comprehension skills (also see, Willms & Murray, 2007). A similar pattern of findings was found with a sample of younger, and necessarily less experienced readers (Cunningham & Stanovich, 1991). Thus, it seems that modest reading comprehension skills may not necessarily lead to low levels of knowledge and abilities as long as students invest time in leisure reading (Cunningham & Stanovich, 1991; Stanovich & West, 1989).

More recently, a longitudinal study following children from age 5 to 15 assessed the association of print exposure with reading fluency and comprehension (van Bergen et al., 2021). With a sample size of 200 children, the authors found that early childhood reading skills exerted a positive effect on reading practice. Meaning that reading skills predicted time spend reading

(and not the other way around). They also found that by adolescence, time spent reading positively correlated with reading skills (cf. Erbeli et al., 2018; van Bergen et al., 2018). This pattern was corroborated with a sample of 2500 children from Grades 1 to 9 who were assessed on print exposure and reading fluency and comprehension (Torppa et al., 2019), and the pattern was limited to literature and did not extend to emails, social media, blogs etc. (see also Mar et al., 2006; McGeown et al., 2015; Pfost et al., 2012). These new studies reinforce the notion that reading in context is time well spent.

Striking a Balance with Word Study

What might be the long-term gains of reading in isolation? Theoretically speaking, the isolation condition was originally conceived as a control condition to compare against the context condition. It is now apparent that even though participants learned fewer words from isolation training, the words they did learn were better specified in memory. Nonetheless, learning is iterative and reciprocal. It is an ongoing process and the more you learn, the easier it is to keep learning. Although this was not directly tested here, there may be value in alternating between reading in context and examining words in isolation. Practically speaking, I would be hard pressed to recommend that children read long lists of words as a classroom activity. But, my work extends the notion that there may be value in asking children to study a spelling list that ties into themed classroom lesson.

Of note, studying words in isolation may not be as effective or interesting as something called word study. As far as I am aware, there is no study comparing activities that would fall under the umbrella of word study against reading words in isolation and context. However, there are a number of studies that touch upon word study activities with positive outcomes (Atkinson et al., 2014; Moats, 2005; Williams et al., 2009; Zutell, 1998). Again, I suspect a great value could be extracted by combining word work with reading in context.

Typically, word study activities ask students to examine words and their constituting parts based on phonological and orthographic features. For example, through activities such as words sorts, students initially categorize words or pictures according to their rhyming features (e.g., head, bed, said) and eventually sorting can be refined by separating words that rhyme but that have different orthographic rimes (e.g., fed, bed, wed, led versus head, dead, spread, read). Word study activities for more advanced learners can focus on morphological roots, instead. For example, students are given words like, *redo*, *hopeful*, *disagreement*, and *unselfishness*, and asked to circle the prefixes and suffixes, underline the vowels and vowel teams in the base/root of the word. This can be followed up by asking children to pronounce the parts and then the whole word. Teachers can also base word study activities around themes and vocabulary encountered in storybooks or class novels (Tompkins, 2014).

Limitations and Future Directions

The findings of this dissertation open new avenues for research. Across this dissertation, children and adults self-taught large sets of words that varied across multiple dimensions. Stimulus set size, word regularity, and word length were manipulated across six experiments. This permitted examining the moderating effect of stimulus set size on contextual facilitation and orthographic learning. Doing so demonstrated that more is sometimes better, but this could be furthered by directly comparing diversity of input (i.e., many different words) to repeated exposure to the same target words. One study that also implemented an artificial orthography suggests that participants are sensitive to frequency at the grapheme level (Taylor & Perfetti, 2016). Here, graphemes were presented in a large variety of target words but, it would be of value to determine if training the same graphemes in the same words would have a similar impact on learning.

We found that training in context better supported later reading and that training in isolation better supported spelling outcomes. When it came to the children, this pattern was

clearest when assessing performance from Time 1 to Time 2 and was limited to regular long words and exception words and when children trained with 30 words or less. With the current design it is not possible to comment how word work would compare to training in isolation. Some work suggests this would be a fruitful avenue to investigate (e.g., Joshi et al., 2008). It is likely that actively thinking and playing with word phonology and orthography would lead to superior spelling accuracy compared to reading in isolation.

Future research should also investigate individual differences. Recent work suggests that those with strong reading-related skills, like working memory and spelling, can more easily form lexical representations in memory (Andrews & Bond, 2008; Hersch & Andrews, 2012; Taylor & Perfetti, 2016; Veldre & Andrews, 2014). Also of value, would be to determine if similar effects would be detected with adults with less formal education or students identified as struggling readers. Seminal work found that poor readers and English learners may exhibited greater facilitation from context reading (Nation & Snowling, 1998; Nicholson, 1991; Stanovich & Stanovich, 1995; Wong & Underwood, 1996). Thus, it would be of value to extend outcome measures to include spelling recognition and accuracy.

Further insight might be gained from eye tracking data. To ensure that reading was fluent, we instructed participants to skip words when decoding laboured beyond 2 s. It is possible that time spent on decoding a word can mediate the formation of lexical representations. A self-paced design may yield a different pattern of results, possibly increasing the isolation advantage for spelling. Making use of eye tracking may be a solution to evaluate whether time spent visually inspected a word correlates with learning.

Conclusion

When situating the findings reported here in relation to the two plausible theories outlined in the introduction—self-teaching in context (Share, 1995) and the paradoxical outcome of isolation training (Landi et al., 2006)—I am confident in reporting that both viewpoints can be

harmonized. My data showed that when word reading accuracy is the outcome of interest, context training supports the formation of lexical representations in memory, especially when training with a small set of words. At the same time, isolation training is advantageous when spelling outcomes are privileged, and here more is better. Participants' orthographic knowledge was best and generalized when they trained with the large stimulus set.

So, what does this mean and why do we care? If you want to be a better reader, go ahead, pick up whatever book interests you, be it romance, fantasy, mystery. When you come across a challenging word, the other words in the context will help you to correctly identify and pronounce the challenging word. My data also showed that quantity matters; similar performance benefits were obtained when reading a large list of words and a story. This likely means that seeing the same combination of graphemes in many different words helps to secure the words in memory. So, if you really wanted to, you could read the dictionary. But, stories with dragons and wizards sounds a lot more fun. At the same time, if you want to improve spelling, then more focused work is required. When reading words in isolation, you likely spend more time analyzing the graphemes and the sounds they make, and this helps increase spelling accuracy. If you want to learn a tricky word, spend some time with it one-on-one. I must ask HER-MY-NEE how to spell her name the next time I see her.

References

- Aiken, H. H., Bratsch-Hines, M., Amendum, S., & Vernon-Feagans, L. (2021). Targeted Reading Instruction: Four Guiding Principles. *The Reading Teacher*, 74(5), 505-515.
<https://doi.org/10.1002/trtr.1975>
- Alexander, J. C. (1998). Reading skill and context facilitation: A classic study revisited. *The Journal of Educational Research*, 91(5), 314-318.
<https://www.tandfonline.com/doi/abs/10.1080/00220679809597559>
- Allington, R. L., & McGill-Franzen, A. M. (2021). Reading volume and reading achievement: A review of recent research. *Reading Research Quarterly*.
<https://doi.org/10.1002/rrq.404>
- Andrews, S., & Bond, R. (2008). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing*, 22(6), 687–711.
<https://doi.org/10.1007/s11145-008-9137-7>
- Archer, N., & Bryant, P. (2001). Investigating the role of context in learning to read: A direct test of Goodman's model. *British Journal of Psychology*, 92(4), 579–59.
<https://doi.org/10.1348/000712601162356>
- Atkinson, T. S., Zhang, G., Phillips, S. F., & Zeller, N. (2014). Using word study instruction with developmental college students. *Journal of Research in Reading*, 37(4), 433-448.
<https://doi.org/10.1111/1467-9817.12015>
- Barton, J. J., Hanif, H. M., Eklinder Björnström, L., & Hills, C. (2014). The word-length effect in reading: a review. *Cognition Neuropsychology*, 31(5–6), 378–412.
<https://doi.org/10.1080/02643294.2014.895314>
- Black, J. E., & Barnes, J. L. (2021). Fiction and morality: Investigating the associations between reading exposure, empathy, morality, and moral judgment. *Psychology of Popular Media*, 10(2), 149-164. <https://doi.org/10.1037/ppm0000281>

- Boushey, G., & Moser, J. (2014). *The Daily 5: Fostering literacy independence in the elementary grades* (Second). Stenhouse.
- Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson–Denny Reading Test*. Riverside.
- Castles, A., & Nation, K. (2006). How does orthographic learning happen? In S. E. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 181–209). Psychology Press.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19, 5–51.
doi:10.1177/1529100618772271
- Cedrus Corporation. (2014). SuperLab Pro (5.0). [Computer software]. San Pedro, CA.
www.superlab.com
- Chambrè, S. J., Ehri, L. C., & Ness, M. (2020). Phonological decoding enhances orthographic facilitation of vocabulary learning in first graders. *Reading and Writing*, 33(5), 1133–1162. <https://doi.org/10.1007/s11145-019-09997-w>
- Coltheart, V., & Leahy, J. (1992). Children’s and adults’ reading of nonwords: Effects of regularity and consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(4), 718. <https://doi.org/10.1037/0278-7393.18.4.718>
- Colenbrander, D., Miles, K. P., & Ricketts, J. (2019). To see or not to see: How does seeing spellings support vocabulary learning. *Language, Speech, and Hearing Services in Schools*, 50(4), 609–628. https://doi.org/10.1044/2019_LSHSS-VOIA-18-0135
- Conrad, N. J. (2008). From reading to spelling and spelling to reading: Transfer goes both ways. *Journal of Educational Psychology*, 100(4), 869–878. <https://doi.org/10.1037/a0012544>
- Conrad, N. J., & Deacon, S. H. (2016). Children’s orthographic knowledge and their word reading skill: Testing bidirectional relations. *Scientific Studies of Reading*, 20(4), 339–347. <http://dx.doi.org/10.1080/10888438.2016.1183128>

- Conrad, N. J., Kennedy, K., Saoud, W., Scallion, L., & Hanusiak, L. (2019). Establishing word representations through reading and spelling: Comparing degree of orthographic learning. *Journal of Research in Reading, 42*(1), 162–177.
- Cowan, N., & Alloway, T. (2009). Development of working memory in childhood. In M. L. C. & N. Cowan (Eds.), *Studies in developmental psychology. The development of memory in infancy and childhood* (pp. 303–342). Psychology Press.
- Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: Do children self-teach. *Journal of Experimental Child Psychology, 95*(1), 56–77.
<https://doi.org/10.1016/j.jecp.2006.03.008>
- Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 82*(4), 733. <https://psycnet.apa.org/record/1991-12234-001>
- Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of educational psychology, 83*(2), 264-274. <https://psycnet.apa.org/record/1991-26175-001>
- Cunningham, A. E. & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934-945.
- Cunningham, A. E., & Stanovich, K. E. (2001). What reading does for the mind. *Journal of Direct Instruction, 1*(2), 137-149.
- Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology, 82*(3), 185–199. [https://doi.org/10.1016/S0022-0965\(02\)00008-5](https://doi.org/10.1016/S0022-0965(02)00008-5)
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition*. Minneapolis, MN: Pearson Assessments.

- Echols, L. D., West, R. F., Stanovich, K. E., & Zehr, K. S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88*(2), 296–304. <https://doi.org/10.1037/0022-0663.88.2.296>
- Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to decoding. In P. E. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–144). <https://doi.org/10.4324/9781351236904>
- Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116–125. <https://doi.org/10.1111/j.1467-9817.1995.tb00077.x>
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading, 18*(1), 5–21. <https://doi.org/10.1080/10888438.2013.819356>
- Ehri, L. C. (2020). The science of learning to read words: A case for systematic phonics instruction. *Reading Research Quarterly, 55*(S1). <https://doi.org/10.1002/rrq.334>
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: evidence from the national reading panel's meta-analysis. *Review of Educational Research, 71*(3), 393–447. <https://doi.org/10.3102/00346543071003393>
- Erbeli, F., Bergen, E., & Hart, S. A. (2020). Unraveling the relation between reading comprehension and print exposure. *Child Development, 91*(5), 1548–1562. <https://doi.org/10.1111/cdev.13339>

- Evans, M. A., Williamson, K., & Pursoo, T. (2008). Preschoolers' attention to print during shared book reading. *Scientific Studies of Reading*, 12(1), 106-129.
<https://doi.org/10.1080/10888430701773884>
- Evans, M. A., Saint-Aubin, J., & Landry, N. (2009). Letter names and alphabet book reading by senior kindergarteners: An eye movement study. *Child development*, 80(6), 1824-1841.
<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8624.2009.01370.x>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. [Computer software]. *Behavior Research Methods*, 39, 175–191.
- Foorman, B. R., Francis, D. J., Davidson, K. C., Harm, M. W., & Griffin, J. (2004). Variability in text features in six Grade 1 basal reading programs. *Scientific Studies of Reading*, 8(2), 167–197. https://doi.org/10.1207/s1532799xssr0802_4
- Gagl, B., Hawelka, S., & Wimmer, H. (2015). On sources of the word length effect in young readers. *Scientific Studies of Reading*, 19(4), 289–306.
<https://doi.org/10.1080/10888438.2015.1026969>
- Gentry, J. R., & Ouellette, G. P. (2019). *Brain Words How the Science of Reading Informs Teaching*. Stenhouse Publishers.
- Goodman, K. S. (1967) Reading: A psycholinguistic guessing game. *Literacy Research and Instruction*, 6(4), 126–135, <https://doi.org/10.1080/19388076709556976>
- Goodwin, A. P., & Jiménez, R. T. (2020). The science of reading: Supports, critiques, and questions. *Reading Research Quarterly*. <https://doi.org/10.1002/rrq.360>
- Gough, P. B. (1983). Context, form, and interaction. In: K. Rayner (ed.), *Eye movements in reading: Perceptual and language processes* (pp. 203–211). Academic Press.

- Grolig, L., Tiffin-Richards, S. P., & Schroeder, S. (2020). Print exposure across the reading life span. *Reading and Writing*, 33(6), 1423–1441. <https://doi.org/10.1007/s11145-019-10014-3>
- Guitard, D., Gabel, A. J., Saint-Aubin, J., Surprenant, A. M., & Neath, I. (2018). Word length, set size, and lexical factors: Re-examining what causes the word length effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(11), 1824-1844. <https://doi.org/http://dx.doi.org/10.1037/xlm0000551>
- Guthrie, J. T., Wigfield, A., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In J. T. Guthrie, A. Wigfield, & W. You (Eds.), *Handbook of Research on Student Engagement* (pp. 601–634). Springer.
- Hersch, J., & Andrews, S. (2012). Lexical quality and reading skill: Bottom-up and top-down contributions to sentence processing. *Scientific Studies of Reading*, 16(3), 240–262. <https://doi.org/10.1080/10888438.2011.564244>
- Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, 49(16), 2078–2086. <https://doi.org/10.1016/j.visres.2009.05.015>
- Juel, C. (2006). The impact of early school experiences on initial reading. In Dickinson, David K. & Neuman, Susan B. (Eds.), *Handbook of early literacy research* (pp. 410–426). Guilford.
- Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Dev Psychol*, 44(3), 855-866. <https://doi.org/10.1037/0012-1649.44.3.855>
- Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, 51(2), 289–299. <https://doi.org/10.1080/00461520.2016.1155457>

- Kim, J. S., Hemphill, L., Troyer, M., Thomson, J. M., Jones, S. M., LaRusso, M. D., & Donovan, S. (2017). Engaging struggling adolescent readers to improve reading skills. *Reading Research Quarterly*, 52(3), 357–382. <https://doi.org/10.1002/rrq.171>
- Kozak, S., & Recchia, H. (2019). Reading and the development of social understanding: Implications for the literacy classroom. *The Reading Teacher*, 72(5), 569–577. doi.org/10.1002/trtr.1760
- Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, 65(1), 42–73. <https://doi.org/10.1016/j.jml.2011.03.002>
- Landi, N. (2013). Learning to read words: Understanding the relationship between reading ability, lexical quality, and reading context. In J-F. Rouet, S. R. Goldman, & M. A. Britt (Eds.), *Reading: From words to multiple texts* (pp. 31–47). Routledge.
- Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology*, 94(2), 114–133. <https://doi.org/10.1016/j.jecp.2005.12.004>
- Locher, F. M., Becker, S., & Pfof, M. (2019). The relation between students' intrinsic reading motivation and book reading in recreational and school contexts. *AERA Open*, 5(2), 1–14. <https://doi.org/10.1177/2332858419852041>
- Locher, F., & Pfof, M. (2019). The relation between time spent reading and reading comprehension throughout the life course. *Journal of Research in Reading*, 43(1), 57–77. <https://doi.org/10.1111/1467-9817.12289>
- Luke, S. G., Henderson, J. M., & Ferreira, F. (2015). Children's eye-movements during reading reflect the quality of lexical representations: An individual differences approach. *Journal*

of Experimental Psychology: Learning, Memory, and Cognition, 41(6), 1675–1683.

<https://doi.org/10.1037/xlm0000133>

MacKinnon, S. (2015). *When the going gets tough, context gets going: Exception word reading via self-teaching* (ID Code: 980776) [Master's thesis, Concordia University]. Spectrum Research Repository.

Mansell, J., Evans, M., & Hamilton-Hulak, L. (2005). Developmental changes in parents' use of miscue feedback during shared book reading. *Reading Research Quarterly*, 40(3), 294-317. <https://doi.org/10.1598/rrq.40.3.1>

Mar, R. A., & Rain, M. (2015). Narrative fiction and expository nonfiction differentially predict verbal ability. *Scientific Studies of Reading*, 19, 419–433
<https://doi.org/10.1080/10888438.2015.1069296>

Martens, V. E. G., & de Jong, P. F. (2008). Effects of repeated reading on the length effect in word and pseudoword reading. *Journal of Research in Reading*, 31(1), 40–54.
<https://doi.org/10.1111/j.1467-9817.2007.00360.x>

Martens, V. E., & de Jong, P. F. (2006). The effect of word length on lexical decision in dyslexic and normal reading children. *Brain and Language*, 98(2), 140–149.
<https://doi.org/10.1016/j.bandl.2006.04.003>

Martin-Chang, S. (2017). Learning to read with and without feedback, in and out of context. *Journal of Educational Psychology*, 109(2), 233–244.
<https://doi.org/10.1037/edu0000131>

Martin-Chang, S., & Gould, O. N. (2012). Reading to children and listening to children read: Mother–child interactions as a function of principal reader. *Early Education & Development*, 23(6), 855-876. <https://doi.org/10.1080/10409289.2011.578911>

- Martin-Chang, S., & Levesque, K. (2013). Taken out of context: Differential processing in contextual and isolated word reading. *Journal of Research in Reading, 36*(3), 330–349. <https://doi.org/10.1111/j.1467-9817.2011.01506.x>
- Martin-Chang, S., & Levesque, K. (2015). Reading words in and out of connected text: The impact of context on semantic and orthographic processing. *Scientific Studies of Reading, 19*(5), 392–408. <https://doi.org/10.1080/10888438.2015.1059839>
- Martin-Chang, S. L., & Levy, B. A. (2005). Fluency transfer: Differential gains in reading speed and accuracy following isolated word and context training. *Reading and Writing, 18*(4), 343–376. <https://doi.org/10.1007/s11145-005-0668-x>
- Martin-Chang, S. L., & Levy, B. A. (2006). Word reading fluency: A transfer appropriate processing account of fluency transfer. *Reading and Writing, 19*(5), 517–542. <https://doi.org/10.1007/s11145-006-9007-0>
- Martin-Chang, S. L., Levy, B. A., & O’Neil, S. (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology, 96*(1), 37–56. <https://doi.org/10.1016/j.jecp.2006.08.004>
- Martin-Chang, S., Ouellette, G., & Bond, L. (2017). Differential effects of context and feedback on orthographic learning: How good is good enough. *Scientific Studies of Reading, 21*(1), 17–30. <https://doi.org/10.1080/10888438.2016.1263993>
- Martin-Chang, S., Ouellette, G., & Madden, M. (2014). Does poor spelling equate to slow reading? The relationship between reading, spelling, and orthographic quality. *Reading and Writing, 27*(8), 1485–1505. <https://doi.org/10.1007/s11145-014-9502-7>
- Martin-Chang, S., Kozak, S., Levesque, K. C., Calarco, N., & Mar, R. A. (2021). What’s your pleasure? Exploring the predictors of leisure reading for fiction and nonfiction. *Reading and Writing, 34*(6), 1387–1414. <https://doi.org/10.1007/s11145-020-10112-7>

- McGeown, S. P., Duncan, L. G., Griffiths, Y. M., & Stothard, S. E. (2015). Exploring the relationship between adolescent's reading skills, reading motivation and reading habits. *Reading and Writing*, 28(4), 545-569. <https://doi.org/10.1007/s11145-014-9537-9>
- Merga, M. K. (2017). What would make children read for pleasure more frequently. *English in Education*, 51(2), 207–223. <https://doi.org/10.1111/eie.12143>
- Miles, K. P., & Ehri, L. C. (2017). Learning to read words on flashcards: Effects of sentence contexts and word class in native and nonnative English-speaking kindergartners. *Early Childhood Research Quarterly*, 41, 103–113. <https://doi.org/10.1016/j.ecresq.2017.06.001>
- Moats, L. C. (2003). *Speech to print workbook*. The Bookwatch.
- Moats, Louisa C. (2005). How spelling supports reading: And why it is more regular and predictable than you may think. *American Educator*, Winter 2005/06, 12-43.
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267–296. doi.org/10.1037/a0021890
- Nation, K. (2017). Nurturing a lexical legacy: Reading experience is critical for the development of word reading skill. *NPJ Science of Learning*, 2(3), 1–4. <https://doi.org/10.1038/s41539-017-0004-7>
- Nation, K. (2019). Children's reading difficulties, language, and reflections on the simple view of reading. *Australian Journal of Learning Difficulties*, 24(1), 47-73. <https://doi.org/10.1080/19404158.2019.1609272>
- Nation, K., & Snowling, M. J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child development*, 69(4), 996–1011. <https://doi.org/10.1111/j.1467-8624.1998.tb06157.x>

- Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, 96(1), 71–84. <https://doi.org/10.1016/j.jecp.2006.06.004>
- Nicholson, T. (1991). Do children read words better in context or in lists? A classic study revisited. *Journal of Educational Psychology*, 83(4), 444–450. <https://doi.org/10.1037/0022-0663.83.4.444>
- Ocal, T., & Ehri, L. C. (2017). Spelling pronunciations help college students remember how to spell difficult words. *Reading and Writing*, 30(5), 947–967. <https://doi.org/10.1007/s11145-016-9707-z>
- Ontario Human Rights Commission. (2022, February). *Right to Read: Public inquiry into human rights issues affecting students with reading disabilities*. <https://www.ohrc.on.ca/en/right-to-read-inquiry-report>
- Orton-Gillingham Approach (n.d.) <https://www.orton-gillingham.com/approach/>
- Ouellette, G. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology*, 107(1), 50–58. <https://doi.org/10.1016/j.jecp.2010.04.009>
- Ouellette, G., & Fraser, J. R. (2009). What exactly is a *yait* anyway: The role of semantics in orthographic learning. *Journal of Experimental Child Psychology*, 104(2), 239–251. <https://doi.org/10.1016/j.jecp.2009.05.001>
- Ouellette, G., & Sénéchal, M. (2008). Pathways to literacy: A study of invented spelling and its role in learning to read. *Child Development*, 79(4), 899–913. <https://doi.org/10.1111/j.1467-8624.2008.01166.x>
- Ouellette, G., Martin-Chang, S., & Rossi, M. (2017). Learning from our mistakes: Improvements in spelling lead to gains in reading speed. *Scientific Studies of Reading*, 21(4), 350–357. <https://doi.org/10.1080/10888438.2017.1306064>

- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Routledge.
<https://doi.org/10.4324/9781351236904>
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven. C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy: Vol. 11 Studies in Written Languages and Literacy* (pp. 189–213). Amsterdam/Philadelphia: John Benjamins.
- Perfetti, C. A., & Roth, S. (1981). Some of the interactive processes in reading and their role in reading skill. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading*. Routledge.
- Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition*, 7(4), 273–282.
- Pfost, M., Dörfler, T., & Artelt, C. (2012). Reading competence development of poor readers in a German elementary school sample: An empirical examination of the Matthew effect model. *Journal of Research in Reading*, 35(4), 411-426. <https://doi.org/10.1111/j.1467-9817.2010.01478.x>
- Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, 52(1), 1–56.
<https://doi.org/10.1016/j.cogpsych.2005.06.001>
- Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human perception and performance*, 25(2), 482–503.
- Rastle, K., & Taylor, J. S. H. (2018). Print-sound regularities are more important than print-meaning regularities in the initial stages of learning to read: Response to Bowers &

- Bowers (2018). *Quarterly Journal of Experimental Psychology (2006)*, 71(7), 1501–1505.
<https://doi.org/10.1177/1747021818775053>
- Rau, A. K., Moeller, K., & Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: An eye-tracking study. *Scientific Studies of Reading*, 18(3), 224–233. <https://doi.org/10.1080/10888438.2013.857673>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506.
<https://doi.org/10.1080/17470210902816461>
- Rayner, K., Pollatsek, A., & Schotter, E. R. (2012). Word identification and eye movements. In A. F. H. I. Weiner, & R.W. Proctor (Ed.), *Handbook of Psychology* (2nd ed., pp. 548–577). Wiley Online Library. <https://doi.org/10.1002/9781118133880.hop204020>
- Rayner, K., Pollatsek, A., & Schotter, E. R. (2013). Reading: Word identification and eye movements. In A. F. Healy, R. W. Proctor, & I. B. Weiner (Eds.), *Handbook of psychology: Experimental psychology* (pp. 548–577). Hoboken, NJ, US: John Wiley & Sons Inc.
- Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, 21(3), 448–465.
<https://doi.org/10.1037/0882-7974.21.3.448>
- Renaissance Learning (2018, January 23). *Daily reading practice and the magic of 15 minutes*.
<https://www.renaissance.com/2018/01/23/blog-magic-15-minutes-reading-practice-reading-growth/>

- Ricketts, J., Bishop, D. V. M., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, 15(1), 47–70. <https://doi.org/10.1080/10888438.2011.536129>
- Rodrigues, M. L., Jing, S. E., & Martin-Chang, S. (revise, resubmit). When size matters: Effect of context and set size on orthographic learning. *Journal of Cognitive Psychology*.
- Rossi, M., Martin-Chang, S., & Ouellette, G. (2019). Exploring the space between good and poor spelling: Orthographic quality and reading speed. *Scientific Studies of Reading*, 23(2), 192–201. <https://doi.org/10.1080/10888438.2018.1508213>
- Schaffner, E., & Schiefele, U. (2016). The contributions of intrinsic and extrinsic reading motivation to the development of reading competence over summer vacation. *Reading Psychology*, 37(6), 917-941. <https://doi.org/10.1080/02702711.2015.1133465>
- Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, 47(4), 427-463. <https://doi.org/10.1002/RRQ.030>
- Schmalz, X., Marinus, E., & Castles, A. (2013). Phonological decoding or direct access? Regularity effects in lexical decisions of Grade 3 and 4 children. *The Quarterly Journal of Experimental Psychology*, 66(2), 338–346. <https://doi.org/10.1080/17470218.2012.711843>
- Schwartz, M., Kahn-Horwitz, J., & Share, D. L. (2014). Orthographic learning and self-teaching in a bilingual and biliterate context. *Journal Experimental Child Psychology*, 117, 45–58. <https://doi.org/10.1016/j.jecp.2013.08.008>
- 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/10.1037/0033-295X.96.4.523>

- Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 23(3), 383–404.
[https://doi.org/10.1016/S0022-5371\(84\)90270-6](https://doi.org/10.1016/S0022-5371(84)90270-6)
- Shahar-Yames, D., & Share, D. L. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading*, 31(1), 22–39.
<https://doi.org/10.1111/j.1467-9817.2007.00359.x>
- Shanahan, T., & North Central Regional Educational Lab., N. I. (2005). The National Reading Panel Report: Practical advice for teachers. *Learning Point Associates / North Central Regional Educational Laboratory*.
- Share, D. L. (1995). Phonological decoding and self-teaching: *sine qua non* of reading acquisition. *Cognition*, 55(2), 151–218. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)
- Share, D. L. (1999). Phonological decoding 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. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, 87(4), 267–298.
<https://doi.org/10.1016/j.jecp.2004.01.001>
- Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, 134(4), 584–615.
<https://doi.org/10.1037/0033-2909.134.4.584>
- 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.

- Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education: Contributions from Educational Psychology*, 1, 1–57.
- Simpson, G. B., Peterson, R. R., Casteel, M. A., & Burgess, C. (1989). Lexical and sentence context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(1), 88–97.
- Sparks, R. L., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of ‘10 years later.’ *Reading and Writing*, 27(1), 189–211. doi.org/10.1007/s11145-013-9439-2
- Stanovich, K. E. (1984). The interactive-compensatory model of reading: A confluence of developmental, experimental, and educational psychology. *Remedial and Special Education*, 5(3), 11–19.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.
- Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, 20(1), 51–68. doi.org/10.3758/BF03208254
- Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*. 85(2), 211-229.
- Stanovich, K. E., & Stanovich, P. J. (1995). How research might inform the debate about early reading acquisition. *Journal of Research in Reading*, 18(2), 87–105.
<https://doi.org/10.1111/j.1467-9817.1995.tb00075.x>
- Stanovich, K. E., & West, R. F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition*, 7(2), 77–85.

- Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, 7(3), 658–672.
- Stanovich, K. E., & West, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, 112(1), 1–36.
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24(4), 402–433. <https://doi.org/10.2307/747605>
- Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Relation between early reading acquisition and word decoding with and without context: A longitudinal study of first-grade children. *Journal of Educational Psychology*, 76(4), 668.
<https://psycnet.apa.org/record/1985-10766-001>
- Stanovich, K. E., West, R. F., & Feeman, D. J. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive-compensatory model. *Journal of Experimental Child Psychology*, 32(2), 185–199.
[https://doi.org/10.1016/0022-0965\(81\)90076-X](https://doi.org/10.1016/0022-0965(81)90076-X)
- Stanovich, K. E., Nathan, R. G., West, R. F., & Vala-Rossi, M. (1985). Children's word recognition in context: Spreading activation, expectancy, and modularity. *Child Development*, 56(6), 1418–1428. <https://doi.org/10.2307/1130461>
- Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language & Linguistics Compass*, 9(8), 311–327.
<https://doi.org/10.1111/lnc3.12151>
- Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal Experimental Psychology: Learning Memory Cognition*, 28(1), 207–214. <https://doi.org/10.1037/0278-7393.28.1.207>

- Taylor, J. N., & Perfetti, C. A. (2016). Eye movements reveal readers' lexical quality and reading experience. *Reading and Writing*, 29(6), 1069–1103. <https://doi.org/10.1007/s11145-015-9616-6>
- Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 60–76. <https://doi.org/10.1037/a0020126>
- Tiffin-Richards, S. P., & Schroeder, S. (2020). Context facilitation in text reading: A study of children's eye movements. *Journal Experimental Psychology: Learning Memory Cognition*, 46(9), 1701–1713. <https://doi.org/10.1037/xlm0000834>
- Tremblay, B., Rodrigues, M. L., & Martin-Chang, S. (2020). From storybooks to novels: A retrospective approach linking print exposure in childhood to adolescence. *Frontiers in Psychology*, 11:571033. <https://doi.org/10.3389/fpsyg.2020.571033>
- Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M. K., Tolvanen, A., & Poikkeus, A. M. (2019). Leisure reading (but not any kind) and reading comprehension support each other-a longitudinal study across grades 1 and 9. *Child Development*. <https://doi.org/10.1111/cdev.13241>
- Tunmer, W. E., & Chapman, J. W. (1998). Language prediction skill, phonological decoding ability, and beginning reading. In R. C. Hulme & M. Joshi (Eds.), *Reading and Spelling: Development and Disorders*. Routledge
- Tunmer, W. E., & Chapman, J. W. (2002). The relation of beginning readers' reported word identification strategies to reading achievement, reading-related skills, and academic self-perceptions. *Reading and Writing*, 15, 341–358. <https://doi.org/10.1023/A:1015219229515>

- van Bergen, E., Vasalampi, K., & Torppa, M. (2021). How are practice and performance related? Development of reading from age 5 to 15. *Reading Research Quarterly*, 56(3), 415–434. <https://doi.org/10.1002/rrq.309>
- van Bergen, E., Snowling, M. J., de Zeeuw, E. L., van Beijsterveldt, C. E. M., Dolan, C. V., & Boomsma, D. I. (2018). Why do children read more? The influence of reading ability on voluntary reading practices. *Journal of Child Psychology and Psychiatry*, 59(11), 1205–1214. <https://doi.org/10.1111/jcpp.12910>
- Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *The Quarterly Journal of Experimental Psychology*, 67(4), 703–727. <https://doi.org/10.1080/17470218.2013.826258>
- Walgermo, B. R., Foldnes, N., Uppstad, P. H., & Solheim, O. J. (2018). Developmental dynamics of early reading skill, literacy interest and readers' self-concept within the first year of formal schooling. *Reading and Writing*, 31(6), 1379-1399. <https://doi.org/10.1007/s11145-018-9843-8>
- Wang, H-C., Castles, A., & Nickels, L. (2012). Word regularity affects orthographic learning. *Quarterly Journal of Experimental Psychology*, 65(5), 856–864. <https://doi.org/10.1080/17470218.2012.672996>
- Wang, H-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, 109(1), 39–57. <https://doi.org/10.1016/j.jecp.2010.11.005>
- Wang, H-C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, 17(5), 369–384. <https://doi.org/10.1080/10888438.2012.749879>

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>

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition*, 13(6), 557–572.

<https://doi.org/10.3758/BF03198326>

West, R. F., & Stanovich, K. E. (1978). Automatic contextual facilitation in readers of three ages.

Child Development, 49(3), 717–727. <https://doi.org/10.2307/1128240>

West, R. F., Stanovich, K. E., Feeman, D. J., & Cunningham, A. E. (1983). The effect of sentence context on word recognition in second- and sixth-grade children. *Reading Research Quarterly*, 19(1), 6–15.

Wilkinson, G. S. (1993). *Wide Range Achievement Test–Third Edition*. Lutz, FL: Psychological Assessment Resources.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test–Fourth Edition*. Lutz, FL: Psychological Assessment Resources.

Williams, C., Phillips-Birdsong, C., Hufnagel, K., Hungler, D. and Lundstrom, R.P. (2009). Word Study Instruction in the K–2 Classroom. *The Reading Teacher*, 62(7), 570-578. <https://doi.org/10.1598/RT.62.7.3>

Willms, D., & Murray, S. (2007). *International adult literacy survey: Gaining and losing literacy skills over the lifecourse*. Statistics Canada. <https://www150.statcan.gc.ca/n1/en/pub/89-552-m/89-552-m2007016-eng.pdf?st=IXDus7RF>.

Wong, M. Y., & Underwood, G. (1996). Do bilingual children read words better in lists or in context? *Journal of Research in Reading*, 19(1), 61–76. <https://doi.org/10.1111/j.1467-9817.1996.tb00087.x>

- Zare, M., Kozak, S., Rodrigues, M. L., & Martin-Chang, S. (2023). The roots of reading for pleasure: Recollections of reading and current habits. *Literacy*.
<https://doi.org/10.1111/lit.12315>
- Zawoyski, A. M., Ardoin, S. P., & Binder, K. S. (2015). Using Eye Tracking to Observe Differential Effects of Repeated Readings for Second-Grade Students as a Function of Achievement Level. *Reading Research Quarterly*, 50(2), 171–184.
<https://doi.org/10.1002/rrq.91>
- Zhao, J., Li, T., Elliott, M. A., & Rueckl, J. G. (2018) Statistical and cooperative learning in reading: An artificial orthography learning study. *Scientific Studies of Reading*, 22(3), 191–208. <https://doi.org/10.1080/10888438.2017.1414219>
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3–29. <https://doi.org/10.1037/0033-2909.131.1.3>
- Zutell, J. (1998). Word sorting: A developmental spelling approach to word study for delayed readers. *Reading and Writing Quarterly*, 14(2), 219-238.
<https://doi.org/10.1080/1057356980140205>

Appendix A

Example of Consent Form: Adults (Study 1)

Minor adjustments were made reflecting exact details of Experiments 1, 2, 3, and 4



INFORMATION AND CONSENT FORM

Study Title: When in Doubt Read in Context

Researcher: Monyka L. Rodrigues, MA

Researcher's Contact Information: literacy.lab.concordia@gmail.com

Faculty Supervisor: Dr. Sandra Martin-Chang

Faculty Supervisor's Contact Information: s.martin-chang@concordia.ca

Source of funding for the study: N/A

You are being invited to participate in the research study mentioned above. This form provides information about what participating would mean. Please read it carefully before deciding if you want to participate or not. If there is anything you do not understand, or if you want more information, please ask the researcher.

A. PURPOSE

The purpose of the research is to study the effect of reading words in passages of text compared to reading words in lists on reading speed and accuracy.

B. PROCEDURES

It is a three-part study. Each session is scheduled 3-5 days apart. During the study, you will do some reading and spelling; some words will be in English while others will be presented in phonemic notation. The reading portion includes reading several words in short passages and lists while some words will be written in phonemic notation. You will then be asked whether you recognize some of these words. For the spelling portion, you will be asked to spell a list of words (in both English and phonemic notation), increasing in difficulty. Following each spelling exercise, you will be asked to send a picture of the words you spelled to the researcher. To remember what was answered, the first and second sessions will be audio recorded while the third session, the screen will be video recorded. Your video feed (i.e., your face) will not be recorded.

Session 1: You will read and spell some words in increasing difficulty. After a short break, you will read one story and one list out loud where some words will be phonemic notation. To measure learning, the story and list will be read three times each. In total, participating will take between 30-45 minutes.

Session 2: You will continue to read the story and list out loud. To measure learning, the story and list will be read three times each. After a short break, you will read a few short sentences and answer some vocabulary questions followed by reading a few short stories and answer some comprehension questions. In total, participating will take approximately 40-60 minutes.

Session 3: In phonemic notation, you will read a list of words, identify correct spellings, make old-new decisions about words, and spell some words in increasing difficulty. In total, participating will take approximately 40-60 minutes (including short breaks).

C. RISKS AND BENEFITS

There are no known risks to participating in the experiments described here. Potential benefits may include learning to read and spell new words and learn the meaning of new words.

D. CONFIDENTIALITY

We will gather the following information as part of this research:

Your date of birth;
Education level;
Primary languages;
Reading and spelling abilities.

We will not allow anyone to access the information, except people directly involved in conducting the research. We will only use the information for the purposes of the research described in this form.

The information gathered will be coded. That means that the information will be identified by a code. The researcher will have a list that links the code to your name.

We will protect the information by keeping the data on a secure server in a locked room.

We intend to publish the results of the research. However, it will not be possible to identify you in the published results.

We will destroy the code linking your name (if you choose to provide one) to the data five years after the completion of this study. The coded data from this study will be permanently archived.

F. CONDITIONS OF PARTICIPATION

You do not have to participate in this research. It is purely your decision. If you do participate, you can stop at any time. You can also ask that the information you provided not be used, and your choice will be respected. If you decide that you don't want us to use your information, you must tell the researcher within one month of participating.

As a compensatory indemnity for participating in this research, you will receive a \$45 transfer. If you withdraw before the end of the research, you will receive a \$15 transfer. To make sure that research money is being spent properly, auditors from Concordia or outside will have access to a coded list of participants. It will not be possible to identify you from this list.

There are no negative consequences for not participating, stopping in the middle, or asking us not to use your information.

G. PARTICIPANT'S DECLARATION

I have read and understood this form. I have had the chance to ask questions and any questions have been answered. I agree to participate in this research under the conditions described.

NAME (optional)

SIGNATURE

DATE

If you have questions about the scientific or scholarly aspects of this research, please contact the researcher. Their contact information is on page 1. You may also contact their faculty supervisor.

If you have concerns about ethical issues in this research, please contact the Manager, Research Ethics, Concordia University, 514.848.2424 ex. 7481 or oor.ethics@concordia.ca.

Appendix B

Example of Consent Form: Parents of children (Study 2)

Minor adjustments were made reflecting exact details of Experiments 3 and 4



INFORMATION AND CONSENT FORM

Study Title: Learning when reading
Researcher: Monyka L. Rodrigues, MA
Researcher's Contact Information: literacy.lab.concordia@gmail.com
Faculty Supervisor: Dr. Sandra Martin-Chang
Faculty Supervisor's Contact Information: s.martin-chang@concordia.ca
Source of funding for the study: NSERC Discovery Grant RGPIN-2018-03758

You are being invited to participate in the research study mentioned above. This form provides information about what participating would mean. Please read it carefully before deciding if you want to participate or not. If there is anything you do not understand, or if you want more information, please ask the researcher.

A. PURPOSE

The purpose of this research is to study how reading words in short stories and lists effects children's ability to store words in memory.

B. PROCEDURES

If your child participates, they will be asked to read and spelling grade appropriate words during different activities. It is a multi-part study taking place over several days. Some of the sessions will be audio recorded for scoring purposes only.

C. RISKS AND BENEFITS

There are no known risks to participating in the experiments described here. Potential benefits may include learning to read and spell new words and learn the meaning of new words.

D. CONFIDENTIALITY

We will gather the following information as part of this research:

Children date of birth;
Parent(s) / guardian Education level;
Primary languages spoken;
Reading and spelling abilities.

By agreeing to have your child participate in this study, you agree to let the researcher use the information gathered during the tasks. We will not allow anyone to access the information, except people directly involved in conducting the research. We will only use the information for the purposes of the research described in this form.

The information gathered will be coded. That means that the information will be identified by a code. The researcher will have a list that links the code to participants name. We will protect the information by ensuring the paper copies of all tasks are kept in a locked laboratory and by keeping the data on a secure server in a locked room.

We intend to publish the results of the research. However, it will not be possible to identify you or your child in the published results.

We will destroy the code linking your name to the data five years after the completion of this study. The coded data from this study will be permanently archived.

F. CONDITIONS OF PARTICIPATION

Your child does not have to participate in this research. It is purely your decision. If they participate, they can stop at any time and elect to have their data removed from the study. If you decide that you don't want us to use their information, you must tell the researcher within one month of participating.

As a compensatory indemnity for participating in this research, your child will receive a choice of stickers or a pencils at each session. If they withdraw before the end of the research, they will still receive the small gift. At the end, your child will receive a storybook and a junior scientist certificate. To make sure that research money is being spent properly, auditors from Concordia or outside will have access to a coded list of participants. It will not be possible to identify your child from this list.

There are no negative consequences for not participating, stopping in the middle, or asking us not to use your information.

G. PARTICIPANT'S DECLARATION

I have read and understood this form. I have had the chance to ask questions and any questions have been answered. I agree to participate in this research under the conditions described.

CHILD'S NAME (please print)

PARENT'S NAME

SIGNATURE

DATE

If you have questions about the scientific or scholarly aspects of this research, please contact the researcher. Their contact information is on page 1. You may also contact their faculty supervisor.

If you have concerns about ethical issues in this research, please contact the Manager, Research Ethics, Concordia University, 514.848.2424 ex. 7481 or oor.ethics@concordia.ca.

Appendix C

Script read to children

Your parents gave permission for you to participate in my study about reading and spelling activities. If you want to work with me, it means I can use your information to help your teachers find the best ways to teach you how to read and write. I will not share your answers with anyone, including your teachers, or your friends. There are no wrong answers, I just want you to do your best.

Appendix D

Demographics Questionnaire: Adults (Study 1)

Please list all of the languages you speak in order of dominance:

- 1.
- 2.
- 3.

Please indicate the highest level of education you have obtained (or the approximate equivalent to a degree obtained in another country):

Less than high school
High school (List grade _____)
Some CEGEP/College (List year _____)
CEGEP DEC/College degree
Some University (List year _____)
Undergraduate degree
Some Graduate School
Masters Degree
Ph.D./M.D./J.D.
Other _____

Date of birth:

Please indicate your gender:

Male Female Other

Please indicate your ethnicity(ies):

Asian
Black or African American
First Nations
Hispanic or Latino
White

Appendix E

Demographics Questionnaire: Parents of Children (Study 2)

Please list the languages your child speaks in order of dominance (strongest to weakest):

- 1.
- 2.
- 3.

Mother or Primary Guardian: Please indicate the highest level of education you have obtained (or the approximate equivalent to a degree obtained in another country):

High school (List grade _____)
CEGEP DEC/College degree
Undergraduate degree
Masters Degree
Ph.D./M.D./J.D.
Other _____

Day / Month / Year

Your child's date of birth:

Please indicate your child's gender:

Male Female Other, specify: _____

Please indicate your child's ethnicity(ies):

Arab
Black or African American
East Asian
Indigenous or First Nations
Latin American
South Asian (e.g., East Indian, Pakistani, Sri Lankan)
Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, Laotian)
West Asian (e.g., Iranian, Afghan)
White (e.g., European decent)

Appendix F

Experiment 1: Target Words

Word Set A (Bob)			Word Set B (Sarah)		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
accident	noun	42.35	able	adjective	285.48
afraid ^a	adjective	69.87	accustomed	adjective	4.86
ambulance ^a	noun	9.74	airplane	noun	13.82
amount ^a	noun	108.33	amazing ^a	adjective	50.05
analyzed	verb	18.22	assisted	verb	6.32
apple	noun	46.07	avid	adjective	5.51
audit	noun	9.60	bankruptcy ^a	noun	21.19
band ^a	noun	70.03	being	verb	700.39
blood	noun	151.26	best	adjective	340.25
brave	verb	1.04	bloom	noun	3.90
brother	noun	119.38	bridesmaid	noun	0.69
calculator	noun	2.24	brought	verb	174.82
case	noun	382.55	camouflage	noun	2.66
caused ^a	verb	69.68	city ^a	noun	475.66
chatty	adjective	0.84	coffin ^a	noun	7.43
cold	adjective	110.12	companions	noun	6.46
concert	noun	27.59	conquered	verb	2.98
consisted	verb	13.10	could	verb	1683.14
cost	noun	102.69	countries	noun	146.26
determine	verb	70.73	dictionary	noun	7.23
dinner	noun	89.89	disappear	verb	17.08
dizzy	adjective	5.04	discovered	verb	66.92
doctors	noun	57.21	divorce	noun	24.12
earthquake	noun	14.12	dream	noun	64.98
eaten ^a	verb	14.05	employment ^a	noun	36.46
enjoying	verb	16.94	escalator	noun	1.43
event	noun	101.38	exciting	adjective	28.73
expecting	verb	20.79	eyes	noun	338.21
expensive	adjective	55.73	flower	noun	19.71
fact	noun	370.02	following	verb	85.37
fear	noun	71.62	food ^a	noun	249.25
fell ^a	verb	91.26	forgot	verb	16.62
fine	adjective	130.13	front	noun	59.34
fishing	noun	42.12	funny	adjective	55.08
guitar	noun	18.98	girl	noun	177.88

Word Set A (Bob)			Word Set B (Sarah)		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
happened	verb	215.07	give ^a	verb	379.57
hospital	noun	125.76	gotten ^a	verb	64.82
illness ^a	noun	32.59	greatest	adjective	58.90
knees	noun	36.53	haunted	verb	5.35
knew ^a	verb	329.73	heights	noun	16.26
laboratory ^a	noun	30.09	huge	adjective	115.34
matters ^a	noun	42.11	industrialized	verb	0.05
meantime	noun	14.65	instance	noun	11.80
miracle ^a	noun	16.97	interesting ^a	adjective	107.25
missed	verb	45.26	jellyfish ^a	noun	1.43
money	noun	420.56	kitten	noun	2.91
mood	noun	32.07	library	noun	69.79
nausea	noun	4.44	little	adjective	390.83
next	adjective	492.24	look	verb	518.73
nightmare	noun	15.15	machines	noun	32.41
numerous	adjective	40.50	made ^a	verb	711.67
nurses ^a	noun	23.48	magician ^a	noun	3.49
original	adjective	93.58	meaningful	adjective	23.20
overcooked ^a	adjective	0.20	months	noun	273.48
pills ^a	noun	13.06	musty ^a	adjective	1.78
placed ^a	verb	69.52	named	verb	98.96
public ^a	adjective	326.79	need	noun	116.81
puppy ^a	noun	6.14	nonexistent	adjective	3.77
radiant	adjective	3.90	ocean ^a	noun	48.02
recommended	verb	26.17	packed	verb	22.69
recuperating ^a	verb	0.48	paid ^a	verb	99.75
returned	verb	88.37	parents	noun	261.34
rotten	adjective	5.55	past ^a	noun	101.90
sauce ^a	noun	37.73	prison	noun	76.44
severe	adjective	40.89	recovering ^a	verb	9.13
shore	noun	17.37	remote	adjective	28.89
sing	verb	33.85	satisfying ^a	adjective	8.78
spaghetti	noun	4.61	schoolwork	noun	1.05
spouse	noun	11.40	scuba diving	noun	0.42
stricken	verb	3.05	shipwreck	noun	1.07
sunshine	noun	9.19	spent	verb	149.52
surgery ^a	noun	50.42	stay	verb	169.12
survived ^a	verb	25.15	thin	adjective	54.75

Word Set A (Bob)			Word Set B (Sarah)		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
symptoms	noun	39.04	ticket ^a	noun	31.43
talk	verb	271.75	took	verb	413.75
tasteless	adjective	1.10	tour	noun	53.47
taxes	noun	66.81	traffic	noun	56.92
television	noun	113.84	translation ^a	noun	14.08
tests	noun	65.08	travelling ^a	verb	1.40
told	verb	464.26	university ^a	noun	359.86
tranquil	adjective	2.17	vacation	noun	26.13
vulgar ^a	adjective	3.57	victims ^a	noun	59.67
wanted	verb	321.94	villages	noun	16.81
weak ^a	adjective	37.55	wedding	noun	41.47
would	verb	2458.60	white ^a	adjective	323.17
years ^a	noun	1215.94	world	noun	766.87

Note. Target words were presented in phonetic symbols (Moats, 2003) in and out of context.

^atarget words part of the small (48) set size condition.

Appendix G

Experiment 1: Story A (Bob) Large Stimulus set size

Ever since Bob had the **æksidənt** on his sailboat, he was **afred** of going to the **haspɪtəl**. He **wʊd** rather do his **tæksɛz**—twice!—then go through an **ɔdɪt**, or even **sɪŋ** in **pʌblɪk**, before going there. The **orɪjɪnəl** **ɪvənt** that had **kæzɪd** his **fɪr** occurred **ʃɪr** before, even though it felt like a reoccurring **nəɪtmər** to him.

It all **hæpənd** one **træŋkwɪl** day when he and his **brʌðə** were out **fɪʃɪŋ** and **enʃɔɪɪŋ** the **redɪənt** **sʌŋʃəɪn**. The day was going very well until Bob was suddenly **strikiŋ** with **nəʒɪə** and felt **dɪzi** and **wɪk** in the **nɪz**. Before he **nu** it, Bob **fɛl** out of the boat and into the **kɒld** water, taking his **ɛkspensɪv** **ɡɪtɑ** with him. Luckily, his **brev** **pʌpɪ** jumped in and helped him swim to **ʃɔr**.

Worried that Bob had a **sɛvɪr ɪlnɛs**, he was rushed to the St. James Medical Center by **æmbjulɛns**. At the centre, Bob was placed in a room with another man **rikupɛrɛtɪŋ** from **sɛrʃəi**. The man was quite **vɒlgə** and too **čæti** for Bob's liking. Even though Bob was in no **mud** to **tɔk**, this man **told** him about his days in the army and how he had **sɛrvɔɪvd** an **ɛθkwɛk**. In the **mintajm**, **numəʊs blɒd tɛsts** were taken by inexperienced **nɜrsɜz** and **ænələɔɪzd** in the **læbrətɔri**. To make **mætəz** even worse, Bob's room did not have a **tɛlɛvɪʒən** and **dɪnə kɑnsɪstɛd** of **ovəkʊkt spagɛti** with **tɛstlɛs sɑs**.

ɛkspɛktɪŋ the worst, Bob's **spəʊs** believed it was a **mɪrəl** that he was just **fɔɪn** after all. The **daktəz** believed Bob had simply **ɪtɛn** a **rɑtɪn æpəl** and **rɛkɒmɛndɛd** that he take some medication for the **nɛkst** few days just in **kɛs** his **sɪmptəmz rɪtərnd**. With the **kɑst** of these **pɪlz** and the **fækt** that he **mɪst** his favourite **bænd** in **kɑnsə-t** that night, Bob **wɒntɛd** a **kælkjuletə** to **dɪtɜrmin** the **æmɒwnt** of **mʌni** he had lost on that very bad day.

Experiment 1: Story B (Sarah) Large Stimulus set size

Sarah had a **drim** of **trævliŋ** the **wærlð**. **biŋ** an **ævið** explorer, she planned the **gretest**, most **iksajtŋ** **vekešlŋ** ever **fallowiŋ** her graduation from **juniværsiti**. She was leaving behind the **siti** **træfɪk** as well as the **mæsti** **lajbreri**, which she had become so **æklæstlæmd** to. Though her **pærents** had **gaten** a **divors** **mænθs** before, they were both enthusiastic about the trip and even **ped** for the **erplen** **tiket**.

With graduation now over, **skulwærk** was a thing of the **pæst**. Sarah was off to explore **rɪmot** **vɪleʃɪz** in underdeveloped **kæntɹɪz**. Places where **ɪndæstriælajzd** **mæšɪnz**—such as an **ɛskaletə**—were **nænɛgzɪstənt**. Sarah **pækt** very **lɪtəl**; she **bɹət** **kæmufləʃ** garments, a Swiss Army knife, and a hat. In her rush to leave, she **fɔrgat** to take a **trænzlešlŋ** **dɪkʃlənɛri**.

Nevertheless, Sarah had **æmeziŋ** experiences during her travels. She went deep water **skubʌ dajvɪŋ** in the **oʃɪn** where she saw a great **wajt** shark and a **hjuʃ** red **ʃɛlifɪʃ**. She **kaŋkə-d** her fear of **hajtɪs** by jumping off a 30-foot waterfall in Zambia. In Malawi, she even **tʊk** a **tʊr** of a **hantɛd prɪzɪn** and **spɛnt** the night there in a bed **mɛd** to **lʊk** like a **kafɪn**. Sarah met some great people during her trip. For **ɪnstɪns**, she met an **ɪntɪrɛstɪŋ** and **fʌni mʌʃɪʃɪn** who **kʊd** make a **kɪtɪn dɪsʌpɪr** into **θɪn** air and make a **flæwə blum** in **frʌnt** of your **ɑjz**. She also became **bɛst kʌmpænjʌnz** with a **gɜrl nɛmɔd** Viola and was a **brɔjdzmɛd** at her **wɛdɪŋ**.

Sarah's travels were also very **sætɪsfajɪŋ** because she was **ɛbəl** to **ɡɪv** back to people in **nɪd**. In Portugal, she helped **vɪktɪmz** of a **ʃɪprɛk** find **fʊd** and places to **steɪ**. And in Tunisia, she **ʌsɪstɛd** families **rɪkʌvəɪŋ** from **bæŋkrʌptsi** by helping them find **mɪnɪŋfʊl ɛmplɔjmənt**. Sarah **dɪskʌvə-d** many things during her travels, and she is currently planning another trip.

Appendix H

Experiment 2: Target Words

Word Set C (Bob)			Word set D (Sarah)		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
advised	verb	11.02	able	adjective	281.45
afraid	adjective	79.53	aided	verb	3.99
amount	noun	107.45	airplanes	noun	4.99
analyzed	verb	13.65	amazing	adjective	70.58
army	noun	91.54	appealing	adjective	11.41
assigned	verb	22.76	avid	adjective	4.57
attack	noun	104.92	baseball	noun	52.65
audit	noun	8.14	became	verb	159.74
backpack	noun	6.37	best	adjective	453.69
band	noun	63.84	bloom	noun	9.44
bland	adjective	0.08	brought	verb	152.84
blood	noun	2.28	cap	noun	26.37
bomb	noun	0.43	city	noun	377.99
brother	noun	119.38	close	adjective	217.13
car	noun	246.17	coffin	noun	6.09
case	noun	363.13	conquered	verb	3.70
caused	verb	62.97	crisis	noun	67.74
cold	adjective	109.51	decent	adjective	22.88
concert	noun	22.50	deep	adjective	113.77
cost	noun	145.10	disappear	verb	16.31
countless	adjective	12.34	diving	verb	6.67
determine	verb	57.19	divorce	noun	25.99
disease	noun	89.40	dream	noun	78.92
doctor	noun	107.34	English	adjective	93.28
dog	noun	99.53	explorer	noun	6.73
eaten	verb	13.23	eyes	noun	253.53
enjoying	verb	17.36	find	verb	517.47
equipped	verb	10.78	flight	noun	52.21
evening	noun	78.58	following	verb	154.70
event	noun	96.92	food	noun	231.86
fact	noun	351.89	forgot	verb	25.86
faint	adjective	11.45	front	adjective	203.85
favorite	adjective	85.72	funny	adjective	76.63
fear	noun	104.20	garment	noun	3.81
fell	verb	76.24	girl	noun	199.99

Word Set C (Bob)			Word set D (Sarah)		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
fishing	verb	30.93	give	verb	465.71
happened	verb	229.37	gotten	verb	58.31
health	noun	302.68	heights	noun	13.02
horrible	adjective	27.68	homework	noun	14.37
initial	adjective	46.14	huge	adjective	110.71
knees	noun	27.16	invited	verb	31.62
knew	verb	279.97	jobs	noun	119.31
lab	noun	32.03	knife	noun	29.47
mango	noun	3.01	leaving	verb	90.74
meals	noun	18.88	library	noun	59.30
meantime	adverb	14.43	little	adjective	759.13
miracle	noun	16.30	look	adjective	781.39
missed	verb	53.43	machines	noun	27.86
money	noun	440.16	magician	noun	3.64
mood	noun	27.57	modern	adjective	84.58
nausea	noun	3.66	months	noun	233.89
next	adjective	492.19	musty	adjective	1.21
nightmare	noun	15.12	named	verb	84.68
noisy	adjective	5.51	need	noun	762.40
nurses	noun	17.18	ocean	noun	41.68
peaceful	adjective	17.09	organized	verb	34.17
pills	noun	13.03	packed	verb	22.02
prepared	verb	57.28	paid	verb	105.81
pricey	adjective	2.71	parents	noun	216.08
public	noun	376.70	past	noun	295.86
returned	verb	66.09	people	noun	1795.82
rotten	adjective	5.77	pigeon	noun	2.69
rude	adjective	11.84	planning	verb	68.84
sailboat	noun	1.19	prison	noun	65.04
sauce	noun	26.81	regions	noun	25.55
severe	verb	33.34	remote	adjective	23.72
shared	verb	52.77	rewarding	adjective	6.15
shore	noun	21.27	rose	noun	70.20
sing	verb	33.78	rush	noun	35.69
soggy	adjective	1.79	shaped	verb	15.02
spaghetti	noun	3.98	shark	noun	9.02
stricken	verb	2.41	shipwreck	noun	0.84
sunshine	noun	8.74	spent	verb	129.42

Word Set C (Bob)			Word set D (Sarah)		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
surgery	noun	41.77	stay	verb	205.10
talk	verb	363.31	ticket	noun	31.43
taxes	noun	70.69	tour	noun	55.56
tests	noun	52.99	towns	noun	19.31
things	noun	635.36	traffic	noun	51.53
TV	noun	116.57	turtle	noun	7.14
waking	verb	10.02	used	adjective	490.49
wanted	verb	324.77	vacation	noun	25.01
warm	adjective	59.80	victims	noun	50.39
weak	adjective	36.68	waterfall	noun	2.73
wife	noun	183.59	wedding	noun	44.39
would	verb	2366.03	witty	adjective	3.62
years	noun	1039.65	world	noun	720.33

Note. Target words were presented in phonetic symbols (Moats, 2003) in and out of context.
^atarget words part of the small (48) set size condition.

Appendix I

Experiment 2: Story C (Bob) Large Stimulus Set Size

Ever since Bob had a misfortunate day on his **selbot**, he became **afred** of going to the **helθ** care center. He **wud** rather do his **tæksɛz** then go through an **ɔdit**, or even **sɪŋ** in **pʌblɪk**, before revisiting the center. The **ɪnɪʃəl ɪvɛnt** that had **kazd** his **fɪr** occurred many **jɪrɪz** before, even though it felt like a reoccurring **nəjtmə** to him.

It all **hæpɛnd** one **pɪsfəl** day when he and his **brʌθə** were out **fɪʃɪŋ** and **ɛnʃɔɪŋ** the **wɔrm sʌnʃəɪn**. The day was going very well until Bob was suddenly **striken** with **nəziə** and felt **fɛnt** and **wɪk** at the **nɪz**. Before he **nju** it, Bob **fɛl** out the ship and into the **kold** water, taking his **prəɪsɪ bækpæk** with him. Fortunately, his brave **dɔg** jumped in and helped him swim to **ʃɔr**.

Worried that Bob had a **sɛvɪr dɪzɪs**, he was rushed to the medical center by **kar**. Later, he was **asəɪnd** to a room with another man **wɛkɪŋ** up from **sə-ʃəɪ**. The man was quite **rʊd** and too **nɔɪzi** for Bob's liking. Even though Bob was in no **mud** to **tak**, the man **ʃə-d** stories from his days in the **armi** and how he had survived an **ətæk** from a cluster **bam**.

In the **mintajm**, Bob had **kantlis blad tests** taken by inexperienced **nəsɛz** and that took forever to be **ænælojzd** in the **læb**. What made **θɪŋz** worse, Bob's room was not **ɪkwɪpt** with **tivi** and his **mils** was **sagi spageti** with **blænd sas**.

priɪpəd for the worst, Bob's **wɔɪf** believed it was a **mɪrɪkəl** that he was fine after all. The **daktə** concluded that Bob had **itən** a **ratɪn mæŋgo** and **advɔɪzd** that he takes medication for the **nɛkst** few days just in **kes** the symptoms **rɪtənd**. With the **kast** of these **pɪlz** and the **fækt** that he **mɪst** his **fevərɪt bænd** in **kansət** that **ɪvniŋ**, Bob **wantəd** a calculator to **dɪtəmiːn** the **æmæwnt** of **mʌni** he had wasted on that ridiculously, **hɔrəbəl** day (End).

Experiment 2: Story D (Sarah) Large Stimulus Set Size

Sarah had a **drim** of travelling around the **wæld**. Being an **ævid** **iksplorə**, she had **orgənəjzd** herself the **bɛst** and most **əpɪlɪŋ** **vekeʃɪn** ever **fəlowɪŋ** graduation from university. She was **lɪvɪŋ** behind the **sɪti træfɪk** as well as the **mʌsti lajbrəri**, which she had become so **juzd** to. Though her **pærɛnts** had **ɡatən** a **dɪvɔrs** few **mʌnθs** before, they were both enthusiastic about her trip and even **ped** for the **floʊjt tɪkɛt**.

With graduation now accomplished, **homwɜk** was a thing of the ancient **pæst**. Sarah was off to explore some **rɪmot tæwns** in underdeveloped **rɪʒɪns**. Places where **mʌdən mʌʃɪnz**—such as an **əplən**—were nonexistent. Sarah **pækt** very **lɪtəl**; she **brɔt** camouflage **ɡarmɛnt**, a Swiss army **nɑjf**, and **bɛsbɔl kæp**. In her **rʌʃ** to leave, she **fɔrgat** to take her **ɪŋɡlɪʃ** translation dictionary.

Nevertheless, Sarah had **əmezɪŋ** experiences during her travels. She went **dɪp** water *scuba* **dəjvɪŋ** in the **oʃɪn** where she saw a great white **ʃɑrk** and a **hjuː** leatherback **tɔtəl**. She **kʌŋkəd** her fear of **haɪts** by jumping off a 30-foot **wɔtə-fɔl** in Zambia. In Malawi, she even took a **tʊr** of a haunted **prɪzɪn** and **spɛnt** the night there in a bed **ʃɛpt** to **lʊk** like a **kafɪn**.

Sarah met some lovely **pipəl** during her trip. For instance, she met a **witi** and **fani maʃišin** who could make a **piʃin disəpir** into thin air and make a **roz blum** in **frant** of your **ajz**. She also **bikem kloz** companions with a **gəl nemd** Viola and was **invajtɪd** to be the bridesmaid at her **wɛdɪŋ**.

Sarah's travels were also very **riwardɪŋ** because she was **ebəl** to **gɪv** back to people in **nɪd**. In Portugal, she helped **viktimz** of a **šiprɛk fajnd fud** and apartments to **ste**. And in Budapest, she **edɪd** families recovering from financial **kraʃɪs** by helping them find **disənt ʃabs**. Sarah discovered many things during her travels, and she is now **plænɪŋ** an expedition in east Asia (End).

Appendix J

Experiment 3: Large Set Size Target Words

Word Set E (Bob)			Word Set F (Sarah)		
<i>/i/ eagle; /æ/ apple; /u/ ooze; /č/ chat; /z/ zoo</i>			<i>/ɛ/ echo; /e/ ape; /ʌ/ up; /š/ she; /j/ juice</i>		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
analyzed	verb	13.65	able	adjective	281.45
angry	adjective	52.23	agent	noun	56.59
apple	noun	61.53	ahead	adjective	123.08
bad	adjective	288.85	airplane	noun	11.12
band	noun	63.84	assisted	verb	7.11
blues	noun	16.40	availed	verb	0.24
cartoons	noun	4.94	bake	verb	12.88
catching	verb	13.75	best	adjective	453.69
caused	verb	62.97	bread	noun	33.46
chapter	noun	37.62	bridge	noun	44.63
chart	noun	19.79	budget	noun	81.68
chatty	adjective	1.02	chef	noun	20.27
cheery	adjective	1.42	crush	verb	10.65
cheese	noun	36.24	danger	noun	41.31
chewy	adjective	1.02	dialect	noun	2.22
chief	adjective	108.05	divulge	verb	1.18
children	noun	445.51	fresh	adjective	76.13
chomped	verb	0.19	fun	noun	129.70
choose	verb	79.54	garbage	noun	17.71
crack	noun	23.73	getaway	noun	2.87
crappy	adjective	3.68	giant	adjective	42.09
dizzy	adjective	4.45	greatest	adjective	55.65
doctors	noun	57.24	indulged	verb	1.27
fact	noun	351.89	journal	noun	68.08
flat	adjective	40.18	jumping	verb	14.56
freezing	adjective	9.26	jungle	noun	11.19
gloomy	adjective	2.70	knowledge	noun	115.97
goose	noun	5.84	large	adjective	229.11
handy	adjective	8.63	lush	adjective	5.10
itching	verb	2.18	made	verb	711.93
knees	noun	27.16	magician	noun	3.64
knew	verb	279.97	major	adjective	198.17
launched	verb	27.72	met	verb	129.05
monsoons	noun	0.24	named	verb	84.68

Word Set E (Bob)			Word Set F (Sarah)		
<i>/i/ eagle; /æ/ apple; /u/ ooze; /č/ chat; /z/ zoo</i>			<i>/ε/ echo; /e/ ape; /ʌ/ up; /š/ she; /j/ juice</i>		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
mood	noun	27.57	nutshell	noun	1.89
moons	noun	3.46	ocean	noun	41.68
moved	verb	120.89	page	noun	106.56
nausea	noun	3.66	paid	verb	105.81
pal	noun	10.92	porridge	noun	0.84
pills	noun	13.03	reddish	adjective	1.72
pooch	noun	0.90	rush	verb	35.69
room	noun	325.29	shark	noun	9.02
rude	adjective	11.84	shelter	noun	20.66
sandwich	noun	13.92	shepherd	noun	7.65
snooze	verb	0.92	shipwreck	noun	0.84
spoon	noun	10.74	shirts	noun	9.70
statue	noun	11.21	shorts	noun	10.20
stitches	noun	2.70	shut	verb	81.91
stupid	adjective	67.15	take	verb	870.07
tally	verb	3.43	ticket	noun	31.43
taxes	noun	70.69	up	adjective	2348.92
times	noun	356.59	vacation	noun	25.01
torches	noun	1.49	wage	noun	18.84
tv	noun	116.13	wait	verb	212.72
weak	adjective	36.68	wished	verb	15.62

Note. Target words were presented in phonetic symbols (Moats, 2003).

Experiment 3: Small Set Size Target Words

Word Set E (Bob)			Word Set F (Sarah)		
<i>/ε/ echo; /e/ ape; /ʌ/ up; /š/ she; /j/ juice</i>			<i>/i/ eagle; /æ/ apple; /u/ ooze; /č/ chat; /z/ zoo</i>		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
angel	noun	23.79	bloom	noun	9.44
blood	noun	137.80	busy	adjective	50.49
brave	adjective	17.89	charming	adjective	11.66
cottage	noun	7.66	cheap	adjective	34.02
favourite	noun	4.63	dream	noun	78.92
fell	verb	76.24	duo	noun	5.22
fish	noun	83.01	family	noun	443.66
jello	noun	0.41	food	noun	231.86
juggle	verb	1.32	fruit	noun	33.55

Word Set E (Bob)			Word Set F (Sarah)		
<i>/ɛ/ echo; /e/ ape; /ʌ/ up; /ʃ/ she; /j/ juice</i>			<i>/i/ eagle; /æ/ apple; /u/ ooze; /č/ chat; /z/ zoo</i>		
Target word	Part of speech	Frequency	Target word	Part of speech	Frequency
lake	noun	63.24	goofy	adjective	3.48
loving	verb	23.14	grab	verb	30.21
next	adjective	492.19	group	noun	321.44
option	noun	52.24	hitchhike	verb	0.30
orange	adjective	38.18	matching	verb	12.32
placed	verb	53.92	past	noun	295.86
precious	adjective	19.14	pleased	adjective	23.18
sergeant	noun	16.26	rose	noun	70.20
severe	adjective	33.34	satchel	noun	1.40
shocked	verb	20.22	topaz	adjective	0.58
shore	noun	21.27	victims	noun	50.39
space	noun	164.89	visit	verb	91.14
sunshine	noun	8.74	watched	verb	66.21

Note. Target words were presented in phonetic symbols (Moats, 2003).

Appendix K

Experiment 4: Story E (Bob)

Ever since Bob had an accident on the **lek** at his **katrj**, he was afraid of going to the hospital. The original event that **kazd** his fear occurred many **munz** ago, even though it felt like a reoccurring nightmare to him. In fact, he would rather do his **tæksiz** a million **tqimz**, be charged by an **ængri gus**, or even **jʌgəl** flaming **tortčiz**, before going to the hospital again.

It all happened one tranquil day when he, his **pæł**, and his **čildrin** were out **kæčɪŋ frš** and just **lʌvɪŋ** the **čiri**, radiant **sʌnšqɪn**. The day was going very well until Bob was suddenly stricken with **nazia** and felt **dizi** and **wik** in the **niz**. Before he **nu** it, Bob banged his head, **fɛl** out of the boat into the **frizɪŋ** water, taking his **prešʌs bluz** guitar with him. Luckily, his **brev puč lančt** herself into the water and helped him swim to **šor**.

Worried that Bob had a **sɛvɪr** illness, he was sent to the St. James Medical Centre by ambulance. At the centre, Bob could not **čuz** his bed and was **plest nɛkst** to another man recuperating from surgery. This fellow was quite **rud** and too **čæti** for Bob's liking. Even though Bob was in no **mud** to talk and just wanted to **snuz**, the man told him about his days as a **sarjɛnt** and how he had weathered **mansunz** in Thailand. Bob asked to be **muvd** to another **spes** but it was of no use.

Meanwhile, **blad** tests were taken by inexperienced nurses and **ænəlɔjzɪd** in the laboratory. To make matters worse, Bob's **rum** did not have a **tivi** so he could not watch **kartunz**. Instead, he looked out his window and staired at a **stæču** of an **enʃəl** in the courtyard. The snacks were also terrible. They served him a soggy **čiz sænwič**, a **flæt** cola, and **čui, ɔɪnʃ ʃɛlo**. When Bob thought things could not get any worse, he heard a **kræk** as he **čampt** down on the **spun**. All this made him feel quite **glumi** and **ičɪŋ** to get out of the hospital.

The **čif daktǽrz** looked at Bob's **čǽrt** and thought he had simply eaten a **bæd æpəl**. He had the **apšʌn** to take some medication just in case the symptoms returned—at least he did not need **stičǽz** for his head. Bob's spouse was **šakt** and believed it was a miracle that he was fine after all. But with the cost of the **pǽlz** and the **fækt** that he missed his **fevrǽt bænd** in concert that night, Bob wanted his **hændi** calculator to **tæli** the sum of money he lost on that very **kræpi, stupid** day. And that closes the **čæptǽr** on why Bob fears hospitals.

Experiment 4: Story F (Sarah)

Sarah had a **drim** of travelling the world. Being an avid explorer, she planned the **gretɛst**, most exciting **vekešin** ever following her graduation from university. She would soon forget the **bizi** city, the **mejɔ** traffic jams, and the musty library. Her parents were both enthusiastic about the trip. They helped Sarah plan her trip and even **ped** for the travel **eĵint** and **ɛrplen tɪkɛt**.

With graduation now over, homework was a thing of the **pæst**. Sarah could not **wet** to **tek** off to explore villages in remote regions. Sarah planned to **hičhɔjk** so she packed very little. She brought a Swiss Army knife, camouflage **šɔrts** and **mæčɪŋ šɔrts**, and a **ĵɔɛnt** hat. Although, she did take some maple syrup with her, which took extra space. In her excitement and **rʌš** to leave, she forgot to **græb** her translation dictionary. So, she had to buy a **čip** one on the way and **krʌš** it in her bag.

Despite having a small **baɪt**, and often eating **porɪʃ** as a main meal, Sarah had an amazing time during her travels. She went diving in the **oʃɪn** where she saw a **reɪʃ ʃɑrk** and a **grʊp** of turtles. She conquered her fear of heights by **ʃʌmpɪŋ** off a 30-foot **briʃ** in a **lʌʃ** Zambian **ʃʌŋgəl**. In Malawi, she even took a tour of a haunted prison and spent the night in a bed **med** to look like a coffin that **ʃʌt** with a lid. In Kenya, a **ʃɛf** taught her how to **bek bɛd**. Luckily, many people spoke a **dʒʌlɛkt** of English. It was all amazing and every night she **ɪndʌɪʃd** by adding a **peɪ** in her **ʃəneɪl**.

Sarah **mɛt** some marvelous people during her **vɪzɪt** across the globe. For instance, she talked with a **fʌn** and interesting **mʌɪʃɪn** hoping he would **dʒɪvʌɪʃ** his **nəɪʃ** about magic. He could make a **lɑɪ** bowl with **frɛʃ frʊt** disappear into thin air and make a **roz blum** in front of your eyes. She also met a **ʧɑrmiŋ ʃɛpərd** who gave her a **sæʧəl** beaded with **topæz**. She even became **bɛst** companions with a **gʊfi** girl **nɛmd** Viola. Sarah was a bridesmaid at her wedding and **wɑʧt** Viola walk down the aisle.

Sarah's travels were also satisfying because she was **ebəl** to help those in need. In Portugal, she **Λveld viktɪmz** of a **šɪprɛk** by finding them **fud** and **šɛltɔ̃**. Sarah also joined a **duo** in Morocco to pick **Λp gɔ̃bɪj** along a riverbank but **wɪʃt** she could do more. Her trip ended in Tunisia where she **Λsɪstɛd** women of a **fæmli** in **denjɔ̃** of bankruptcy find employment with a good **wej**. In a **nΛtšɛl**, Sarah discovered many things during her **gɛtΛwe**, and was very **plɪzd** with how it turned out. Now, she is looking **Λhɛd** to her next trip.

Appendix L

Experiment 5: Irregular, Regular Long, and Regular Short Target Words

Word set G (Animal Diary)			Word set H (Cold Night)		
Target word	Type	Frequency	Target word	Type	Frequency
Aching	Irregular	2.81	Choir	Irregular	7.80
Comb	Irregular	3.43	Deaf	Irregular	8.28
Cough	Irregular	6.51	Dreamt	Irregular	1.92
Flood	Irregular	18.05	Flown	Irregular	6.12
Gauge	Irregular	6.93	Ghoul	Irregular	0.47
Hearth	Irregular	2.11	Hearse	Irregular	0.77
Leapt	Irregular	2.82	Mauve	Irregular	0.57
Pour	Irregular	15.50	Pear	Irregular	2.92
Scarce	Irregular	5.81	Sword	Irregular	16.13
Shone	Irregular	3.72	Ton	Irregular	10.89
Acrobat	Regular Long	0.65	Chortle	Regular Long	0.14
Comfort	Regular Long	29.64	Decide	Regular Long	62.51
Compass	Regular Long	4.82	Flinch	Regular Long	1.45
Crocodiles	Regular Long	0.94	Floated	Regular Long	5.68
Flaunt	Regular Long	0.81	Gossip	Regular Long	8.00
Gargle	Regular Long	0.22	Grizzly	Regular Long	2.58
Hundred	Regular Long	73.37	Hermit	Regular Long	1.23
Lesson	Regular Long	35.30	Manicure	Regular Long	1.25
Poacher	Regular Long	0.28	Pencil	Regular Long	8.89
Scrape	Regular Long	4.01	Tonsil	Regular Long	0.20
Crabs	Regular Short	2.99	Drank	Regular Short	12.19
Cot	Regular Short	2.02	Chat	Regular Short	13.88
Cove	Regular Short	3.26	Dew	Regular Short	2.34
Flee	Regular Short	6.16	Drab	Regular Short	1.58
Gala	Regular Short	3.43	Gown	Regular Short	6.76
Heap	Regular Short	4.93	Grin	Regular Short	9.38
Leak	Regular Short	8.83	Hem	Regular Short	2.33
Pond	Regular Short	11.03	Pests	Regular Short	2.65
Scabs	Regular Short	0.54	Swig	Regular Short	1.00
Shed	Regular Short	16.67	Tone	Regular Short	34.11

Appendix M

Experiment 5: Story G (Animal Diary)

February 15: I read that snakes shed their skin. A snake shed one near the cove. The skin was a big heap on the ground. Its scales shone in the sun. I am going to use my compass to go back to the same spot at the cove and search the area for another one. Snakeskins are scarce. It's hard to get them before they are blown away by the Arctic wind.

February 16: After what I found yesterday, I probably have a hundred live crabs now. I also started collecting crow feathers yesterday. I put them in a heap near the crabs.

February 17: Sam says I talk too much about my animals. It's hard to gauge if he is right. I try not to flaunt my cool things. He's one to talk! His aunt hates it when he brags. Once, his aunt made him gargle with soap to teach him a lesson because he would not stop bragging.

February 28: Did you know when a person hunts crocodiles illegally he or she is called a poacher. If we had crocodiles in the Arctic I would protect them. I would give everyone a lesson about how wrong it is to hunt illegally. I would make the punishment for hunting so terrible, every poacher would leave. They would flee when they saw me coming.

March 14: When I grow I am going to throw a gala to raise money for animals that are scarce. The gala is going to be named after me, but I will try not to flaunt it.

March 18: I saw a nest knocked down during a flood at the pond a few days ago. I leapt to catch it but I missed. Even an acrobat would have had a hard time catching it, and I'm no acrobat! You should see the scrape on my knee! I have lots of scabs. Let me start at the beginning...

The baby crow was not hurt from the fall because the nest landed in part of the pond that wasn't frozen. The nest was like a boat. It was hard to gauge what to do. But then the nest started to leak. The leak got bigger and bigger and I saw the bird shiver in the freezing water. I decided to take the nest home without touching the baby bird. The water was rising so I knew I had to flee. I used my compass to find my way.

When I got close, the glow from the fire shone brightly through the window promising the comfort of home. I set the nest on the hearth by the fire. I started to shiver after the flood, too. Then I started to cough and my scrape started aching.

My mom set up a cot for me by the hearth, too. She started to comb my hair, but then noticed I had a temperature. She put down the comb. She said my temperature was one hundred degrees. From the comfort of my cot, I watched her pour some medicine for me. Then I watched her pour some water for the baby bird. The flames leapt in the fireplace. The aching stopped. My mom still made me gargle with salt because of my cough (but at least it wasn't soap!). What a great day. The big scabs on my knees make me look cool.

Experiment 5: Story H (Cold Night)

It was a drab, chilly night. Mist floated off the swamp. In the morning, the sparkly dew looked beautiful, but in the moonlight the dew set a creepy tone. There was a gross, moldy stench in the air from the swamp. Someone had searched for treasure, but all they had found was a ton of pests.

A grizzly old man took a swig of coffee. He was a hermit. The coffee made him flinch as he drank because he had sore throat. His right tonsil was fine, but his left tonsil had hurt for a week. He took another swig of coffee. He needed to decide whether or not see a doctor. He grabbed a pencil to make a note about it. The pencil broke. Maybe it was a sign that he should not decide right now, he thought with a grin. After that, he drank his coffee more slowly. He did not flinch again.

He looked out the window into the drab evening. Suddenly he saw a hearse drive towards him. The other cars had flown by, but the hearse floated as slowly as a ghoul. He turned away from the window. At first, the man thought that he had dreamt it, but on a night like this anything was possible.

The man was partly deaf but he still knew people were coming. Normally, the hermit didn't like visitors. Tonight was an exception.

Waiting was hard. The old man started to read a gossip magazine and eat a pear to pass the time. He dozed off and dreamt of a beautiful choir that sang with a perfect tone. He awoke to find the gossip magazine and pear were real but the choir was not.

Finally, a few boys and girls arrived at the old man's home. One girl had a mauve gown and a matching mauve manicure. She got the hem of her dress caught in the door. The hem was slightly torn but the rest of the dress was fine. Then two boys came. The first had a grizzly bear mask, the second had sword. The girl with the gown showed off her manicure as she held out her bag. The other boy put down his sword to do the same. With a chortle and a sweep of his arm, the old man gave them a ton of sweet candy. They had a short chat and then said goodbye.

The doorbell rang again. This time it was a little girl with a ghoul mask. He gave her a friendly grin as he handed out more candy with a sweep of his hands. She started to chat. The man pointed to his hearing aid to show he was partially deaf. The child gave a chortle and then said "thank you" very loudly.

Soon the evening was over. It had flown by. The moldy stench no longer seemed gross. The pests were asleep. It had been a perfect Halloween.

Appendix N

Experiment 5: Preexposure Phase Vocabulary Teaching

Word Set G and H	
Target	Phrase
Aching	Sore; his back is aching/sore; you can also have an aching head or tooth or other body parts
Acrobat	A person who does tricks up high at the circus
Chat	Talk with your friends, have a short conversation
Choir	A group of singers, like at a church or some schools
Chortle	A big laugh
Comb	A tool for your hair; it takes away tangles
Comfort	make someone feel better when they're sad or hurt
Compass	a tool that helps you find your way, especially when you're in the woods
Cot	A small bed you can fold up to put away, like at a hotel or camp
Cove	A curve in the land by the water
Crocodiles	A big, green animal with sharp teeth
Deaf	People who have a hard time hearing or can't hear at all.
Decide	Make a choice about which way to go or what you want
Dew	Little drops of water, like on leaves and grass in the morning
Drab	Dark, boring colours, not bright
Drank	Drank is the past of drink
Dreamt	He is dreaming, but last night he dreamt, the past of dream
Flaunt	Show off; he likes to flaunt his big muscles
Flee	Caption Sparrow is scared and running away from the cows; he wants to flee the cows
Flinch	Make a quick, nervous movement; reaction to surprise, fear or pain. Someone is throwing sticks at him, and he knows it, so he flinches before getting hit.
Flood	When there's so much water it overflows (gets high) and spills into the streets and onto land. It can happen after a big storm or when snow melts.
Flown	It means to fly in the past. There birds have flown south many times.
Gala	A fancy party
Gauge	When you need to figure out something, like a reaction to something
Ghoul	A spooky creature
Gossip	The girls in the back are talking about the girl in pink. She feels bad because they are saying things about her that could be true or not true.
Gown	A fancy dress
Grin	A big smile
Grizzly	A type of bear
Hearse	A long car with curtains in the windows, it has a funny shape in the back; sometimes seen in spooky movies

Word Set G and H	
Target	Phrase
Hearth	It goes around the fireplace to protect the floor
Hem	The bottom of a shirt, skirt, or pants where it is sewn
Hermit	Some people are called hermits because When you need to figure out something, like a reaction to something they like to live alone and they do not like to have many visitors; they might live in a shack in the woods or mountains, far away from people
Hundred	The last number on the number chart
Leak	A hole in a pipe where water comes out
Leapt	The past of leap; a big jump
Manicure	Putting on nail polish and shaping nails to look pretty
Mauve	A light, pink-purple colour
Pests	Animals and insects that you do not want in your house
Poacher	A person who hunts animals even though he is not allowed to do it. A poacher sells the animals to make money. If he gets caught, he will be in big trouble.
Scabs	When you scrape your skin
Scarce	Not enough of something; water is scarce in the desert
Scrape	It is like using a knife or tool to take off or scrape old paint
Shed	When old skin peels off, like lizards and snakes; dogs and cats also shed fur
Shone	Is the past of shine, the stars shone last night
Swig	Means to take a big drink of something, like when you are very thirsty
Ton	Something that is very heavy can weight a ton. A ton means a lot of something.
Tone	This boss looks angry, his tone of voice is loud and angry
Tonsil	In the back of your throat you have tonsils, you have one on each side, they can get red and sore

Appendix O

Experiment 5: Spelling Test

Word Set G and H	
Target	Phrase
Pond	Next to my cottage there is a pond where my frogs live.
Ghoul	The ghoul in that scary movie was creepy.
Can	I can do it!
Lesson	I am going to my violin lesson after school today.
Grizzly	The grizzly bear came out of his den.
Acrobat	The brave acrobat walked the tightrope.
Flinch	The loud noise made me flinch.
Crabs	I saw some crabs on the beach.
Pencil	I need to sharpen my pencil.
Leak	After a storm, the old pipes usually leak.
Mauve	Her favourite colour dress is mauve.
Flaunt	He drives fast to flaunt his fancy car.
Run	I run very fast.
Flown	I have never flown in a hot air balloon.
Shone	The stars shone brightly last night.
Pests	He used to spray to kill the pests in the kitchen.
Comfort	When I was sad, my mom would always comfort me.
Gown	Cinderella wore a beautiful gown at the ball.
Poacher	The poacher was caught and arrested.
Tone	Her voice had a clear tone.
Pour	Pour your milk carefully.
Gossip	They liked to gossip about the other kids.
Cough	I have a cough and a sore throat.
His	His name is John.
Chat	Let's chat about our party plans.
Scarce	In the desert, water is scarce.
Swig	After the race, I took a swig of water.
Flee	The villagers were forced to flee from the invaders.
Gargle	I gargle with warm water and salt.
Drank	He drank a whole carton of juice.
Gauge	It was hard to gauge their reaction to the prime minister's speech.
Decide	She will decide where to have her party.
Scabs	I got these scabs from falling off my bike.
Dreamt	He dreamt about flying.
Scrape	After lunch, scrape your dish.

Word Set G and H	
Target	Phrase
Sword	The pirate held a shiny sword.
Deaf	He uses sign language because he is deaf.
Cove	The beach at the cove is beautiful.
Chortle	She gave a big chortle at the joke.
Leapt	I leapt across the puddles.
Drab	The uniform was a drab green.
Shed	The lizard shed its skin.
Hem	I sewed the hem on your dress.
Hearth	Let's warm up by the hearth.
Manicure	Her manicure made her hands look pretty.
Crocodiles	The river is home to many crocodiles.
Hearse	The hearse drive by slowly.
Flood	The heavy rains caused a flood.
Tonsil	His left tonsil was red.
Hundred	We had a hundred day party at school.
Floated	The ghost floated through the air.
Compass	She used her compass to find the way.
Grin	The Cheshire Cat has a big grin.
Aching	My feet were aching after walking for hours.
Hermit	The hermit lived in a small shack in the woods.
Help	I need some help with my zipper.
Cot	At camp, I slept on a cot.
Dew	The dew on the grass glistened.
Heap	I left my dirty clothes in a heap on the floor.
Pear	He ate a juicy pear for snack.
Comb	She should comb her tangled hair.
Ton	The teacher gave us a ton of homework!
Gala	Tonight, we are going to a gala for the children's hospital.
Choir	I enjoy listening to the choir sing.

Appendix P

Experiment 6: Target Words

Word Set I (Animal Diary)		Word set J (Cold Night)	
Target word	Frequency	Target word	Frequency
acrobat	.65	awful	28.18
adore	3.27	babysitter	2.33
agree	126.43	bag	63.24
animals	67.25	balls	22.01
atlas	4.70	best	453.71
balance	56.87	better	498.91
beast	15.95	black	312.07
became	278.44	boys	104.37
bundle	5.88	candy	20.94
change	329.23	cash	60.52
chatter	4.33	children	445.46
circus	8.29	chilly	4.25
compass	4.81	chuckled	3.71
could	1541.10	comfort	29.63
crab	6.05	coming	288.17
crocodile	1.93	counted	15.53
cub	2.71	creepy	8.39
cute	27.37	crook	2.81
darling	15.15	dew	2.34
exact	29.42	different	416.50
example	241.18	digging	12.03
explorer	6.73	dirty	31.40
facts	54.70	drab	1.58
family	443.77	drifted	6.32
father	301.33	drops	18.85
film	139.20	enjoy	72.41
flames	11.17	evening	78.57
flexible	13.14	farmer	14.72
flips	2.78	filled	61.41
fox	48.86	found	445.57
giving	115.85	front	203.85
happens	98.07	gate	25.76
humans	46.58	gave	189.66
hunting	27.06	gems	2.98
including	263.62	golf	36.59

Word Set I (Animal Diary)		Word set J (Cold Night)	
Target word	Frequency	Target word	Frequency
jobs	119.31	good	1138.30
land	140.33	grinned	7.08
lessons	31.48	grizzly	2.58
lion	13.08	Halloween	12.95
made	711.97	handful	20.63
mammals	5.28	happy	159.46
mean	518.85	hid	8.15
member	102.24	jug	2.34
months	233.99	kids	246.29
need	762.47	left	423.57
nice	184.02	life	724.78
now	1665.68	little	759.26
odd	29.61	luck	51.12
online	99.15	man	749.35
other	1464.49	mist	6.19
people	1795.79	moonlight	4.84
place	478.75	morning	254.64
possible	215.35	muddy	5.62
prize	24.98	next	492.24
rabbit	10.29	night	425.21
rope	14.35	objects	39.28
scales	13.43	off	751.23
searched	10.84	old	428.74
seems	216.16	porch	17.36
sense	226.92	robbers	2.00
serpent	2.47	round	70.53
shed	16.67	rumor	7.26
shiver	2.32	shuffled	2.37
should	927.45	silent	34.23
showed	94.83	sipping	3.44
similar	130.21	smirked	1.02
sister	91.93	snickered	.53
skin	81.21	sparkled	1.06
snakes	6.89	sports	75.41
spotted	14.09	stack	11.19
stopped	88.68	started	236.08
stream	24.55	steaming	3.38
swallow	8.36	stench	2.63

Word Set I (Animal Diary)		Word set J (Cold Night)	
Target word	Frequency	Target word	Frequency
switched	11.35	swamp	5.96
teacher	107.59	tapping	46.53
tells	64.20	tea	46.53
think	1503.28	thrust	9.53
three	574.50	time	1711.88
travel	74.00	towards	182.06
types	60.48	treat	44.36
valley	48.64	trick	22.94
watch	168.10	uncover	4.13
week	328.83	visitors	29.96
win	145.45	wetlands	4.50

Appendix Q

Experiment 6: Story I (Animal Diary)

I adore facts. For example, all types of snakes shed their skin. It happens every three months. This week I spotted a bundle of scales in the valley by the stream. With my compass and atlas, I will travel to that exact place when I sense I need a change.

My sister tells our father I chatter too much about animals. How is that possible? Humans are mammals. So that is similar to saying I think too much about her. I mean, other people do not agree that a crocodile or a crab is as cute as a fox or a rabbit. But that seems odd. Is there a prize for the most darling beast? If there is, the lion cub should win over any member in my family, including her.

My teacher showed us a film about an explorer who stopped hunting. He is now giving lessons about being nice. I searched online, he switched jobs again. He became a circus acrobat. He could swallow flames, balance on a rope, and land flips. It made me shiver to see that he is as flexible as a serpent.

Experiment 6: Story J (Cold Night)

It was a drab chilly night and mist drifted off the swamp. In the morning, the dew drops sparkled, but in the moonlight they were creepy. There was an awful stench coming from the wetlands. Kids had been digging for gems, but all they found were muddy golf balls.

“Better luck next time,” smirked a grizzly old farmer from the comfort of his porch. He was sipping steaming black tea when the tapping started on his front gate. He did not enjoy visitors. Life was best when it was silent. But this evening was different.

“Trick or Treat,” said the little robbers.

The man grinned. “Happy Halloween,” he said as he gave out a handful of candy.

“Uncover anything good?” The boys thrust a dirty jug filled with round objects towards him.

“There was a rumor that a crook hid out here,” said the babysitter. He snickered.

After the children left, he shuffled to his sports bag, still chuckling to himself as he counted his stack of cash.

